WCC 2024 Proceedings

Public-Key Encryption from the Lattice Isomorphism Problem

– Extended Abstract –

Léo Ackermann[1], Adeline Roux-Langlois[1], and Alexandre Wallet[2]

[1] Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France
{leo.ackermann, adeline.roux-langlois}@cnrs.fr
[2] Inria, Univ Rennes, CNRS, IRISA, Rennes, France
alexandre.wallet@inria.fr

Abstract. We build a public-key encryption scheme relying on the Lattice Isomorphism Problem, which is the problem of deciding whether two lattices are rotations of each other. We generalize a restricted-to-$\mathbb{Z}^n$ scheme from *Benett et al.* using the quadratic form formalism. Our proposal benefits from more versatility, no floating points arithmetics, and relies on a plausibly falsifiable assumption.

Keywords: Public-key Encryption · Lattice Isomorphism Problem

Modern cryptography is challenged by the advent of quantum computers with enough quantum-bits and efficient error correction. In this still hypothetical setup both factorization and discrete logarithm problem are no longer hard as Shor's algorithm [24] can solve them in polynomial time. Lattices, or discrete subgroups of a real multidimensional space, have proven themselves as strong candidates for quantum-resistant cryptography. Besides conjectured quantum-resilient, the *average-case-worst-case* connection of lattice problems accounts for their attractivity. Decades of lattice-based cryptography gave birth to well understood hypotheses — eg. NTRU, the Learning With Errors (LWE) and the Small Integer Solution (SIS) problems — and a thick variety of constructions, ranging from public-key encryption scheme [23] and signatures [15] to fully homomorphic encryption [14] going through anonymous credentials [7,18]. Algebraically structured variants of those problems, relying on number theoretic structures, yield fast and compact schemes. Recent lines of works consider agressive variations of standard hypotheses to reach attractive performances [1,7,8].

At a high level, the current lattice-based schemes generate a public *random* lattice together with a trapdoor that forms the secret key. Typically, a random basis of the lattice is made public while a particular one, made of short and as orthogonal as possible vectors is kept secret. Breaking these schemes reduces to solving well-identified and well-studied hard problems over random lattices. One such problem is Bounded Distance Decoding (BDD): given a point very close to a lattice Λ, one is asked to return the closest lattice point. Heuristically, the concrete hardness of this problem is driven by the gap $\mathrm{gap}(\Lambda, \rho)$ between ρ,

the distance between Λ and the target, and half the Gaussian heuristic, which predicts the length of the shortest vectors in "random enough" lattices. It turns out that solvers perform much better when $\mathrm{gap}(\rho, \Lambda)$ gets large. For LWE, SIS, NTRU schemes, one expects a $O(\sqrt{n})$-gap, but other classes of lattices can reach much smaller gaps. For example, Barnes-Sloane lattices [3] have decoding gap as small as $\Theta(\sqrt{\log n})$. Consequently, these lattices give a much better concrete BDD security at a given dimension; equivalently, they require quite smaller dimension to reach a given security level, leading to efficiency improvements. This is an unfortunate aspect of the current "standard" lattice-based cryptography: hypotheses on random lattices and their subsequent constructions barely connect with the luxuriant litterature on remarkable lattices.

■ Minimal preliminaries

Vectors and matrices. Matrices are denoted by bold capital letters (eg. $\mathbf{B}$), the transpose operator by $\cdot^{\mathsf{T}}$ and the dual by $\cdot^{\vee}$. (Column) vectors are denoted by bold letters (eg $\mathbf{x}$).

Spaces. Let $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ denote respectively the set of natural numbers, integers and reals. The discretized n-dimensional hypercube is $\mathcal{H}_q^n = \{0, 1/q, \ldots, (q-1)/q\}^n$. We denote the general linear group of degree n over $\mathbb{Z}$ by $\mathrm{GL}_n(\mathbb{Z})$, the set of symmetric positive definite matrices of dimension $n \times n$ over $\mathbb{R}$ by $\mathrm{S}_n^{++}(\mathbb{R})$.

Lattices. A (full-rank) lattice Λ is an n-dimensional discrete subgroup of $\mathbb{R}^n$. As such, it admits a smallest non-zero vector of length $\lambda_1(\Lambda)$. The gaussian heuristics $gh(\Lambda)$ gives an estimate of $\lambda_1(\Lambda)$ and is accurate for random lattices[3].

Quadratic forms. Quadratic forms can be represented by real symmetric matrices. In this work we will only be interested in the positive definite case, i.e. elements of $\mathrm{S}_n^{++}(\mathbb{R})$. A quadratic form Q represented by a matrix $\mathbf{Q} \in \mathrm{S}_n^{++}(\mathbb{R})$ defines a scalar product $\langle \mathbf{x}, \mathbf{y} \rangle_Q = \mathbf{x}^{\mathsf{T}}\mathbf{Q}\mathbf{y}$, with its associated Euclidean norm $\|\mathbf{x}\|_Q^2 = \mathbf{x}^{\mathsf{T}}\mathbf{Q}\mathbf{x}$. The relation $\mathcal{R}$ that relates $\mathbf{Q}$ to $\mathbf{Q}'$ whenever there exists $\mathbf{U} \in \mathrm{GL}_n(\mathbb{Z})$ such that $\mathbf{Q}' = \mathbf{U}^{\mathsf{T}}\mathbf{Q}\mathbf{U}$ is an equivalence relation, and $[Q]$ denotes the class of Q. The smallest length of a vector of $\mathbb{Z}^n$ through the induced norm, ie. $\min_{\mathbf{x} \in \mathbb{Z}^n \setminus \{0\}} \sqrt{\mathbf{x}^{\mathsf{T}}\mathbf{Q}\mathbf{x}}$, only depends on the class and is written $\lambda_1([Q])$. The n-th minima is the infimum of the radii of 0-centered balls[4] containing i linearly independent $\mathbb{Z}^n$ vectors.

Gaussians. The gaussian function for the quadratic form Q with width parameter σ is $\rho_{Q,\sigma}(\mathbf{x}) = \exp(-\pi\|\mathbf{x}\|_Q^2/\sigma^2)$, for all $\mathbf{x} \in \mathbb{R}^n$. When $\sigma = 1$, the subscript is omitted. Elliptic Gaussians behave much like spherical ones: if $\mathbf{Q} = \mathbf{L}^{\mathsf{T}}\mathbf{L}$, then $\|\mathbf{x}\|_Q = \|\mathbf{L}\mathbf{x}\|$. The discrete Gaussian distribution $\mathcal{D}_{Q,\sigma}$ with width parameter σ is defined by the probability density function $\Pr_{\mathbf{x} \hookleftarrow \mathcal{D}_{Q,\sigma}}(\mathbf{x} = \mathbf{y}) = \frac{\rho_{Q,\sigma}(\mathbf{y})}{\rho_{Q,\sigma}(\mathbb{Z}^n)}$,

[3] We refer to [2] for an intelligible definition of a random lattice.
[4] Note that *balls* are implicitly defined using the quadratic norm.

for any $\mathbf{y} \in \mathbb{Z}^n$. Sampling along this distribution can be done efficiently [11]. We denote by $\eta_\epsilon([Q])$ the smoothing parameter of a class. Informally, above this threshold, the reduction *mod 1* of a (continuous) Gaussian vector of covariance $Q' \in [Q]$ is indistinguishable from a uniformly distributed point in $[0,1)^n$.

■ The Lattice Isomorphism Problem

The Lattice Isomorphism Problem (LIP) is a tentative at giving more attention to remarkable lattices and their strong geometric properties. It was first introduced in [21] and further studied in [17]. In its search variant, it asks to find an isomorphism between two lattices given as input, should it exist. In [11], Ducas and van Woerden bring the LIP problem on the cryptographic frontline and showed how to build core primitives founded on this problem. In an independent work [4], Bennett *et al.* study similar ideas but restricted to the $\mathbb{Z}^n$ lattice. This new approach for building primitives from lattices shares the flavour of first code-based and multivariate constructions [5]. In the former, illustrated for example by McEliece's public key encryption scheme, a code G with an efficient decoder is hidden by permutation S, P as $G^{pub} = SGP$; in the latter, an easy-to-invert quadratic map $\mathcal{Q}$ is masked by affine transformations T, S as $\mathcal{Q}^{pub} = T \circ \mathcal{Q} \circ S$. In the LIP paradigm, a lattice of known good basis is hidden by $\mathbf{B}^{\mathsf{pub}} = \mathbf{OBU}$ where $\mathbf{U}$ is unimodular (integer entries and determinant ± 1) and $\mathbf{O}$ is a rotation. Only someone knowing $(\mathbf{U}, \mathbf{O})$ can benefit from the good properties of $\mathbf{B}$. Should one only use $\mathbf{U}$, as in the GGH encryption scheme [16], then attacks exist, so on the first look, one would need to deal with rotation matrices having irrational entries.

LIP-based cryptography can fortunately be rephrased using quadratic forms instead of lattices bases. Therefore, rather than considering rotated lattices as in [4], one can essentially work *modulo rotation*. We have $(\mathbf{B}^{\mathsf{pub}})^\mathsf{T}\mathbf{B}^{\mathsf{pub}} = \mathbf{U}^\mathsf{T}(\mathbf{B}^\mathsf{T}\mathbf{B})\mathbf{U}$ which gives a nice reformulation over the Gram matrix of $\mathbf{B}$ and $\mathbf{B}^{\mathsf{pub}}$, that are quadratic forms. In other words, the lattices described by the bases $\mathbf{B}$ and $\mathbf{B}'$ are isomorphic if and only if the corresponding Gram matrices $Q_\mathbf{B}, Q_{\mathbf{B}'}$ are congruent *using unimodular matrices*. With such a reformulation, a natural playground for representing the space is to work within $\mathbb{Z}^n$ but with a norm that reflects the geometry of lattices of one's choice. This observation gives rise to the following formulation of the search version of LIP.

Definition 1 (wc-LIP^{Q_0}, quadratic form version). *Given quadratic forms Q and Q_0 from $\mathrm{S}_n^{++}(\mathbb{R})$, find $\mathbf{U} \in \mathrm{GL}_n(\mathbb{Z})$ such that $Q = \mathbf{U}^T Q_0 \mathbf{U}$, should it exists.*

Staying concretely within $\mathbb{Z}^n$ seems to significantly ease implementations of LIP-based scheme, compared to their equivalent in standard lattice-based cryptography. The recent `Hawk` proposal [10] to NIST's second call for standardization can be compared performance-wise to `Falcon` [22], boasting much simpler[5] implementation constraints on top.

[5] Implementing the Gaussian sampler of `Falcon` is a notoriously difficult task.

As many cryptographic problems, LIP is unlikely to be NP hard. Nevertheless, it benefits from worst-case to average-case self-reduction *within* an instantiation *class*. Informally, for a fixed equivalence class $[Q_0] = \{\mathbf{U}^\mathsf{T} Q_0 \mathbf{U} \mid \mathbf{U} \in \mathrm{GL}_n(\mathbb{Z})\}$ of quadratic forms, there is an efficient way [11, lem. 3.9] of generating a random member of the class with corresponding LIP instance being *as hard as possible*[6]. In this document, we denote by $\mathtt{QFS}_{Q,s}$ the (non-deterministic) sampler within the class of $[Q]$ with parameter s, and by $\mathcal{D}_s([Q])$ its output distribution. Additionally, the connection of LIP with the Graph Isomorphism Problem (GIP) [25] accounts for its assumed theoritical hardness.

Besides the work of [4], the LIP problem restricted to the $\mathbb{Z}^n$ lattice has been the focus of [12,6], to improve understanding of LIP hardness in what is arguably the most simple lattice one can think of. In particular, it helps driving throughout hazardous choices when instantiating the LIP problem on concrete quadratic form classes.

For now, despite the exciting connection with other isomorphism problems, such as the GIP problem, only low-level constructions exist: an identification scheme, two hash-and-sign signatures and a key-exchange mechanism, all described in [11].

■ A LIP-based public-key encryption scheme

We propose the *first* public-key encryption scheme *founded on LIP*. More precisely, it relies on a mild restriction of its distinguishing variant that we note $\Delta\mathsf{LIP}_{\mathsf{pke}}$. Under the assumption that the aforementionned problem is hard for the considered classes of quadratic forms, our scheme achieves IND-CPA security. The latter essentially means that any adversary has negligible probability of guessing the encrypted bit. This completes the set of fundamental primitives that can be built from the LIP assumption.

It should be noted that encryption of communications can already be done relying on LIP, using the key-exchange mechanism from [11]. In such a scheme, two parties agree upon a common private key by decoding a small (Gaussian) element. Private communications follow using symmetric encryption. Our scheme's target is beyond: besides encryption, a PKE scheme is a first step toward more advanced cryptography. For example, one could think of identity-based or attribute-based encryption as future objectives.

The new ΔLIP$_{pke}$ security assumption. This new assumption stems from the distinguishing variant of LIP that appears in [11], and consists in guessing to which which class a quadratic form belongs to given two proposals. At a high level our assumption $\Delta\mathsf{LIP}_{\mathsf{pke}}$ states that this variant remains secure when one restricts the set of possible instances to those where the smoothing parameter differ significantly between classes. Formally, the corresponding cryptographic game is defined as such.

[6] This is an example of worst-case to average-case reduction

Definition 2 ($\Delta\mathsf{LIP}^{Q_0,Q_1}_{\mathsf{pke},s}$[7]). *Given Q_0 and Q_1 from $\mathrm{S}_n^{++}(\mathbb{R})$ such that there exists an efficient algorithm for bounded distance decoding at distance $r \leq \lambda_1([Q_0])/2$ in $[Q_0]$ and $\eta_\varepsilon([Q_1]) < r$, and a quadratic form sampled as $(Q,\cdot) \leftarrow \mathtt{QFS}_{Q_b,s}$ where $b \hookleftarrow_\$ \{0,1\}$, guess b.*

The scheme. At a high-level, our PKE scheme can be seen as an adaptation of the Dual-Regev PKE from [15] with a LIP flavour. The design is inspired from [4], but our approach 1) is not restricted to (rotations of) the $\mathbb{Z}^n$ lattice (ie. with the class of equivalence of the n-dimensional identity); 2) is founded on a concrete assumption such as LIP — recall that rotations involves irrational numbers, while LIP can be dealt with mostly with rationals. The encryption correctness relies on the strong concentration of high-dimensional Gaussian vectors. Ciphertexts live in the unit discretised n-cube $\mathcal{H}_q^n$ with lower vertice at the origin; if the ciphertext corresponds to 0, it is uniformly distributed; if it corresponds to a 1, it is the reduction of an Gaussian modulo $\mathbb{Z}^n$, with a covariance matrix corresponding to the public quadratic form, and the closeness is therefore measured in the induced norm. On the one hand, once reduced modulo the cubic lattice, the Gaussian distribution will strongly concentrate around the vertices of $\mathcal{H}^n$, as shown in Figure 1. On the other hand, uniformly distributed vectors are much more likely to be in the inner part of the domain, since this is where most of the mass is. In the full version, we show that encryptions of 1 are in fact so close to the vertices that there is essentially no chance that they be mistaken for an element sampled uniformly at random within $\mathcal{H}_q^n$ or a discretized version of it, thanks to the properties of the smoothing parameter. Using the decoding algorithm $\mathtt{Dec}$ associated to the public form, the secret key owner can know how far the ciphertext was from the $\mathbb{Z}^n$ lattice, and conclude: a ciphertext close to the lattice corresponds to a 1, and a ciphertext far away from the lattice to a 0.

Theorem 1. *Restricted to instantiation where $\lambda_1(S_n)$ is smaller that 2 – which can always be achieved by rescaling[8] – any generated keypair (pk, sk) is such that for any bit b it holds with overwhelming probability that $\mathtt{Dec}(sk, \mathtt{Enc}(pk, b)) = b$.*

An intuition for the hardness, illustrated in Figure 1, is the following: an adversary ignoring the secret key is unable to observe $\mathcal{H}^n$ through the good norm, and distribution of points of $\mathcal{H}^n$ that are close to the vertices of the hypercube is mixed up in their point of view. In other words, given a ciphertext, they cannot compute efficiently the closeness of the cipher to a vertice without the secret key. More formally, we rely on the $\Delta\mathsf{LIP}_{\mathsf{pke}}$ hypothesis. Indeed, if the adversary cannot find which class the public form belongs to, in their view everything happens as if the smoothing parameter is big enough that the reduced Gaussian distribution mimics uniform distribution.

Theorem 2. *If the ΔLIP_{pke} problem is hard, then the scheme is IND-CPA secure.*

[7] When $Q_0, Q_1$and s are clear from context, sub/super-scripts are omitted.

[8] Recall that one of the purpose of LIP is to instantiate scheme on remarkable lattices: the first minimum of such lattices is likely to be known.

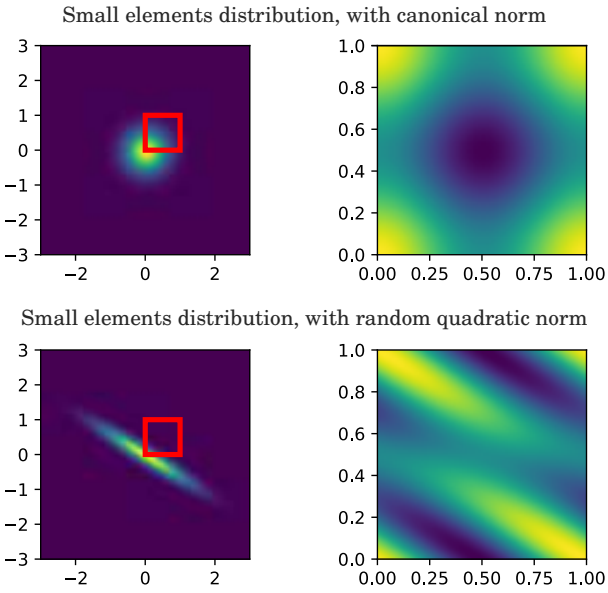


Fig. 1. Distribution of 2-dimensional small elements reduced modulo $\mathcal{H}^2$, either for the Euclidean norm or the norm induced by a random quadratic form.

The scheme is fully specified in Figure 2. At the core of the proof of Theorem 1 is a counting argument. We describe the number of points that are likely to be output by the reduced-gaussian sampler as the cardinal of the intersection of $\mathbb{Z}^n$ and an ellipsoid. Such estimates are the topic of classic mathematical problems, and we find a good-enough approximation thanks to a result of Landau [19]. The proof of Theorem 2 is then a game-based proof that turns the original IND-CPA game into a variant where encryptions of 0 and 1 follow the same distribution.

Looking for a concrete scheme, one can deviate from the parameter regime deduced from our proof, as is standard in cryptography. Then, the scheme can be instantiated on any particular lattice with efficient decoder, such as the hypercubic lattice $\mathbb{Z}^n$ as before, but also (again) Barnes-Wall lattices, and many more, enjoying possibly strong properties. For example, having a lattice minima closer to the Gaussian heuristics gives better concrete security at given dimension.

Encryption of bits may seems limiting but this is the case of many unstructured lattice encryption schemes (eg. Regev and Dual-Regev schemes). Pratical schemes in fact considered algebraically structured lattices to achieve m bits messages at the cost of possibly weaker assumptions, that restrict standard assumptions to specific classes[9] of lattices.

[9] The typical choices nowadays are lattices coming from ideals in cyclotomic number field field. See eg. [20] for details.

Protocol 1. Public-key encryption

Let $(S_n)_{n\in\mathbb{N}} = (\mathbf{B}_{S_n}^\mathsf{T}\mathbf{B}_{S_n})_{n\in\mathbb{N}}$ be a family of n-dimensional quadratic forms with an efficient decoding algorithm $\mathtt{Dec}_{\mathbf{S}_n}$ of known decoding radius $r_n < \lambda_1(S_n)/2$, and such that

$$\psi(n) = 1/\sqrt{\pi n} \cdot \det(S_n)^{-1/2} \cdot (2e\pi r_n^2)^{n/2}$$

is negligible. Let $n \in \mathbb{N}$, $s \geq \max\left\{\lambda_n(S), \|\mathbf{B}_S^*\| \cdot \sqrt{\ln(2n+4)/\pi}\right\}$. Further, let $q = \left\lceil \frac{sn}{r} \cdot \sqrt{\ln(2n+4)/\pi} \right\rceil$.

— **Key Generation** —

1: Sample $(P, \mathbf{U}) \leftarrow \mathtt{QFS}_{S,s}$ ▷ $P = \mathbf{U}^\mathsf{T} S \mathbf{U}$, see [11, lem. 3.9] for details
2: Return $(pk, sk) = (P, \mathbf{U})$

— **Encryption of 0** —

1: Sample $\mathbf{e} \hookleftarrow 1/q \cdot D_{P,qr/\sqrt{n}}$
2: Compute $\mathbf{c} \leftarrow \mathbf{e} \mod \mathbb{Z}^n$
3: Return $\mathbf{c}$, living in $\mathcal{H}_q^n$

— **Encryption of 1** —

1: Sample $\mathbf{c} \hookleftarrow \mathcal{U}(\mathcal{H}_q^n)$
2: Return $\mathbf{c}$, living in $\mathcal{H}_q^n$

— **Decryption** —

1: Compute $\mathbf{y} \leftarrow \mathtt{Dec}_S(\mathbf{Uc})$ ▷ Algorithm of protocol requierements
2: Compute $\mathbf{z} = \mathbf{c} - \mathbf{U}^{-1}\mathbf{y}$
3: If $\mathbf{z} \in \mathcal{H}_q^n$ and $\|\mathbf{z}\|_P \leq r$:
4: Return 0
5: Return 1

Fig. 2. Public-key encryption scheme

■ Security and cryptanalysis discussion

As observed in prior works, easy instances of $\Delta\mathsf{LIP}_{pke}$ exist: any pair of forms that do not have the same determinant; or the same parity; or more generally, that do not belong to the same genus, that is, the set of all equivalence classes for the relation $\mathcal{R}$ over p-adic integers for all prime p, can be distinguished in polynomial time. When restricted to pairs sharing at least these identified invariants, there are reasons to believe that $\Delta\mathsf{LIP}_{\mathsf{pke}}$ is a difficult problem. While our efforts unfortunately could not go beyond the current state of the art of the hardness of $\Delta\mathsf{LIP}$, we provide below other supporting observations that there are many hard instances. More can be found in the full version of this article.

Another invariant of a lattice is its *theta series*, a power series used to encode all lattice points by sorting them by (squared) norm. We however observe that theta series does not seem to be useful in the context of breaking $\Delta\mathsf{LIP}_{\mathsf{pke}}$. On the one hand, while the theta series can give accurate estimates of the smooth-

ing parameter of lattices [13], computing the first terms of the series amounts to solving the shortest vector problem by enumeration. This certainly requires exponential time at current state of knowledge. On the other hand, starting with dimension 4 [9], the theta series does not carry enough information to completely characterize a lattice. Lattices sharing the same theta series but not equivalent to one another are called *isospectral*. Knowing one such pair, one can build many other: if (L_1, L_2) is isospectral, then $(\Lambda \oplus L_1, \Lambda \oplus L_2)$ is also isospectral, for any lattice Λ.

Restricted to unimodular lattices, that is, self-dual lattices, the IND-CPA security of our scheme relies on the following mild assumption, which is reminiscent of [11].

Conjecture 1 (Mild version). For any (Q_0, Q_1) instance of $\Delta\mathsf{LIP}_{\mathsf{pke}}$ of dimension n, with equal polytime computable $\mathcal{R}$-invariants arithmetic quantities, $1 \leq \max\{gh(Q_i)/\lambda_1(Q_1), gh(Q_i^\vee)/\lambda_1(Q_i^\vee)\}$, the problem wc-$\Delta\mathsf{LIP}_{\mathsf{pke}}^{Q_0,Q_1}$ is $2^{\Theta(n)}$-hard.

While we gave clues that distinguishing between quadratic classes should not be easier if they differ by their smoothing parameter it may seem quite tricky to find family of lattices that could give an instance of $\Delta\mathsf{LIP}_{\mathsf{pke}}$ in the regime of parameter we need. Therefore, falsifying our security assumption seems tough at first sight. This is not surprising: the authors of [4] had already observed that it was quite unclear how to instantiate their $\mathbb{Z}^n$, rotation-based scheme with parameters reasonably close to their proof's regime (if possible at all!). Similar observation appear in [11]. This is partly due to the difficulty of understanding the genus of high-dimensional lattices. Nevertheless, in the full-version we support[10] our new assumption by showing that the famous Barnes-Wall lattices actually provide candidates for its plausibility, only missing the exact requirement by small *constant* factors. In dimension 2^m with m odd, these are unimodular even lattices, which are also known to form a single genus on top of many of fascinating properties: this may suggest to look more into these class of lattices.

■ Open questions

The first research direction we want to highlight is further reductionist effort. It seems reasonable to think that $\Delta\mathsf{LIP}$ problems restricted to classes that mainly differ by a gap on some quantity $\chi(\Lambda)$ is a problem simpler than a Gap problem (eg. GapSVP is a famous problem Gap on lattices, regarding the shortest vector's length) on this quantity. Is this even true? What can be said of the opposite direction? See that if those problems were in fact equivalent, $\Delta\mathsf{LIP}$ could be seen as a generic way to consider Gap problems in cryptography while easing space requirements.

The second one concerns falsifiability of the assumption: can the $\Delta\mathsf{LIP}_{\mathsf{pke}}$ assumption we rely on be effectively instantiated, and concretely cryptanalysed? As priorly stressed, gaps of $\mathbb{Z}^n$ are not that small, and one could expect better performances at fixed security level with our scheme instantiated on lattices with

[10] More precisely, besides cryptanalysis, we give clues for its falsifiability.

smaller gaps, such as Barnes-Wall lattices [3]. As highlighted in the previous paragraph, this question is open since [4]. A way of tackling this problem is a further study of the existence of optimal unimodular even quadratic form — a form is optimal when its minimum is the largest possible in a genus, and extremal when it reaches a known upper bound for these forms, of $2\lfloor n/24 \rfloor + 2$. It is known that extremal forms cannot exist above dimensions 163264, but they may exist for cryptographic sizes; nonetheless, we propose the conjecture that optimal forms could have minima large enough to answer the problem. Another promising direction is to use the Siegel-Weill mass formula that give (efficiently) the *average* theta series of a given genus: a mean argument could suffice to deduce the existence of quadratic forms that fit our requirements[11].

■ Acknowledgments

We thank the reviewers for their valuable suggestions and remarks, as well as Wessel Van Woerden for insightful discussions. The authors were supported by PEPR quantique France 2030 programme (ANR-22-PETQ-0008) and by ANR ASTRID project AMIRAL (ANR-21-ASTR-0016).

References

1. Shweta Agrawal, Mélissa Rossi, Anshu Yadav, and Shota Yamada. Constant input attribute based (and predicate) encryption from evasive and tensor lwe. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology – CRYPTO 2023*, pages 532–564, Cham, 2023. Springer Nature Switzerland.
2. Yoshinori Aono, Thomas Espitau, and Phong Q. Nguyen. Random lattices: Theory and practice, 2019. Available at https://espitau.github.io/bin/random_lattice.pdf.
3. E. S. Barnes and N. J. A. Sloane. New lattice packings of spheres. *Canadian Journal of Mathematics*, 35(1):117–130, 1983.
4. Huck Bennett, Atul Ganju, Pura Peetathawatchai, and Noah Stephens-Davidowitz. Just how hard are rotations of $\mathbb{Z}^n$? algorithms and cryptography with the simplest lattice. In *Advances in Cryptology–EUROCRYPT 2023: 42nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Lyon, France, April 23-27, 2023, Proceedings, Part V*, pages 252–281. Springer, 2023.
5. Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen, editors. *Post-Quantum Cryptography*. Springer Berlin, Heidelberg, 2008.
6. Tamar Lichter Blanks and Stephen D Miller. Generating cryptographically-strong random lattice bases and recognizing rotations of $\mathbb{Z}^n$. In *Post-Quantum Cryptography: 12th International Workshop, PQCrypto 2021, Daejeon, South Korea, July 20–22, 2021, Proceedings 12*, pages 319–338. Springer, 2021.
7. Jonathan Bootle, Vadim Lyubashevsky, Ngoc Khanh Nguyen, and Alessandro Sorniotti. A framework for practical anonymous credentials from lattices. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology – CRYPTO 2023*, pages 384–417, Cham, 2023. Springer Nature Switzerland.

[11] This would not necessarily give *concrete* falsifiability as the matching form may remain unknown.

8. Valerio Cini, Russell W. F. Lai, and Giulio Malavolta. Lattice-based succinct arguments from vanishing polynomials. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology – CRYPTO 2023*, pages 72–105, Cham, 2023. Springer Nature Switzerland.
9. J. H. Conway and N. J. A. Sloane. Four-dimensional lattices with the same theta series. *International Mathematics Research Notices*, 1992(4):93–96, 02 1992.
10. Léo Ducas, Eamonn W Postlethwaite, Ludo N Pulles, and Wessel van Woerden. Hawk: Module lip makes lattice signatures fast, compact and simple. In *Advances in Cryptology–ASIACRYPT 2022: 28th International Conference on the Theory and Application of Cryptology and Information Security, Taipei, Taiwan, December 5–9, 2022, Proceedings, Part IV*, pages 65–94. Springer, 2023.
11. Léo Ducas and Wessel van Woerden. On the lattice isomorphism problem, quadratic forms, remarkable lattices, and cryptography. In *Advances in Cryptology–EUROCRYPT 2022: 41st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Trondheim, Norway, May 30–June 3, 2022, Proceedings, Part III*, pages 643–673. Springer, 2022.
12. Léo Ducas. Provable lattice reduction of $\mathcal{F}^n$ with blocksize $n/2$. Cryptology ePrint Archive, Paper 2023/447, 2023. https://eprint.iacr.org/2023/447.
13. Thomas Espitau, Alexandre Wallet, and Yang Yu. On gaussian sampling, smoothing parameter and application to signatures. In Jian Guo and Ron Steinfeld, editors, *Advances in Cryptology - ASIACRYPT 2023 - 29th International Conference on the Theory and Application of Cryptology and Information Security, Guangzhou, China, December 4-8, 2023, Proceedings, Part VII*, pages 65–97. Springer, 2023.
14. Craig Gentry. Fully homomorphic encryption using ideal lattices. In *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, STOC '09, page 169–178, New York, NY, USA, 2009. Association for Computing Machinery.
15. Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. *Proceedings of the fortieth annual ACM symposium on Theory of computing*, 2008.
16. Oded Goldreich, Shafi Goldwasser, and Shai Halevi. Public-key cryptosystems from lattice reduction problems. In *CRYPTO'97: Proceedings of the 17th Annual International Cryptology Conference on Advances in Cryptology*, pages 112–131, 1997.
17. Ishay Haviv and Oded Regev. On the lattice isomorphism problem. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 391–404. SIAM, 2014.
18. Corentin Jeudy, Adeline Roux-Langlois, and Olivier Sanders. Lattice signature with efficient protocols, application to anonymous credentials. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology – CRYPTO 2023*, pages 351–383, Cham, 2023. Springer Nature Switzerland.
19. E. Landau. Über gitterpunkte in mehrdimensionalen ellipsoiden. *Mathematische Zeitschrift*, 21:126–132, 1924.
20. Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In Henri Gilbert, editor, *Advances in Cryptology – EUROCRYPT 2010*, pages 1–23, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
21. W. PLESKEN and B. SOUVIGNIER. Computing isometries of lattices. *Journal of Symbolic Computation*, 24(3):327–334, 1997.
22. Thomas Prest, Pierre-Alain Fouque, Jeffrey Hoffstein, Paul Kirchner, Vadim Lyubashevsky, Thomas Pornin, Thomas Ricosset, Gregor Seiler, William Whyte, and Zhenfei Zhang. Falcon. *Post-Quantum Cryptography Project of NIST*, 2020.

23. Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6), sep 2009.
24. Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, oct 1997.
25. Mathieu Dutour Sikirić, Achill Schürmann, and Frank Vallentin. Complexity and algorithms for computing voronoi cells of lattices. *Mathematics of Computation*, 78(267):1713–1731, sep 2009.

A Niederreiter public-key cryptosystem using a convolutional approach

Paulo Almeida[1], Miguel Beltrá[2], and Diego Napp[2]

[1] University of Aveiro palmeida@ua.pt
[2] University of Alicante miguel.beltra@ua.es, diego.napp@ua.es

Abstract. The Niederreiter public-key cryptosystem is believed to be secure against attackers having access to a large quantum computer when the parameters are set adequately. As a main drawback, the scheme needs to use very large public keys to reach acceptable levels of security. In this work we explore a variation of the classical scheme that allows to reduce the size of this key. We replace the scrambling matrix S and the permutation matrix P used in the original scheme with polynomial matrices $S(D)$ and $R(D)$. The public key is then the parity-check matrix of a convolutional code. Instead of encrypting a constant vector $\mathbf{e}$ of information we encrypt a polynomial vector $\mathbf{e}(D)$. Thus, the scheme can be seen as a convolutional version of the Niederreiter public-key cryptosystem. We analyze its security against general decoding and structural attacks and conclude that this approach can be of interest in order to reduce the size of the key.

1 Introduction

Nowadays all widely used public-key cryptosystems (PKC) are based on the problem of integer factoring [23] or the computation of discrete logarithms [5,6,18,11]. These problems can be easily solved with a quantum computer using Shor's algorithm [27,28]. Due to the continuous progress in this area, the National Institute of Standards and Technology (NIST) emphasized the necessity of having alternative PKCs and started a standardization project to define quantum-resistant schemes. Most of the promising schemes are based on problems coming from lattice theory [25,14,22] and coding theory [4,2,17].

The McEliece cryptosystem [16] was introduced in 1978 and became the first PKC based on a coding theory problem. Even though no feasible attack is known when the parameters are chosen adequately, the scheme has not been used in practice. The reason is mainly due to the large public key of the scheme.

The Niederreiter cryptosystem is a variation of the McEliece cryptosystem proposed in 1986 by Niederreiter [19]. This variant is equivalent to the McEliece cryptosystem in terms of security [12], but it can reduce the size of the public key. In fact, the NIST standardization project proposals *Classic McEliece* [4] and *BIKE* [2] are Niederreiter schemes using binary Goppa and QC-MDPC codes respectively.

In this work we propose a variation of the original Niederreiter scheme using polynomial matrices to hide a binary Goppa code. Instead of sending a large vector $\mathbf{e}$, we send a finite sequence of smaller vectors $\mathbf{e}_i$, $i = 0, 1, \ldots \ell - 1$, represented by a polynomial vector $\mathbf{e}(D) = \mathbf{e}_0 + \mathbf{e}_1 D + \cdots + \mathbf{e}_{\ell-1} D^{\ell-1}$. This idea was explored in the context of the McEliece cryptosystem in [1]. We study the security of the new scheme against general decoding and structural attacks and compare the size of the keys with those keys of *Classic McEliece*. In all cases the proposed scheme uses smaller keys. The reduction is relevant in the cases of 192-bit and 256-bit security.

This work is structured as follows: In Section 2 we recall some basic definitions from coding theory and describe the Niederreiter PKC. In Section 3 we introduce the new scheme. Section 4 contains an explanation of possible attacks against the scheme. In Section 5 we propose parameters to reach the standard security levels and compare the obtained public key sizes with those used in *Classic McEliece* PKC. Finally, in Section 6 we provide some conclusions and future work.

2 Preeliminaries

In this section we introduce the basic definitions we need to describe the proposed variation of the Niederreiter cryptosystem. Throughout the whole paper, vectors are column vectors. We introduce the following notation: given an integer j and a positive integer ℓ, we denote by $[j]_\ell$ the canonical representative of j modulo ℓ, i.e., the smallest non negative integer that is congruent with j modulo ℓ. It is worthwhile to point out that, for $-\ell \leq j \leq 2\ell - 1$, we have

$$[j]_\ell = \begin{cases} j + \ell & \text{when} \quad -\ell \leq j \leq -1; \\ j & \text{when} \quad 0 \leq j \leq \ell - 1; \\ j - \ell & \text{when} \quad \ell \leq j \leq 2\ell - 1. \end{cases}$$

2.1 Coding theory

Let $\mathbb{F}_q$ be the finite field of q elements, where q is a prime power. An $[n, k]$ **linear code** $\mathcal{C}$ is a subspace of $\mathbb{F}_q^n$ of dimension k. Its vectors are called **codewords**. This space can be defined either by a **generator matrix** $G \in \mathbb{F}_q^{k \times n}$, that is, $\mathcal{C} = \{\mathbf{u}^\top G : \mathbf{u} \in \mathbb{F}_q^k\}$ or by a **parity-check matrix** $H \in \mathbb{F}_q^{(n-k) \times n}$, that is, $\mathcal{C} = \{H\mathbf{c} = \mathbf{0}_{n-k} : \mathbf{c} \in \mathbb{F}_q^n\}$. Given a vector $\mathbf{v} \in \mathbb{F}_q^n$, its **syndrome** is $\mathbf{s} = H\mathbf{v}$. The **Hamming weight** of a vector $\mathrm{wt}(\mathbf{v})$ is the number of non-zero components of that vector. The **distance** between two vectors $d(\mathbf{v}, \mathbf{w})$ is the number components where they differ. The **minimum distance** $d(\mathcal{C})$ of a code is the minimum distance between any two different codewords, and it coincides with the minimum weight among the nonzero codewords of the code. The value $\left\lfloor \frac{d(\mathcal{C})-1}{2} \right\rfloor$ is called the **error correction capability of the code**.

The **Syndrome Decoding Problem** (SDP) is an important problem in coding theory. It can be stated as follows:

SDP: *Given a parity-check matrix* $H \in \mathbb{F}_q^{(n-k)\times n}$ *of a code, a vector* $\boldsymbol{s} \in \mathbb{F}_q^{n-k}$ *and a non negative integer* t *find, if exists, a vector* $\boldsymbol{e} \in \mathbb{F}_q^n$ *such that* $\boldsymbol{s} = H\boldsymbol{e}$ *and* $\mathrm{wt}(\boldsymbol{e}) \leq t$.

If t is the error correction capability of the code, then the solution of the SDP is unique. The SDP problem is known to be NP-complete [3] but there exist families of codes admitting efficient algorithms to solve it in polynomial time, for example the family of Goppa codes.

Definition 1. *Let* q *be a prime power,* m *a positive integer,* $g(z) \in \mathbb{F}_{q^m}[z]$ *and* $L = (\alpha_1, \alpha_2, \ldots, \alpha_n) \subset \mathbb{F}_{q^m}$ *be an ordered subset of* $\mathbb{F}_{q^m}$ *of different elements that are not roots of* $g(z)$*. The* ***Goppa code*** *of parameters* g *and* L *is defined as*

$$\Gamma(g, L) = \left\{ (c_1, c_2, \ldots, c_n) \in \mathbb{F}_q^n : \sum_{i=1}^{n} \frac{c_i}{z - \alpha_i} \equiv 0 \pmod{g(z)} \right\}.$$

L *is called the* ***support*** *of* $\Gamma(g, L)$.

The dimension of a Goppa code is lower bounded by $n - m\deg(g)$. Binary Goppa codes defined using a squarefree polynomial are of interest due to the following theorem.

Theorem 1. *[15, Ch. 12, §3, Thm. 6] Let* $\Gamma(g, L)$ *be a Goppa code defined over* $\mathbb{F}_2$*. If* $g(z)$ *is squarefree then* $d(\Gamma(g, L)) \geq 2\deg(g) + 1$.

Irreducible polynomials defined over a finite field are squarefree [9, pag. 289, Ex. 13], thus, from now on, we restrict our attention to binary Goppa codes defined by a irreducible monic polynomial of dimension $n - m\deg(g)$. The error correction capability of these codes is at least $\deg(g)$ and we can correct that number of errors in polynomial time using Patterson's algorithm [21].

A **convolutional code** $\mathcal{C}$ of rate k/n is an $\mathbb{F}_q[D]$-submodule of $\mathbb{F}_q^n[D]$ of rank k given by a polynomial **encoder matrix** $G(D) \in \mathbb{F}_q^{k\times n}[D]$,

$$\mathcal{C} = \{\mathbf{u}^\top(D)G(D) : \mathbf{u}(D) \in \mathbb{F}_q^k[D]\}.$$

Not all convolutional codes admit a dual representation with a parity-check matrix. If a code has a basic generator matrix $G(D)$, i.e., a generator matrix with polynomial right inverse, then it can can be defined in terms of a parity-check matrix $H(D) \in \mathbb{F}_q^{(n-k)\times n}[D]$ [24]

$$\mathcal{C} = \{\mathbf{c}(D) \in \mathbb{F}_q^n[D] : H(D)\mathbf{c}(D) = \mathbf{0}_{n-k}\}.$$

For more details about convolutional codes we refer to [10].

2.2 The Niederreiter cryptosystem

The Niederreiter PKC [19] is a code-based cryptosystem based on the hardness of the SDP. Its public key is a disguised version of the parity-check matrix of some

linear code having an efficient syndrome decoding algorithm while the private key is this algorithm together with the elements used to disguise that partiy-check matrix. The ciphertext is the syndrome of some low weight vector that can be efficiently recovered with the syndrome decoding algorithm. The scheme is described as follows:

Public parameters: a prime power q and three positive integers n, k and t.

Private and public keys: the private key is composed by three matrices:

- a non singular matrix $S \in \mathbb{F}_q^{(n-k)\times(n-k)}$;
- the parity-check matrix $H \in \mathbb{F}_q^{(n-k)\times n}$ of an $[n,k]$ linear code over $\mathbb{F}_q$ admitting an efficient syndrome decoding algorithm capable to correct t errors;
- and a permutation matrix $P \in \mathbb{F}_q^{n\times n}$.

The public key is the matrix $H' = SHP$.

Encryption procedure: the message to be transmitted is a vector $\mathbf{e} \in \mathbb{F}_q^n$ of $\mathrm{wt}(\mathbf{e}) \leq t$. The ciphertext is $\mathbf{s} = H'\mathbf{e}$.

Decryption procedure: multiply $\mathbf{s}$ by S^{-1} to obtain $S^{-1}\mathbf{s} = HP\mathbf{e}$. As P is a permutation, $\mathrm{wt}(P\mathbf{e}) \leq t$ so the efficient decoding algorithm can be applied to $HP\mathbf{e}$ to recover $P\mathbf{e}$. Finally, multiply by the inverse of P to recover $\mathbf{e}$.

3 A convolutional version of the Niederreiter cryptosystem

In this section we present a variation of the Niederreiter cryptosystem. We multiply the parity-check matrix of a binary Goppa code by two polynomial matrices to hide its structure. The resulting matrix is the parity-check matrix of some convolutional code. The plaintext is a finite sequence of vectors of low weight and the ciphertext is the syndrome of this sequence. We impose a restriction on the weights of the sequence in order to guarantee that the message can be correctly decrypted.

Public parameters: The message length ℓ and five positive integers n, k, t, s, r, with $s + r < \ell$.

Private key: It is composed by three matrices:

1) $H \in \mathbb{F}_2^{(n-k)\times n}$ is the parity-check matrix of an $[n,k]$ binary Goppa code able to correct t errors.

2) $S(D) \in \mathbb{F}_2^{(n-k)\times(n-k)}[D]$ is a polynomial matrix of the form

$$S(D) = S_0 + S_1 D + S_2 D^2 + \cdots + S_s D^s$$

that is non singular over $\mathbb{F}_2[D]/(D^\ell - 1)$.

3) $R(D) \in \mathbb{F}_2^{n\times n}[D]$ is a polynomial matrix of the form

$$R(D) = R_0 + R_1 D + R_2 D^2 + \cdots + R_r D^r$$

that is non singular over $\mathbb{F}_2[D]/(D^\ell - 1)$ and such that each column of each coefficient matrix R_i, $i \in \{0, 1, \ldots, r\}$, has no more than one nonzero element.

Public key: A polynomial matrix $H'(D) \in \mathbb{F}_2^{(n-k)\times n}[D]$ defined as

$$H'(D) = S(D)HR(D) = H'_0 + H'_1 D + H'_2 D^2 + \cdots + H'_{s+r} D^{s+r}.$$

Encryption: The message $\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_{\ell-1}$ is represented as a polynomial vector

$$\mathbf{e}(D) = \mathbf{e}_0 + \mathbf{e}_1 D + \cdots + \mathbf{e}_{\ell-1} D^{\ell-1} \in \mathbb{F}_2^n[D]$$

satisfying the condition

$$\mathrm{wt}([\mathbf{e}_{[i]_\ell}\ \mathbf{e}_{[i-1]_\ell}\ \mathbf{e}_{[i-2]_\ell}\ \cdots\ \mathbf{e}_{[i-r]_\ell}]) \le t, \qquad \text{for all } i \in \{0, 1, \ldots, \ell-1\}, \tag{1}$$

that is, the weight of any r consecutive vectors $\mathbf{e}_i$ adds up to t. The ciphertext is the vector

$$\mathbf{s}(D) = H'(D)\mathbf{e}(D) \pmod{D^\ell - 1}.$$

Decryption: To decrypt a ciphertext $\mathbf{s}(D)$, we first compute the inverse of $S(D)$ over $\mathbb{F}_2[D]/(D^\ell - 1)$, namely $S^{-1}(D)$, and the vector

$$\hat{\mathbf{s}}(D) = S^{-1}(D)\mathbf{s}(D) \pmod{D^\ell - 1}.$$

This vector is of the form $\hat{\mathbf{s}}_0 + \hat{\mathbf{s}}_1 D + \cdots + \hat{\mathbf{s}}_{\ell-1} D^{\ell-1}$. It happens that for each coefficient $\hat{\mathbf{s}}_i$ there exists some vector $\hat{\mathbf{e}}_i \in \mathbb{F}_2^n$ with $\mathrm{wt}(\hat{\mathbf{e}}_i) \le t$ such that $\hat{\mathbf{s}}_i = H\hat{\mathbf{e}}_i$. Hence, we can recover each $\hat{\mathbf{e}}_i$ using an efficient decoding algorithm, e.g. Patterson's algorithm [21]. Let $\hat{\mathbf{e}}(D) = \hat{\mathbf{e}}_0 + \hat{\mathbf{e}}_1 D + \cdots + \hat{\mathbf{e}}_{\ell-1} D^{\ell-1}$. We recover the message $\mathbf{e}(D)$ by computing

$$\mathbf{e}(D) = R^{-1}(D)\hat{\mathbf{e}}(D) \pmod{D^\ell - 1}.$$

We can prove that this decryption procedure is correct.

Theorem 2. *Given a ciphertext* $\mathbf{s}(D) = H'(D)\mathbf{e}(D) \pmod{D^\ell - 1}$*, with* $\mathbf{e}(D)$ *satisfying condition* (1)*, the decryption procedure described above computes* $\mathbf{e}(D)$ *correctly.*

Proof. Let us define the polynomials

$$\begin{aligned} \hat{\mathbf{s}}(D) &= S^{-1}(D)\mathbf{s}(D) \pmod{D^\ell - 1}; \\ \hat{\mathbf{e}}(D) &= R(D)\mathbf{e}(D) \pmod{D^\ell - 1}. \end{aligned}$$

These are polynomials of degree $\ell - 1$ and so, we can write

$$\hat{\mathbf{s}}(D) = \hat{\mathbf{s}}_0 + \hat{\mathbf{s}}_1 D + \cdots + \hat{\mathbf{s}}_{\ell-1} D^{\ell-1} \in \mathbb{F}_2^{n-k}[D];$$
$$\hat{\mathbf{e}}(D) = \hat{\mathbf{e}}_0 + \hat{\mathbf{e}}_1 D + \cdots + \hat{\mathbf{e}}_{\ell-1} D^{\ell-1} \in \mathbb{F}_2^{n}[D],$$

and the coefficients of $\hat{\mathbf{e}}(D)$ are given by

$$\hat{\mathbf{e}}_i = R_0 \mathbf{e}_{[i]_\ell} + R_1 \mathbf{e}_{[i-1]_\ell} + R_2 \mathbf{e}_{[i-2]_\ell} + \cdots + R_r \mathbf{e}_{[i-r]_\ell},$$

for each $i \in \{0, 1, \ldots, \ell - 1\}$. Since each column of each R_j, $j \in \{0, 1, \ldots, r\}$ has at most one nonzero element, (1) implies $\mathrm{wt}(\hat{\mathbf{e}}_i) \leq t$. The first step is to compute $\hat{\mathbf{s}}(D)$. We have

$$S^{-1}(D)\mathbf{s}(D) \pmod{D^\ell - 1} = HR(D)\mathbf{e}(D) \pmod{D^\ell - 1},$$

so it holds that $\hat{\mathbf{s}}(D) = H\hat{\mathbf{e}}(D)$. Then each coefficient of $\hat{\mathbf{s}}(D)$ is of the form $\hat{\mathbf{s}}_i = H\hat{\mathbf{e}}_i$. Since each $\mathrm{wt}(\hat{\mathbf{e}}_i) \leq t$, the use of the syndrome decoding algorithm to each $\hat{\mathbf{s}}_i$ in the second step of the decryption yields the vectors $\hat{\mathbf{e}}_i$. Once $\hat{\mathbf{e}}(D)$ is computed, then the message $\mathbf{e}(D)$ can be recovered by computing

$$R^{-1}(D)\hat{\mathbf{e}}(D) \pmod{D^\ell - 1} = R^{-1}(D)R(D)\mathbf{e}(D) \pmod{D^\ell - 1} = \mathbf{e}(D).$$

since $R(D)$ is invertible over $\mathbb{F}_2[D]/(D^\ell - 1)$.

4 Analysis of the possible attacks

In this section we study the security of the scheme against possible message recovery and key recovery attacks.

4.1 Message recovery attacks

Equation $\mathbf{s}(D) = H'(D)\mathbf{e}(D) \pmod{D^\ell - 1}$ can be written using matrices as

$$\begin{bmatrix} \mathbf{s}_0 \\ \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_{\ell-1} \end{bmatrix} = \underbrace{\begin{bmatrix} H'_0 & & & & & & H'_{s+r} & \cdots & H'_2 & H'_1 \\ H'_1 & H'_0 & & & & & & H'_{s+r} & \cdots & H'_2 \\ H'_2 & H'_1 & \ddots & & & & & & \ddots & \vdots \\ \vdots & H'_2 & \ddots & \ddots & & & & & & H'_{s+r} \\ H'_{s+r} & \vdots & \ddots & \ddots & \ddots & & & & & \\ & H'_{s+r} & & \ddots & \ddots & \ddots & & & & \\ & & \ddots & & \ddots & \ddots & \ddots & & & \\ & & & H'_{s+r} & \cdots & H'_2 & H'_1 & H'_0 & & \\ & & & & H'_{s+r} & \cdots & H'_2 & H'_1 & H'_0 \end{bmatrix}}_{\mathcal{H}} \begin{bmatrix} \mathbf{e}_0 \\ \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_{\ell-1} \end{bmatrix}.$$

This can be seen as an instance of the SDP. Assuming $\mathrm{wt}(\mathbf{e}_i) \approx \frac{t}{(r+1)}$ in average (see (1)) we have

$$\mathrm{wt}([\mathbf{e}_0\ \mathbf{e}_1\ \cdots\ \mathbf{e}_{\ell-1}]) \approx \frac{\ell t}{(r+1)},$$

For adequate values of ℓ this becomes a huge linear system which is unexpected to be solved by any of the classical ISD algorithms. See for instance [7].

This naive approach does not take into account that the ciphertext is a sequence. An attacker can consider intervals of the sequence rather than the whole sequence, and try to recover individual blocks in an iterative attack instead of attacking the whole sequence at once. Notice that the ciphertext is computed modulo $D^\ell - 1$, so without loss of generality, the attacker must find the interval $[\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{\ell_0}]$, $\ell_0 \in \{0, 1, \ldots, \ell - 1\}$, such that the work factor of solving the linear system obtained from the first $(\ell_0 + 1)(n - k)$ rows of the above linear system viewed as an instance of the SDP is minimal when using one of the ISD algorithms. Let $\mathcal{H}(\ell_0)$ be matrix formed by the first $(\ell_0 + 1)(n - k)$ rows of $\mathcal{H}$. Notice that in this case

- The length of the code with parity-check matrix $\mathcal{H}(\ell_0)$ is

$$N = \min\{(\ell_0 + s + r + 1)n, \ell n\};$$

- The dimension of the code with parity-check matrix $\mathcal{H}(\ell_0)$ is

$$N - (\ell_0 + 1)(n - k);$$

- The part of $[\mathbf{e}_0\ \mathbf{e}_1\ \cdots\ \mathbf{e}_{\ell-1}]$ associated to $[\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{\ell_0}]$ can be assumed to have weight

$$\approx \min\left\{(\ell_0 + s + r + 1)\frac{t}{(r+1)}, \ell\frac{t}{(r+1)}\right\}.$$

4.2 Key recovery attacks

The best known key recovery attack against the Niederreiter cryptosystem based on binary Goppa codes is due to Loindreau and Sendrier [13], using the Sendrier's *support-splitting algorithm* (SSA) [26]. In this attack, the knowledge of g and the elements of the support $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, but not their order, can be used to recover $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. If $n = 2^m$, then $\{\alpha_1, \alpha_2, \ldots, \alpha_n\} = \mathbb{F}_{2^m}$, so the set is known. However, for $n < 2^m$, the attacker must guess $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ among the $\binom{2^m}{n}$ subsets of $\mathbb{F}_{2^m}$ of cardinality n. On the other hand, since g is not known, an exhaustive search has to be performed to find g. Recall that g is chosen to be monic and irreducible, and there are about $\approx \frac{2^{mt}}{t}$ possible polynomials of this class [15, Ch. 4, §8, Th. 15]. It is pointed out in [13] that this number can be reduced to $\approx \frac{2^{m(t-3)}}{mt}$ by exploiting certain group of automorphisms. Table 1 contains these values for certain parameters m and t.

The next theorem establishes a relation between the security of the original Niederreiter scheme with the security of the proposed scheme.

m	t	$\frac{2^{m(t-3)}}{mt}$
10	42	$\approx 2^{381.28}$
10	48	$\approx 2^{441.09}$
11	51	$\approx 2^{518.86}$
11	57	$\approx 2^{584.70}$
11	72	$\approx 2^{749.37}$
11	75	$\approx 2^{782.31}$
11	84	$\approx 2^{881.14}$

Table 1. Values of $\frac{2^{m(t-3)}}{mt}$ for certain parameters m and t.

Theorem 3. *Any attack aiming to recover the private key of the convolutional Niederreiter cryptosystem with parameters n, k, s, r can be transformed to an attack aiming to recover the private key of the original Niederreiter cryptosystem with parameters n and k.*

Proof. Assume an attacker knows the public key $H' = SHP$ of some instance of the original Niederreiter cryptosystem. Then the attacker can select $S(D) \in \mathbb{F}_2^{(n-k)\times(n-k)}[D]$ of degree ℓ and non singular over $\mathbb{F}_2[D]/(D^\ell - 1)$ at random and $R(D) \in \mathbb{F}_2^{n\times n}[D]$ of degree r, non singular over $\mathbb{F}_2[D]/(D^\ell - 1)$ and such that each column of each coefficient matrix R_i, $i \in \{0, 1, \ldots, r\}$, has at most one element. Let $S'(D) = S(D)S$ and $R'(D) = PR(D)$. It is clear that

$$H'(D) = S(D)H'R(D) = S'(D)HR'(D)$$

is a valid public key for the convolutional Niederreiter cryptosystem. Hence, any attack that can recover $S'(D)$, H and $R'(D)$ is an attack that can recover S, H and P.

Under the assumption that attacks based on the SSA are actually the best attacks against the original Niederreiter cryptosystem to recover the private key, the above theorem states that attacks against the proposed convolutional Niederreiter cryptosystem cannot be better than SSA attacks. Of course, this does not prove that the convolutional Niederreiter scheme is secure but any structural attack against the proposed scheme would have an important impact in the security of the original Niederreiter PKC.

5 Proposed parameters

In Table 2 we propose a set of parameters for the new scheme reaching different security levels. In all cases $s = r = 2$. Column Work Factor contains the expected number of operation of the attack according to the attack described in Subsection 4.1. Notice that all the values in Table 1 related with the attacks based on the SSA are higher than those obtained for the message recovery attacks. This number has been estimated using the tool `Cryptographic Estimators`

[8]. NIST establishes five categories for public-key cryptosystems [20]: category 1 (resp., 3 and 5) requires that the best attack against a particular instance of the cryptosystem needs at least 2^{128} bit operations (resp., 2^{192} and 2^{256}). Categories marked with an * have a slightly lower number of expected bit operations. These values have been considered in purpose to compare them with those proposed for *Classic McEliece* [4]. See Table 3. The reason of considering them is that most of the ISD attacks require an important use of the memory of the computer, and accessing to the memory has a non-negligible cost. Taking this cost into account, the number of operations increases, reaching the stated security category. See [7] for more details on the cost of these memory accesses. We can observe that one of the main drawbacks of the cryptosystem is the large ciphertexts we need to send to attain the same security level. For instance, to obtain a WF of 2^{180} we need to send a message which is $\frac{6732}{1248} \approx 5.4$ times larger than in *Classic McEliece* and the public key is $\frac{2793780}{4193280} \approx 0.66$ times smaller, that is, the key is reduced in a 0.34 %.

Table 2. Work factors and key sizes (in bits) for the proposal in Section 3.

Convolutional Niederreiter cryptosystem with binary Goppa codes							
m	n	t	ℓ	Work Factor	Security Level	Key size	Ciphertext
10	726	42	11	2^{129}	Category 1	1524600	4620
10	831	48	10	2^{140}	Category 1	1994400	4800
11	996	51	12	2^{182}	Category 3*	2793780	6732
11	1092	57	12	2^{195}	Category 3	3423420	7524
11	1396	72	13	2^{254}	Category 5*	5528160	10296
11	1455	75	13	2^{266}	Category 5	6001875	10725
11	1611	84	12	2^{278}	Category 5	7442820	11088

Table 3. Work factors and key sizes (in bits) for the Classic McEliece scheme [4].

Classic McEliece [4,7]						
m	n	t	Work factor	Security level	Key size	Ciphertext length
12	3488	64	2^{141}	Category 1	2088960	768
13	4608	96	2^{180}	Category 3*	4193280	1248
13	6688	128	2^{246}	Category 5*	8359936	1664
13	6960	119	2^{246}	Category 5*	8373911	1547
13	8192	128	2^{276}	Category 5	10862592	1664

6 Conclusions and future work

The new proposed scheme reduces the size of the keys of *Classic McEliece* scheme presented to the NIST standardization project, especially when compared with parameters of category 3 and 5. We expect that key size can be further reduced if we are able to adapt the cryptosystem to admit public keys in systematic form $[I_{n-k} \mid A(D)]$ with the condition that the entries in $A(D)$ are polynomials of certain small degree.

Acknowledgments

The first author was supported by The Center for Research and Development in Mathematics and Applications (CIDMA) through the Portuguese Foundation for Science and Technology (FCT Fundação para a Ciência e Tecnologia), references UIDB/04106/2020 (https://doi.org/10.54499/UIDB/04106/2020). The second and third authors were partially supported by the Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital (Grant No. CIAICO/2022/167) and by the Spanish Ministerio de Ciencia e Innovación via the grant with ref. PID2022-142159OB-I00.

References

1. P. Almeida, M. Beltrá, D. Napp, and C. Sebastião. Smaller keys for the McEliece cryptosystem: a convolutional variant with GRS codes. Submitted to IEEE Transactions on Information Theory.
2. N. Aragon, P. L. Barreto, S. Bettaieb, L. Bidoux, O. Blazy, J.-C. Deneuville, P. Gaborit, S. Ghosh, S. Gueron, T. Güneysu, C. A. Melchor, R. Misoczki, E. Persichetti, J. Richter-Brockmann, N. Sendrier, J.-P. Tillich, V. Vasseur, and G. Zémor. *BIKE - Bit Flipping Key Encapsulation, October 2022. Round 4 submission to the NIST post-quantum cryptography call.* https://bikesuite.org/.
3. E. R. Berlekamp, R. J. McEliece, and H. C. Van Tilborg. On the inherent intractability of certain coding problems. *IEEE Transactions on Information Theory*, 24(3):384–386, 1978.
4. D. J. Bernstein, T. Chou, T. Lange, I. von Maurich, R. Mizoczki, R. Niederhagen, E. Persichetti, C. Peters, P. Schwabe, N. Sendrier, J. Szefer, and W. Wen. *Classic McEliece: conservative code-based cryptography, October 2022. Round 4 submission to the NIST post-quantum cryptography call.*
5. W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976.
6. T. ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Transactions on Information Theory*, 31(4):469–472, 1985.
7. A. Esser and E. Bellini. Syndrome decoding estimator. In G. Hanaoka, J. Shikata, and Y. Watanabe, editors, *Public-Key Cryptography – PKC 2022*, pages 112–141, Cham, 2022. Springer International Publishing.
8. A. Esser, J. A. Verbel, F. Zweydinger, and E. Bellini. `CryptographicEstimators`: a software library for cryptographic hardness estimation. *IACR Cryptol. ePrint Arch.*, page 589, 2023. https://eprint.iacr.org/2023/589.

9. T. Hungerford. *Algebra*. Graduate Texts in Mathematics. Springer-Verlag New York, New York, 2003.
10. R. Johannesson and K. S. Zigangirov. *Fundamentals of Convolutional Coding*. Wiley-IEEE Press, New York, 2nd edition, 2015.
11. N. Koblitz. Elliptic curve cryptosystems. *Mathematics of Computation*, 48:203–209, 1987.
12. Y. X. Li, R. Deng, and X. M. Wang. On the equivalence of McEliece's and Niederreiter's public-key cryptosystems. *IEEE Transactions on Information Theory*, 40(1):271–273, 1994.
13. P. Loidreau and N. Sendrier. Weak keys in the McEliece public-key cryptosystem. *IEEE Transactions on Information Theory*, 47(3):1207–1211, 2001.
14. V. Lyubashevsky, L. Ducas, E. Kiltz, T. Lepoint, P. Schwabe, G. Seiler, D. Stehle, and S. Bai. *CRYSTALS-DILITHIUM*. https://pq-crystals.org/.
15. F. MacWilliams and N. Sloane. *The Theory of Error-Correcting Codes*. North-Holland Mathematical Library 16. Elsevier Science, 1977.
16. R. J. McEliece. A public-key cryptosystem based on algebraic coding theory. *Deep Space Network Progress Report*, 44:114–116, 1978.
17. C. A. Melchor, N. Aragon, S. Bettaieb, L. Bidoux, O. Blazy, J. Bos, J.-C. Deneuville, A. Dion, P. Gaborit, J. Lacan, E. Persichetti, J.-M. Robert, P. Véron, and G. Zémor. *HQC (Hamming Quasi-Cyclic)*. https://pqc-hqc.org/.
18. V. S. Miller. Use of elliptic curves in cryptography. In H. C. Williams, editor, *Advances in Cryptology — CRYPTO '85 Proceedings*, pages 417–426, Berlin, Heidelberg, 1986. Springer Berlin Heidelberg.
19. H. Niederreiter. Knapsack-type cryptosystems and algebraic coding theory. *Problems of Control and Information Theory*, 15(2):159–166, 1986.
20. NIST. Submission requirements and evaluation criteria for the post-quantum cryptography standardization process, December 2016.
21. N. Patterson. The algebraic decoding of goppa codes. *IEEE Transactions on Information Theory*, 21(2):203–207, 1975.
22. T. Prest, P.-A. Fouque, J. Hoffstein, P. Kirchner, V. Lyubashevsky, T. Pornin, T. Ricosset, G. Seiler, W. Whyte, and Z. Zhang. *CRYSTALS-KYBER*. https://falcon-sign.info/.
23. R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, feb 1978.
24. J. Rosenthal, J. M. Schumacher, and E. V. York. On behaviors and convolutional codes. *IEEE Trans. Automat. Control*, 42(6, part 1):1881–1891, 1996.
25. P. Schwabe, R. Avanzi, J. Bos, L. Ducas, E. Kiltz, T. Lepoint, V. Lyubashevsky, J. M. Schanck, G. Seiler, D. Stehle, and J. Ding. *CRYSTALS-KYBER*. https://pq-crystals.org/.
26. N. Sendrier. Finding the permutation between equivalent linear codes: the support splitting algorithm. *IEEE Transactions on Information Theory*, 46(4):1193–1203, 2000.
27. P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994.
28. P. W. Shor. Polynomial time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1481–1509, 1997.

Distance Distribution of Cyclic Orbit Flag Codes*

Clementa Alonso-González[1] and Miguel Ángel Navarro-Pérez[2]

[1] Dpto. de Matemáticas, Universidad de Alicante, Spain
clementa.alonso@ua.es
[2] Dpto. de Matemáticas, Universidad Carlos III de Madrid, Spain
mignavar@math.uc3m.es

Abstract. A *flag code* in network coding consists of a set of sequences of nested subspaces of $\mathbb{F}_q^n$ (flags), where $\mathbb{F}_q$ is the finite field with q elements. In this paper, we deal with *cyclic orbit flag codes*, that is, orbits of a Singer cycle of the general linear group acting on flags of $\mathbb{F}_q^n$. Inspired by the results in [15] and [10] about cyclic orbit codes, we completely characterize those cyclic orbit flag codes attaining both the best possible size and distance, that is, *optimal full-length cyclic orbit flag codes*. As a consequence, we can show that, for this family of codes, the distance distribution depends only on q, n and the dimensions of the subspaces in the generating flag.

Keywords: Cyclic orbit codes · Sidon spaces · Cyclic orbit flag codes · Distance distribution of a code.

1 Introduction

In [17], Trautmann *et al.* introduced the concept of *orbit codes* as those that arise from the action of subgroups of the general linear group $\mathrm{GL}(n,q)$ on subspaces of $\mathbb{F}_q^n$. In case the acting group is cyclic, we speak about *cyclic orbit codes* (see [8], [16]). This family of codes has been widely studied in the last years due to their interesting structure and properties.

On the other hand, in [9], Gluesing-Luerssen *et al.* consider subspace codes as collections of $\mathbb{F}_q$-vector subspaces of $\mathbb{F}_{q^n}$ by means of the natural isomorphism with $\mathbb{F}_q^n$, and study orbit codes given by the natural action of the multiplicative group $\mathbb{F}_{q^n}^*$ on $\mathbb{F}_q$-vector spaces. It is well known that, if $\mathcal{U}$ is a generating subspace of dimension k, and the cyclic orbit code $\mathrm{Orb}(\mathcal{U})$ attains the maximum possible distance, i.e. $2k$, then k must be a divisor of n and $\mathrm{Orb}(\mathcal{U})$ is a k-spread code of size $(q^n-1)/(q^k-1)$. This is the smallest cardinality for cyclic orbit codes generated by a k-dimensional subspace. Clearly, the largest size of such codes is $(q^n-1)/(q-1)$, and codes attaining it will be called *full-length orbit codes*. For dimensions $1 < k \leqslant \lfloor \frac{n}{2} \rfloor$, full-length orbit codes with the best possible distance value, that is $2k-2$, will be called *optimal full-length orbit codes*. In [9] and [16], it was conjectured the existence of optimal full-length orbit codes for any dimension k with $2k \leqslant n$.

Different constructions of optimal full-length orbit codes were proposed in [6], [7] and [14]. In 2018, Roth *et al.* (see [15]) gave a characterization result and showed that all of these codes are generated by subspaces known as *Sidon spaces*. In [10], the distance distribution of optimal full-length orbit codes was completely described.

Flag codes can be seen as a generalization of subspace codes whose codewords are sequences of nested subspaces (flags) of prescribed dimensions. In the network coding framework (see [11]), the first work handling flag codes is [12]. In that paper, the authors propose a natural extension of the multiplicative action of $\mathrm{GL}(n,q)$, previously used in [16, 17] on subspaces, to flags and provide several constructions

* Supported by Ministerio de Ciencia e Innovación (PID2022-142159OB-I00) and Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital (CIAICO/2022/167).

of *orbit flag codes*. From this seminal work, several related papers have dealt with different characterizations of flag codes of the maximum distance and also with different constructions, orbital or not, of them (see [3–5, 13]).

Following the approach in [9], it is possible to consider flags on $\mathbb{F}_{q^n}$ given by nested $\mathbb{F}_q$-subspaces of the field $\mathbb{F}_{q^n}$ and to study *cyclic orbit flag codes*, that is, orbits of $\mathbb{F}_{q^n}^*$ acting on the flag variety (see [2]). The main point in these papers is to address their study by taking into account the *best friend* of a flag, that is, the largest subfield of $\mathbb{F}_{q^n}$ over which every subspace in a flag is a vector space and the way it is obtained. As for the case of subspaces, cyclic orbit flag codes attaining both the best cardinality, that is again $(q^n-1)/(q-1)$, and the best possible distance are called *optimal full-length cyclic orbit flag codes*.

In this paper we completely characterize optimal full-length cyclic orbit flag codes by describing the set of suitable generating flags. Moreover, once we fix a flag $\mathcal{F}$ as a seed of the orbit, we can provide the distance distribution of $\mathrm{Orb}(\mathcal{F})$ and demonstrate that it depends just on n, q and the dimensions of the subspaces appearing in $\mathcal{F}$.

2 Preliminaries

Let q be a prime power and $\mathbb{F}_q$ the finite field with q elements. For any natural number $n \geqslant 1$, the field $\mathbb{F}_{q^n}$ is an n-dimensional vector space over $\mathbb{F}_q$. The set of k-dimensional subspaces of $\mathbb{F}_{q^n}$, that is, the *Grassmannian*, will be denoted by $\mathcal{G}_q(k,n)$. The Grassmannian can be endowed with a metric called the *subspace distance*: for any pair $\mathcal{U}, \mathcal{V} \in \mathcal{G}_q(k,n)$, we set

$$d_S(\mathcal{U},\mathcal{V}) = 2(k - \dim(\mathcal{U} \cap \mathcal{V})). \tag{1}$$

A *constant dimension code* $\mathcal{C}$ of dimension k and length n is a nonempty subset of $\mathcal{G}_q(k,n)$. In case $|\mathcal{C}| \geqslant 2$, its *minimum subspace distance* is given by

$$d_S(\mathcal{C}) = \min\{d_S(\mathcal{U},\mathcal{V}) \mid \mathcal{U},\mathcal{V} \in \mathcal{C},\ \mathcal{U} \neq \mathcal{V}\}.$$

Otherwise, we put $d_S(\mathcal{C}) = 0$. For further information on constant dimension codes, consult [18] and the references inside.

On the other hand, the multiplicative group $\mathbb{F}_{q^n}^*$ acts on $\mathcal{G}_q(k,n)$: for every k-dimensional subspace $\mathcal{U}$ of $\mathbb{F}_{q^n}$ and every nonzero element $\alpha \in \mathbb{F}_{q^n}$, we have that $\mathcal{U}\alpha = \{u\alpha \mid u \in \mathcal{U}\} \in \mathcal{G}_q(k,n)$. This action allows to build constant dimension codes, called *cyclic subspace codes*, as its orbits. Given an $\mathbb{F}_q$-subspace $\mathcal{U}$ of $\mathbb{F}_{q^n}$, the set

$$\mathrm{Orb}(\mathcal{U}) = \{\mathcal{U}\alpha \mid \alpha \in \mathbb{F}_{q^n}^*\}$$

is called the *cyclic orbit code* generated by $\mathcal{U}$ and the *stabilizer (subgroup)* of $\mathcal{U}$ is

$$\mathrm{Stab}(\mathcal{U}) = \{\alpha \in \mathbb{F}_{q^n}^* \mid \mathcal{F}\alpha = \mathcal{F}\}.$$

The code $\mathrm{Orb}(\mathcal{U})$ contains exactly $\frac{q^n-1}{q^m-1}$ elements if, and only if, $\mathbb{F}_{q^m}^*$ is the stabilizer of $\mathcal{U}$, for some divisor m of n. In particular, the largest possible orbit size is $\frac{q^n-1}{q-1}$. Cyclic orbit codes with this cardinality are called *full-length cyclic orbit codes*.

Concerning cyclic orbit codes with maximum distance, if $1 \leqslant k \leqslant \frac{n}{2}$ and $\mathcal{U}$ is a k-dimensional $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^n}$, then $\mathrm{Orb}(\mathcal{U})$ attains the maximum possible minimum distance (the value $2k$) if, and only if, k divides n and $\mathrm{Orb}(\mathcal{U}) = \mathrm{Orb}(\mathbb{F}_{q^k})$. In this situation, the orbit size is $\frac{q^n-1}{q^k-1}$, which is the smallest one for any cyclic orbit code generated by a k-dimensional subspace. In view of this, also in [9, 16], the

authors conjectured whether it was possible to build full-length cyclic orbit codes (of dimension $k > 1$) with distance $2k - 2$, which is the second best value for the minimum distance for dimension k, but the best one after prescribing the orbit size $\frac{q^n-1}{q-1}$. Such codes are denominated *optimal full-length cyclic orbit codes*.

A positive answer to the conjecture was given firstly in [6], [7] and [14], where the authors proposed different constructions of optimal full-length cyclic orbit codes. However, Roth *et al.* went further in [15] and characterized the class of suitable generating subspaces: they must be *Sidon spaces*.

Definition 1. *A subspace $\mathcal{U}$ of $\mathbb{F}_{q^n}$ is said to be a* Sidon space *if for every nonzero elements $a, b, c, d \in \mathcal{U}$ such that $ab = cd$, then $\{a\mathbb{F}_q, b\mathbb{F}_q\} = \{c\mathbb{F}_q, d\mathbb{F}_q\}$, where $e\mathbb{F}_q = \{e\lambda \,|\, \lambda \in \mathbb{F}_q\}$.*

They show that finding optimal full-length cyclic orbit codes is equivalent to constructing Sidon spaces of the desired dimension.

Theorem 1. *([15, Lemma 34]) Let $\mathcal{U}$ be a k-dimensional $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^n}$. The code* $\mathrm{Orb}(\mathcal{U})$ *is an optimal full-length cyclic orbit code if, and only if, $\mathcal{U}$ is a Sidon space.*

In [10] it is completely determined the distance distribution of codes generated by Sidon spaces.

2.1 Flag codes

Now, we recall the basic definitions on flag codes that already appear in [4, 5, 12], specially those ones concerning the particular class of cyclic orbit flag codes introduced in [2].

Definition 2. *A* flag $\mathcal{F} = (\mathcal{F}_1, \ldots, \mathcal{F}_r)$ *on the extension field $\mathbb{F}_{q^n}$ is a sequence of nested $\mathbb{F}_q$-vector subspaces*

$$\{0\} \subsetneq \mathcal{F}_1 \subsetneq \cdots \subsetneq \mathcal{F}_r \subsetneq \mathbb{F}_{q^n}.$$

*The subspace $\mathcal{F}_i$ is called the i-*th subspace *of $\mathcal{F}$ and the type of $\mathcal{F}$ is the vector* $(\dim(\mathcal{F}_1), \ldots, \dim(\mathcal{F}_r))$. *We say that a flag $\mathcal{F}'$ is a* subflag *of $\mathcal{F}$ if each subspace of $\mathcal{F}'$ is a also subspace of $\mathcal{F}$.*

The *flag variety* of type $(t_1, \ldots, t_r)$ on $\mathbb{F}_{q^n}$, that is, the set of flags of this type, will be denoted by $\mathcal{F}_q((t_1, \ldots, t_r), n)$. In this variety, we can define a metric that extends the subspace distance in (1). Given two flags $\mathcal{F} = (\mathcal{F}_1, \ldots, \mathcal{F}_r)$ and $\mathcal{F}' = (\mathcal{F}'_1, \ldots, \mathcal{F}'_r)$ in $\mathcal{F}_q((t_1, \ldots, t_r), n)$, their *flag distance* is

$$d_f(\mathcal{F}, \mathcal{F}') = \sum_{i=1}^{r} d_S(\mathcal{F}_i, \mathcal{F}'_i).$$

Definition 3. *A* flag code *of type $(t_1, \ldots, t_r)$ on $\mathbb{F}_{q^n}$ is a nonempty subset $\mathcal{C} \subseteq \mathcal{F}_q((t_1, \ldots, t_r), n)$. Its* minimum distance *is given by*

$$d_f(\mathcal{C}) = \min\{d_f(\mathcal{F}, \mathcal{F}') \mid \mathcal{F}, \mathcal{F}' \in \mathcal{C},\ \mathcal{F} \neq \mathcal{F}'\}.$$

whenever $|\mathcal{C}| \geqslant 2$. In case $|\mathcal{C}| = 1$, we put $d_f(\mathcal{C}) = 0$.

For type $\mathbf{t} = (t_1, \ldots, t_r)$, the previous distance is an even integer satisfying $0 \leqslant d_f(\mathcal{C}) \leqslant D^{(\mathbf{t},n)}$, where

$$D^{(\mathbf{t},n)} = 2\left(\sum_{t_i \leqslant \frac{n}{2}} t_i + \sum_{t_i > \frac{n}{2}} (n - t_i)\right) \tag{2}$$

and flag codes attaining this upper bound are called *optimum distance flag codes*.

There are constant dimension codes naturally associated to a flag code $\mathcal{C}$.

Definition 4. *Given a flag code $\mathcal{C}$ of type $(t_1, \ldots, t_r)$, we define the* i-projected code *of $\mathcal{C}$ as the set*

$$\mathcal{C}_i = \{\mathcal{F}_i \mid (\mathcal{F}_1, \ldots, \mathcal{F}_i, \ldots, \mathcal{F}_r) \in \mathcal{C}\} \subseteq \mathcal{G}_q(t_i, n).$$

The size of a flag code is clearly related with the ones of its projected codes as follows: $|\mathcal{C}_i| \leqslant |\mathcal{C}|$ for every $i = 1, \ldots, r$. In case $|\mathcal{C}_1| = \cdots = |\mathcal{C}_r| = |\mathcal{C}|$, we say that $\mathcal{C}$ is *disjoint*.

3 Optimal full-length cyclic orbit flag codes

Let us remember the concept introduced in [2] of *cyclic orbit flag code* constructed as the orbit of the multiplicative action of (cyclic) subgroups of $\mathbb{F}_{q^n}^*$ on flags on $\mathbb{F}_{q^n}$.

The cyclic group $\mathbb{F}_{q^n}^*$ acts on flags on $\mathbb{F}_{q^n}$ in a natural way: if $\mathcal{F} = (\mathcal{F}_1, \ldots, \mathcal{F}_r)$ is a flag of type $(t_1, \ldots, t_r)$ on $\mathbb{F}_{q^n}$ and $\alpha \in \mathbb{F}_{q^n}^*$, the flag $\mathcal{F}\alpha$ is

$$\mathcal{F}\alpha = (\mathcal{F}_1\alpha, \ldots, \mathcal{F}_r\alpha). \tag{3}$$

The orbit

$$\mathrm{Orb}(\mathcal{F}) = \{\mathcal{F}\alpha \mid \alpha \in \mathbb{F}_{q^n}^*\}. \tag{4}$$

is called the *cyclic orbit flag code* generated by $\mathcal{F}$ and it has projected codes $(\mathrm{Orb}(\mathcal{F}))_i = \mathrm{Orb}(\mathcal{F}_i)$, for all $1 \leqslant i \leqslant r$. The *stabilizer* of $\mathcal{F}$ is the subgroup

$$\mathrm{Stab}(\mathcal{F}) = \{\alpha \in \mathbb{F}_{q^n}^* \mid \mathcal{F}\alpha = \mathcal{F}\}. \tag{5}$$

Clearly, the cardinality of $\mathrm{Orb}(\mathcal{F})$ is

$$|\mathrm{Orb}(\mathcal{F})| = \frac{q^n - 1}{|\mathrm{Stab}(\mathcal{F})|} = \frac{q^n - 1}{q^m - 1} \tag{6}$$

for some divisor m of n. The minimum distance can be calculated as

$$d_f(\mathrm{Orb}(\mathcal{F})) = \min\{d_f(\mathcal{F}, \mathcal{F}\alpha) \mid \alpha \in \mathbb{F}_{q^n}^* \setminus \mathrm{Stab}(\mathcal{F})\}. \tag{7}$$

Following the terminology used in [9], let us start by giving a special name to the biggest possible orbits under this action.

Definition 5. *Let $\mathcal{F}$ be a flag on $\mathbb{F}_{q^n}$. We say that* $\mathrm{Orb}(\mathcal{F})$ *is a* full-length (cyclic orbit) flag code *if its size is the maximum possible one, that is, $\frac{q^n-1}{q-1}$.*

Concerning full length cyclic orbit codes that are also optimum distance flag codes, in [2] it was proved the following:

Theorem 2. *Given a flag $\mathcal{F}$ of type $\boldsymbol{t}$ on $\mathbb{F}_{q^n}$, a cyclic orbit flag code* $\mathrm{Orb}(\mathcal{F})$ *containing $\frac{q^n-1}{q-1}$ elements attains the maximum possible distance for their type, $D^{(\mathbf{t},n)}$, if and only if, the dimensions in the type vector are* 1 *and/or* $n-1$.

As a consequence, the possible type vectors for such codes are $(1), (n-1)$ and $(1, n-1)$. Clearly, if $\mathrm{Orb}(\mathcal{F})$ is full-length with type $\mathbf{t} \notin \{(1), (n-1), (1, n-1)\}$ then we have

$$d_f(\mathrm{Orb}(\mathcal{F})) \leqslant D^{(\mathbf{t},n)} - 2. \tag{8}$$

Hence, for flags $\mathcal{F}$ of type $\mathbf{t} \notin \{(1), (n-1), (1, n-1)\}$, the best possible value for the distance $d_f(\mathrm{Orb}(\mathcal{F}))$ is precisely $D^{(\mathbf{t},n)} - 2$.

Definition 6. *Let* $\mathbf{t} \neq (1), (n-1), (1, n-1)$ *be a type vector. A full-length flag code* $\mathrm{Orb}(\mathcal{F})$ *is said to be* optimal *if*

$$d_f(\mathrm{Orb}(\mathcal{F})) = D^{(\mathbf{t},n)} - 2.$$

We show first that for a type $\mathbf{t} \neq (1), (n-1), (1, n-1)$, a cyclic orbit flag code $\mathrm{Orb}(\mathcal{F})$ that attains the best possible distance must also have the best possible size. Moreover, we can show that cyclic orbit flag codes with $d_f(\mathrm{Orb}(\mathcal{F})) = D^{(\mathbf{t},n)} - 2$ must be disjoint.

Theorem 3. *Let* $\mathcal{F}$ *be a flag code of type* $\mathbf{t}$ *on* $\mathbb{F}_{q^n}$*. If* $d_f(\mathrm{Orb}(\mathcal{F})) = D^{(\mathbf{t},n)} - 2$*, then* $\mathrm{Orb}(\mathcal{F})$ *is full-length and disjoint.*

Let us now state a result that permits to us completely determine what is the admissible set of type vectors for optimal full length cyclic orbit flag codes. Consider $\mathcal{F}$ a flag of type $\mathbf{t}$ on $\mathbb{F}_{q^n}$. Whenever they appear in $\mathbf{t}$, we will denote the special dimensions

$$t_L = \max\{t_i \mid 2t_i \leqslant n\} \quad \text{and} \quad t_R = \min\{t_i \mid 2t_i \geqslant n\}. \tag{9}$$

Observe that, if $\frac{n}{2}$ is a dimension in the type vector, then $t_L = t_R = \frac{n}{2}$. Moreover, if every dimension is upper (resp. lower) bounded by $\frac{n}{2}$, then $L = r$ and R is not defined (resp $R = 1$ and L is not defined). In any other case, these dimensions t_L and t_R exist, they are different and, in fact, they are consecutive. With this notation, the next result holds.

Theorem 4. *Let* $\mathcal{F}$ *be a flag of type* $\mathbf{t} = (t_1, \ldots, t_r)$ *on* $\mathbb{F}_{q^n}$*. If* $\mathrm{Orb}(\mathcal{F})$ *is an optimal full length cyclic orbit flag code, then the type vector can only contain dimensions* $1, t_L, t_R$ *and* $n-1$*, with* $t_L \leqslant \frac{n}{2} \leqslant t_R$*, defined as in (9). Moreover, in the type vector it must appear at least one dimension different from* 1 *and* $n-1$*.*

Taking into account the previous result, we restrict our study to type vectors $\mathbf{t} = (1, k_1, k_2, n-1)$ with k_1 and k_2 satisfying either

$$k_1 < \frac{n}{2} < k_2 \quad \text{or} \quad k_1 = \frac{n}{2} = k_2 \tag{10}$$

and we determine the possible generating flags of an optimal full-length cyclic orbit flag code.

Theorem 5. *Let* $\mathcal{F} = (\ell, \mathcal{F}_1, \mathcal{F}_2, h)$ *be a flag of type* $\mathbf{t} = (1, k_1, k_2, n-1)$ *on* $\mathbb{F}_{q^n}$*. Then* $\mathrm{Orb}(\mathcal{F})$ *is an optimal full-length cyclic orbit flag code if, and only if, for every* $\alpha \in \mathbb{F}_{q^n} \setminus \mathbb{F}_q$*, it holds*

$$\dim(\mathcal{F}_1 \cap \mathcal{F}_1\alpha) + \dim(\mathcal{F}_2^{\perp} \cap (\mathcal{F}_2\alpha)^{\perp}) \leqslant 1 \tag{11}$$

and the equality holds for some choice of α*.*

Corollary 1. *Let* $\mathcal{F} = (\ell, \mathcal{F}_1, \mathcal{F}_2, h)$ *be a flag of type* $\mathbf{t} = (1, k_1, k_2, n-1)$ *on* $\mathbb{F}_{q^n}$*. Assume that* $\mathrm{Orb}(\mathcal{F})$ *is an optimal full-length cyclic orbit flag code, then both subspaces* $\mathcal{F}_1$ *and* $\mathcal{F}_2^{\perp}$ *are Sidon spaces of* $\mathbb{F}_{q^n}$*.*

Remark 1. It is important to point out that the converse of Corollary 1 is not necessarily true due to the fact that condition (11) must be satisfied. This is consequence of the nested structure of flags that usually produces a strong interdependence between the flag code parameters as pointed out in [1].

4 Distance distribution of optimal full-length cyclic orbit flag codes

Let $\mathcal{F} = (l, \mathcal{F}_1, \mathcal{F}_2, h)$ be a flag of type $(1, k_1, k_2, n-1)$ on $\mathbb{F}_{q^n}$ and assume that it generates an optimal full-length cyclic orbit flag. In particular, both codes $\mathrm{Orb}(\mathcal{F}_1)$ and $\mathrm{Orb}(\mathcal{F}_2^\perp)$ are generated by Sidon spaces and their distance (or intersection) distributions are known (see [10]). Let us write

$$\lambda_j^1 = |\{\mathcal{F}_1\alpha \mid d_S(\mathcal{F}_1, \mathcal{F}_1\alpha) = 2(k_1 - j),\ \alpha \in \mathbb{F}_{q^n} \setminus \mathbb{F}_q\}|,$$

$$\lambda_j^2 = |\{\mathcal{F}_2\alpha \mid d_S(\mathcal{F}_2^\perp, (\mathcal{F}_2\alpha^i)^\perp) = 2(n - k_2 - j),\ \alpha \in \mathbb{F}_{q^n} \setminus \mathbb{F}_q\}|,$$

for $j = 0, 1$. These values have been computed in [10, Th. 3.7] and they are exactly:

$$\begin{aligned} \lambda_1^1 &= \frac{q^{k_1}-1}{q-1}\frac{q^{k_1}-q}{q-1}, \qquad \lambda_0^1 = \frac{q^n-1}{q-1} - \lambda_1^1 - 1, \\ \lambda_1^2 &= \frac{q^{n-k_2}-1}{q-1}\frac{q^{n-k_2}-q}{q-1}, \lambda_0^2 = \frac{q^n-1}{q-1} - \lambda_1^2 - 1. \end{aligned} \tag{12}$$

One can also wonder how many flags in a cyclic orbit flag code are at certain distance from the generating flag. This leads to the next definition.

Definition 7. *Let* $\mathrm{Orb}(\mathcal{F})$ *be a cyclic orbit flag code on* $\mathbb{F}_{q^n}$ *with minimum distance* $d_f(\mathcal{C})$*. If* $\bar{d} = D^{(\mathbf{t},n)} - d_f(\mathcal{C})$*, then the* (flag) distance distribution *of* $\mathrm{Orb}(\mathcal{F})$ *is the sequence*

$$(\lambda_0^f, \lambda_1^f, \ldots, \lambda_{\frac{\bar{d}}{2}}),$$

where

$$\lambda_j^f = |\{\mathcal{F}\alpha \mid d_f(\mathcal{F}, \mathcal{F}') = D^{(\mathbf{t},n)} - 2j\}|,$$

for $0 \leqslant j \leqslant \frac{\bar{d}}{2}$.

In the particular case of considering flag codes with minimum distance $D^{(\mathbf{t},n)} - 2$, then their distance distribution is a sequence of length two and it counts the number of flags λ_0^f giving the maximum possible distance $D^{(\mathbf{t},n)}$ and λ_1^f, giving the value $D^{(\mathbf{t},n)} - 2$.

Theorem 6. *Let* $\mathrm{Orb}(\mathcal{F})$ *be an optimal full-length cyclic orbit flag code of type* $\mathbf{t} = (1, k_1, k_2, n-1)$ *on* $\mathbb{F}_{q^n}$*. Hence, its distance distribution is*

$$(\lambda_0^f, \lambda_1^f),$$

where

$$\lambda_1^f = \lambda_1^1 + \lambda_1^2 = \tfrac{q^{k_1}-1}{q-1}\tfrac{q^{k_1}-q}{q-1} + \tfrac{q^{n-k_2}-1}{q-1}\tfrac{q^{n-k_2}-q}{q-1}$$

$$\lambda_0^f = \lambda_0^1 - \lambda_1^2 = \lambda_0^2 - \lambda_1^1.$$

References

1. Alonso-González, C. and Navarro-Pérez, M.A.: A New Invariant for Cyclic Orbit Flag Codes. Linear and Multilinear Algebra, 1–24. https://doi.org/10.1080/03081087.2024.2304148
2. Alonso-González, C. and Navarro-Pérez, M.A.: Cyclic Orbit Flag Codes. Designs, Codes and Cryptography **89**, 2331-2356 (2021)

3. Alonso-González, C., Navarro-Pérez, M.A. and Soler-Escrivà, X.: An Orbital Construction of Optimum Distance Flag Codes Finite Fields and Their Applications **73**, 101861 (2021)
4. Alonso-González, C., Navarro-Pérez, M.A. and Soler-Escrivà, X.: Flag Codes from Planar Spreads in Network Coding. Finite Fields and Their Applications **68**, 101745 (2020)
5. Alonso-González, C., Navarro-Pérez, M.A. and Soler-Escrivà, X.: Optimum Distance Flag Codes from Spreads via Perfect Matchings in Graphs. Journal of Algebraic Combinatorics **54**, 1279–1297 (2021)
6. Ben-Sasson, E., Etzion, T., Gabizon, A. and Raviv, N.: Subspace Polynomials and Cyclic Subspace Codes. IEEE Transactions on Information Theory **62**, 1157–1165 (2016)
7. Chen, B. and Liu, H.: Constructions of Cyclic Constant Dimension Codes. Designs, Codes and Cryptography **86**(6), 1267–1279 (2018)
8. Etzion, T. and Vardy, A.: Error-Correcting Codes in Projective Spaces. IEEE Transactions on Information Theory **57**, 1165–1173 (2011)
9. Gluesing-Luerssen, H., Morrison, K. and Troha, C.: Cyclic Orbit Codes and Stabilizer Subfields. Advances in Mathematics of Communications **9**(2), 177-197 (2015)
10. Gluesing-Luerssen, H. and Lehmann, H.: Distance Distributions of Cyclic Orbit Codes. Designs, Codes and Cryptography **89**, 447–470 (2021)
11. Koetter, R. and Kschischang, F.: Coding for Errors and Erasures in Random Network Coding. IEEE Transactions on Information Theory **54**, 3579-3591 (2008)
12. Liebhold, D., Nebe, G. and Vázquez-Castro, A.: Network Coding with Flags. Designs, Codes and Cryptography **86**(2), 269-284 (2018)
13. Navarro-Pérez, M.A. and Soler-Escrivà, X.: Flag Codes of Maximum Distance and Constructions using Singer Groups. Finite Fields and their Applications **80**, 102011 (2022)
14. Otal, K., and Özbudak, F.: Cyclic Subspace Codes via Subspace Polynomials. Designs, Codes and Cryptography **85**(2), 191-204 (2017)
15. Roth, R.M., Raviv, N. and Tamo, I.: Construction of Sidon spaces with Applications to Coding. IEEE Transactions on Information Theory **64**(6), 4412–4422 (2018)
16. Trautmann, A.-L., Manganiello, F., Braun, M. and Rosenthal, J.: Cyclic Orbit Codes. IEEE Transactions on Information Theory **59**(11), 7386-7404 (2013)
17. Trautmann, A.-L., Manganiello, F. and Rosenthal, J.: Orbit Codes: A New Concept in the Area of Network Coding In: Proceedings of IEEE Information Theory Workshop, pp. 1–4. IEEE. Dublin, Ireland (2010)
18. Trautmann, A.-L. and Rosenthal, J.: Constructions of Constant Dimension Codes. In: M. Greferath et al. (eds.). Network Coding and Subspace Designs, E-Springer International Publishing AG, 2018, pp. 25–42.
19. Zullo, F.: Multi-orbit Cyclic Subspace Codes and Linear Sets. Finite Fields and their Applications **87**, 102153 (2023)

Extensible Decentralized Secret Sharing and Schnorr Signatures

M. Battagliola[1], R. Longo[2], and A. Meneghetti[3]

[1] University of Trento michele.battagliola@unitn.it
[2] Fondazione Bruno Kessler riccardolongomath@gmail.com
[3] University of Trento alessio.meneghetti@unitn.it

Abstract. Starting from links between Coding Theory and Secret Sharing Schemes, we develop an extensible and decentralized version of Shamir Secret Sharing, that allows the addition of new users after the initial share distribution.
On top of it we design a totally decentralized (t, n)-threshold Schnorr signature scheme that needs only t users online during the key generation phase, while the others join later. Using a classical game-based argument, we prove that if there is an adversary capable of forging the scheme with non-negligible probability, then we can build a forger for the centralized Schnorr scheme with non-negligible probability.

Keywords: Digital Signature, Threshold Cryptography, Secret Sharing, Maximum Distance Separable Code, Diffie-Hellman Assumption.

1 Introduction

Decentralized systems are slowly becoming a desirable alternative to centralized ones, due to the advantages of distributing the management of data, such as avoiding single-points-of-failures or the secure storage of crypto-assets. For them to become a viable alternative, it is necessary to use secure decentralized cryptographic schemes. In particular, digital signature schemes assume a central role in this setting, as hinted by the amount of recent works on multi-user schemes and threshold variants of signature protocols (see e.g. [1, 2, 5, 7]). In this work we present a completely decentralized Extensible and Verifiable Secret Sharing Scheme based on Shamir's one and we enhance it with the possibility of having *offline participants*, firstly introduced in [2]. In particular, our protocol allows for the addition of new parties after the initial secret sharing, a property that can be useful to enhance the resilience of the secret reconstruction, allowing for more protection against share loss.

Our Secret Sharing Scheme is a suitable algorithm for performing the Key Generation in many Discrete Logarithm based threshold signature, such as ECDSA [2], however we decided to focus our attention on the Schnorr's one, due to the increasing interest in this field. Our approach is similar to [3], [9] and [10]. However,

these three signatures work only in the (n, n) case, while ours works for an arbitrary threshold. More recently a general (t, n) Schnorr Signature was proposed, FROST [8], however their assumptions are not classical, while we only rely on the Decisional Diffie Hellman Assumption. Lastly, concurrently with this work, Sparkle [6] was proposed, that require only standard assumptions in the static case and it is very similar to our work. The two works were made independently, we discuss the small difference between them in Section 6.

2 Preliminaries

2.1 From MDS Codes to Secret Sharing

Let $\mathbb{F}_q$ be the finite field with q elements and let α be an agreed-upon primitive element of $\mathbb{F}_q$. Let $\{p^{(i)}\}_{i=1,\dots,\tau} \subseteq \mathbb{F}_q[x]$ be a set of τ polynomials of degree $t-1$, so $p^{(i)} = \sum_{k=0}^{t-1} p_k^{(i)} x^k$, where $p_k^{(i)} \in \mathbb{F}_q$ is the k-th coefficient of the polynomial $p^{(i)}$.

Let $p = \sum_{i=1}^{\tau} p^{(i)}$, with coefficients $p_k = \sum_{i=1}^{\tau} p_k^{(i)}$ for $k = 0, \dots, t-1$, and define $\beta_j = p(\alpha^j)$. Note that, if we define $\beta_{i,j} = p^{(i)}(\alpha^j)$ for $i \in \{1, \dots, \tau\}$ and $j \in \{1, \dots, q-1\}$, then we have that $\beta_j = \sum_{i=1}^{\tau} \beta_{i,j}$.

Definition 1. *Let $J = [j_1, \dots, j_n]$ be a list of $1 \leq n \leq q-1$ distinct integers in $\{1, \dots, q-1\}$. We define G_J as the $(t \times n)$ matrix:*

$$G_J = \left[\alpha^{j \cdot k}\right]_{k \in \{0,\dots,t-1\},\ j \in J}$$

If $n = 1$ then $J = [j]$ and we sometimes simply use G_j instead of $G_{[j]}$.

We remark that the matrix G_J is the generator matrix of a punctured $[n, t]_q$ Reed-Solomon code. A more general approach would be to use any $t \times n$ generator matrix of an MDS code, and, for instance, by using Extended Generalized Reed-Solomon codes we would obtain a broader set of acceptable parameters.

Since p has degree at most $t-1$, given any list $J \subseteq \{1, \dots, q-1\}$ of cardinality at least t, with the list of evaluations $[\beta_j]_{j \in J}$ it is possible to interpolate the polynomial p. That is, the coefficients p_k can be reconstructed and therefore the evaluation $p(\gamma)$ in any element $\gamma \in \mathbb{F}_q$ can be computed. More formally, we have these following propositions, whose proofs are trivial:

Proposition 1. *Let $J = [j_1, \dots, j_t]$ be a list of t distinct integers in $\{1, \dots, n\}$, and let G_J be the square matrix constructed as in Definition 1. Then:*

$$(p_0, \dots, p_{t-1}) = (\beta_{j_1}, \dots, \beta_{j_t}) \cdot G_J^{-1}.$$

Proposition 2. *Let h be any integer in $\{1, \dots, n\}$, let $J = [j_1, \dots, j_t]$ be a list of t distinct integers in $\{1, \dots, n\}$, and let e_ℓ be the ℓ-th element of the standard basis of $\mathbb{F}_q^t$. Then:*

$$\beta_h = \sum_{\ell=1}^{t} f(\beta_{j_\ell}, h, J, \ell),$$

where for any $\ell \in \{1, \ldots, t\}$ *we define the function* f *as:*

$$f(x, h, J, \ell) = x \cdot e_\ell G_J^{-1} G_h. \tag{1}$$

An interesting consequence of Proposition 2 is that t distinct shares are sufficient to compute any other share. However, observe that it is possible to obtain β_{j_ℓ} from $f(\beta_{j_\ell}, h, J, \ell)$, since both G_J and G_h can be easily computed even without knowing anything about the polynomials. This means that Proposition 2 should not be used directly to distribute new shares of a secret, in order to preserve the privacy of the old shares. A simple workaround is to split these secret values. Let $b_{h,J,\ell,k}$ be chosen at random in $\mathbb{F}_q$ for $k \in \{1, \ldots, t\} \setminus \{\ell\}$, and set $b_{h,J,\ell,\ell} = f(\beta_{j_\ell}, h, J, \ell) - \sum_{k=1, k\neq\ell}^{t} b_{h,J,\ell,k}$. If we define $b_{h,J,k} = \sum_{\ell=1}^{t} b_{h,J,\ell,k}$, then we have that:

$$\sum_{k=1}^{t} b_{h,J,k} = \sum_{k=1}^{t} \left(\sum_{\ell=1}^{t} b_{h,J,\ell,k} \right) = \sum_{\ell=1}^{t} \left(\sum_{k=1}^{t} b_{h,J,\ell,k} \right) = \sum_{\ell=1}^{t} f(\beta_{j_\ell}, h, J, \ell) = \beta_h \tag{2}$$

Note that the random values are completely canceled out only when summing all the $b_{h,J,k}$, this means that the values β_{j_ℓ} remain hidden, so exploiting this idea is a safe way to generate new shares.

2.2 Homomorphic Commitment Schemes

A commitment scheme [4] is composed by two algorithms:

- $\texttt{Com}(m, r)$**:** which given the message m to commit and some random value r (sometimes we will omit this randomness in our notation) outputs the commitment `KGC` and the decommitment `KGD`.
- $\texttt{Ver}(\texttt{KGC}, \texttt{KGD})$**:** which given a commitment and its decommitment outputs the committed message m if the verification succeeds, $\perp$ otherwise.

Besides the standard notion of *binding* and *hiding*, we need the following homomorphic property[4]: for all $m_0, m_1, z_0, z_1, \gamma \in \mathbb{F}_q$

$$\texttt{HCom}(m_0; z_0) \cdot \texttt{HCom}(m_1; z_1) = \texttt{HCom}(m_0 + m_1; z_0 + z_1),$$
$$\texttt{HCom}(m_0; z_0)^\gamma = \texttt{HCom}(\gamma \cdot m_0; \gamma \cdot z_0).$$

An example of suitable commitment for our work is Pedersen commitment [11], based on the difficulty of the discrete logarithm.

3 Extensible Decentralized Verifiable Secret Generation and Sharing Protocol

In this section we will give a brief description of our decentralized variant of the Verifiable Secret Sharing Scheme (VSSS) by Pedersen [11], which includes the feature of adding new users.

[4] We use the notation `HCom` when using an homomorphic commitment, while `Com` denotes any binding and hiding commitment scheme.

Let $P_1, \dots, P_n$ be n parties participating in the Secret Sharing Scheme, and let $t \leq n$ be the chosen threshold. We assume that q is a prime big enough that, given n polynomials of degree d sampled uniformly at random from $\mathbb{F}_q[x]$, the probability of their sum to be of degree $d' < d$ is negligible. Finally, let $\texttt{HCom}$ be an homomorphic commitment scheme as per Section 2.2.

3.1 Secret Generation

The distributed secret generation algorithm is carried out by the first $\tau \leq n$ parties $\{P_1, \dots, P_\tau\}$, and proceeds as follows:

1. Each P_i for $i \in \{1, \dots, \tau\}$ generates a secret polynomial $p^{(i)} \in \mathbb{F}_q[x]$ of degree $t-1$, by sampling the coefficients $p_k^{(i)}$ uniformly at random in $\mathbb{F}_q$.
2. The constant term p_0 of the summation polynomial p (see Section 2.1) is implicitly defined as the secret to be shared. Note that no single party P_i for any i knows this secret.
3. Each P_i samples another random polynomial $z^{(i)} \in \mathbb{F}_q[x]$ of degree $t-1$, and uses its coefficients to compute and publish the commitments to the coefficients of their secret polynomial $p^{(i)}$: $C_{0,i,k} = \texttt{HCom}\left(p_k^{(i)}; z_k^{(i)}\right)$.
4. After having received every single commitment $C_{0,j,k}$, for $j \in \{1, \dots, \tau\}$ and $k \in \{0, \dots, t-1\}$, each P_i sends to each P_j the evaluations $\beta_{i,j} = p^{(i)}(\alpha^j)$ and $\gamma_{i,j} = z^{(i)}(\alpha^j)$.
5. Each P_i for $i \in \{1, \dots, \tau\}$ sends the pair $(\beta_{i,j}, \gamma_{i,j})$ also to every party P_j for $j \in \{\tau+1, \dots, n\}$.
6. By exploiting the homomorphic properties of the commitment scheme, each P_i for $i \in \{1, \dots, n\}$ checks the values received against the published commitments:
$$\texttt{HCom}(\beta_{j,i}; \gamma_{j,i}) \stackrel{?}{=} \prod_{k=0}^{t-1} (C_{0,j,k})^{(\alpha^i)^k}, \tag{3}$$
for $j \in \{1, \dots, \tau\}$.
7. If all of these checks pass, each P_i sets its share of the newly generated secret as $\beta_i = \sum_{j=1}^{\tau} \beta_{j,i}$, and saves the checking value $\gamma_i = \sum_{j=1}^{\tau} \gamma_{j,i}$.

Observe that the τ parties involved in the secret generation algorithm are always capable of determining the secret p_0, even if $\tau < t$.

3.2 Secret Reconstruction

If $J \subseteq \{1, \dots, q\}$ is a list of t distinct indexes, then with the vector of shares $(\beta_j)_{j \in J}$ it is possible to reconstruct the secret p_0 as follows:

$$p_0 = (\beta_j)_{j \in J} \cdot G_J^{-1} \cdot e_1^T,$$

which is a direct consequence of Proposition 1. Let $\ell \in \{1, \ldots, t\}$ be the position of j inside the list J, note that the Shamir share β_j can be converted into an additive share ω_j:

$$\omega_j = \beta_j e_\ell \cdot G_J^{-1} \cdot e_1^T; \qquad p_0 = \sum_{j \in J} \omega_j. \tag{4}$$

3.3 Addition of New Parties

Let $J = [j_1, \ldots, j_t] \subseteq \{1, \ldots, n\}$ be a list of t distinct indexes. The parties $\{P_i\}_{i \in J}$ can collaborate to add the new party P_{n+1} (i.e. generate its share β_{n+1}) with the following algorithm:

1. Each P_{j_ℓ} for $\ell \in \{1, \ldots, t\}$ picks randomly $b_{n+1,J,\ell,k}, z_{n+1,J,\ell,k} \in \mathbb{F}_q$ for all $k \in \{1, \ldots, t\} \setminus \{\ell\}$, sets $b_{n+1,J,\ell,\ell} = f(\beta_{j_\ell}, n+1, J, \ell) - \sum_{k=1, k\neq\ell}^{t} b_{n+1,J,\ell,k}$, $z_{n+1,J,\ell,\ell} = f(\gamma_{j_\ell}, n+1, J, \ell) - \sum_{k=1,k\neq\ell}^{t} z_{n+1,J,\ell,k}$, where $f(x, n+1, J, \ell)$ is defined as in Equation (1).
2. Each P_{j_ℓ} publishes the commitments $C_{n+1,J,\ell,k} = \texttt{HCom}\,(b_{n+1,J,\ell,k}; z_{n+1,J,\ell,k})$ for $k \in \{1, \ldots, t\}$.
3. After having received every single commitment $C_{n+1,J,\ell,k}$, for $\ell, k \in \{1, \ldots, t\}$, each P_{j_ℓ} checks the coherence of these commitments with the ones published during the generation phase:

$$\prod_{k=1}^{t} C_{n+1,J,\ell,k} \stackrel{?}{=} \left(\prod_{k=0}^{t-1} \left(\prod_{j=1}^{\tau} C_{0,j,k} \right)^{(\alpha^{j_\ell})^k} \right)^{e_\ell G_J^{-1} G_{n+1}}, \tag{5}$$

for $\ell \in \{1, \ldots, t\}$ (G_J and G_{n+1} are defined as in Definition 1), and:

$$\prod_{k=1}^{t} \prod_{\ell=1}^{t} C_{n+1,J,\ell,k} \stackrel{?}{=} \prod_{k=0}^{t-1} \left(\prod_{j=1}^{\tau} C_{0,j,k} \right)^{(\alpha^{n+1})^k}. \tag{6}$$

If everything checks out, P_{j_ℓ} sends to each P_{j_k} the values $b_{n+1,J,\ell,k}$ and $z_{n+1,J,\ell,k}$, for $\ell, k \in \{1, \ldots, t\}$.
4. Each P_{j_ℓ} checks the consistency of the data received and the committed values:

$$\texttt{HCom}\,(b_{n+1,J,k,\ell}; z_{n+1,J,k,\ell}) \stackrel{?}{=} C_{n+1,J,k,\ell},$$

for $k \in \{1, \ldots, t\}$, sets $b_{n+1,J,\ell} = \sum_{k=1}^{t} b_{n+1,J,k,\ell}$, $z_{n+1,J,\ell} = \sum_{k=1}^{t} z_{n+1,J,k,\ell}$, and sends them to P_{n+1}.
5. P_{n+1} retrieves its share as: $\beta_{n+1} = \sum_{\ell=1}^{t} b_{n+1,J,\ell}$, and the checking value as: $\gamma_{n+1} = \sum_{\ell=1}^{t} z_{n+1,J,\ell}$. Then it checks their consistency with the commitments by verifying:

$$\texttt{HCom}\,(b_{n+1,J,\ell}; z_{n+1,J,\ell}) \stackrel{?}{=} \prod_{k=1}^{t} C_{n+1,J,k,\ell}, \tag{7}$$

for $\ell \in \{1, \ldots, t\}$, and Equations (5) and (6).

At the end of the procedure, P_{n+1} has its own secret values just like the other parties, so it can participate in the secret reconstruction or in the users addition.

3.4 Security of the Secret Sharing

In this section we prove correctness and security of the Secret Sharing Scheme described in Section 3.1, reducing it to the correctness and security of a centralized version, which are a direct consequence of the binding and hiding properties of the commitment scheme. For the correctness we refer to Definition 4.1 of [11] which includes the verifiability, for the security we refer to Theorem 4.4 of [11].

Definition 2 (Centralized Secret Sharing)**.** *The centralized version of the scheme described in Section 3.1 between a dealer D and players $P_1, \dots, P_n$ with threshold t of a secret $s \in \mathbb{F}_q$ proceeds as follows:*

1. *D chooses two random polynomials $p, z \in \mathbb{F}_q[x]$ of degree $t-1$ such that $p_0 = s$;*
2. *D computes and publishes $C_k = \texttt{HCom}(p_k, z_k)$ for $k = 0, \dots, t-1$;*
3. *D sends $\beta_j = p(\alpha^j)$ and $\gamma_j = z(\alpha^j)$ to P_j;*
4. *each P_j checks that their share is correct by verifying:*

$$\texttt{HCom}(\beta_j, \gamma_j) \stackrel{?}{=} \prod_{k=0}^{t-1} C_k^{(\alpha^j)^k} \tag{8}$$

The secret s can be reconstructed as usual by interpolating $\{\beta_j\}_{j \in J}$ where J is a set of at least t indexes.

To prove the security of the proposed secret sharing scheme we need two preliminary Lemmas:

Lemma 1 (Correctness). *If* HCom *is binding then the Secret Sharing Scheme of Definition 2 is correct. If* HCom *is perfectly binding then the Secret Sharing Scheme of Definition 2 is correct even if D has unbounded computational power.*

Lemma 2 (Security). *If* HCom *is hiding then the Secret Sharing Scheme of Definition 2 is secure. If* HCom *is perfectly hiding then the Secret Sharing Scheme of Definition 2 is secure even if the adversary has unbounded computational power.*

From Lemma 2 and Lemma 2 we have the first main result about the secret sharing scheme proposed:

Theorem 1. *If* HCom *is hiding, then the Secret Sharing Scheme described in Section 3.1 is secure.*

Proof. For the sake of simplicity we suppose that $\tau = t$ but the same proof can be adapted for an arbitrary τ.

Since HCom is hiding, then the Secret Sharing Scheme of Definition 2 is secure.

Let us suppose that the adversary controls $P_2, ..., P_t$. We show that after the Secret Generation (Section 3.1) it has no information about the secret p_0.

First of all, notice that $p_0^{(1)}$ is uniformly distributed, thus p_0 is uniformly distributed as well.

Then notice also that steps 1 to 6 are t independent executions of the Verifiable Secret Sharing scheme described in Definition 2 with n participants and threshold t, each having as dealer a different P_i, $i = 1, ..., t$, thus the adversary does not gain any information about $p_0^{(1)}$, the secret of the honest player. Moreover the last step does not involve any new message exchange, thus does not reveal anything. Hence, the adversary has no information about p_0. □

Now we need to prove the security of the Addition of New Parties. Informally, we need to show that an adversary controlling at most $t-1$ participants is not able to learn anything about the secret of the other parties or the secret itself.

Formally, we have the following definition:

Definition 3. *Let $S \subseteq \{1, ..., t, n+1\}$ be a set such that $|S| = t-1$, and let $\texttt{view}_S$ be the set of all the messages that parties in S see during the Addition of New Parties algorithm. We say that the Addition of New Parties is secure if and only if:*

$$\mathbb{P}(P_i \text{ has secret } \omega_i | \texttt{view}_S) = \mathbb{P}(P_i \text{ has secret } \omega_i),$$

for $i \notin S$. Moreover:

$$\mathbb{P}(\text{The shared secret is } p_0 | \texttt{view}_S) = \mathbb{P}(\text{The shared secret is } p_0).$$

Theorem 2. *If* `HCom` *is hiding, then the Addition of New Parties described in Section 3.3 is secure.*

Sketch. Initially we suppose that the adversary does not control P_{n+1}, but only $t-1$ out of the t parties which perform the protocol to add P_{n+1}. WLOG we can suppose that these parties are $P_1, \ldots, P_t$ and that the adversary controls $P_2, \ldots, P_t$.

Since `HCom` is hiding, then the Secret Sharing Scheme of Definition 2 is secure.

We can notice that Step 1 is a (t, t) additive secret sharing of $f(\beta_1, n+1, J, 1)$, with dealer P_1, verified with a homomorphic commitment. This is secure and does not leak any information about β_1 or β_{n+1}.

The following steps do not require any additional computation or communication involving the secret $b_{n+1,J,1,1}$, so the security is trivial.

In the same way we can prove the security when the adversary controls the added user P_{n+1} and $t-2$ among $P_1, \ldots, P_t$.

□

The checks in Equations (5) to (7) also allow to prove the following:

Theorem 3. *If* `HCom` *is binding, then the Addition of New Parties described in Section 3.3 is robust, i.e. an adversary controlling at most $t-1$ parties is not able to corrupt the protocol without being noticed.*

4 Threshold Schnorr Signature

In this section we describe a possible use case of our extensible Secret Sharing Scheme: a (t, n)-threshold variant of Schnorr's digital signature algorithm with offline participants. For our construction we need a group $\mathbb{G}$ of prime order q with generator g where the DLOG problem is assumed to be hard. Moreover the hardness of DLOG implies that the size of q is exponential in the security parameter, thus any practical application necessarily has a number of users $n \ll q$. Finally, we require that at least $\tau \geq t$ users are online for the setup, in the following we suppose there are exactly $\tau = t$ online parties in the key generation phase, namely $P_1, \ldots, P_t$.

For this signature protocol we exploit the Secret Sharing Scheme of Section 3, but note that we have to add some steps to the Key Generation algorithm because we have to publish the public key and check its consistency with the private key shares.

The protocol is divided into four algorithms, with a preliminary Setup Phase where all the common parameters are set:

1. **Key Generation** (Section 4.1): is performed by $P_1, \ldots, P_t$ to create the public key for the signature scheme and the private shares for themselves.
2. **Signature Algorithm** (Section 4.2): performed whenever any group of t parties wants to produce a signature.
3. **Participant Addition**(Section 4.3) performed by any group of at least t parties to create new shares for a new player.
4. **Verification**(Section 4.4) performed by the receiver.

From now on "P_i does something" means that all the parties involved in that phase perform the specified action.

4.1 Key generation

In this phase, the starting parties $P_1, \ldots, P_t$ produce a common public key $\mathcal{A}$ and each obtains a share of the corresponding private key.

1. Secret key generation and commitment:
   (a) P_i randomly picks $a_i \in \mathbb{Z}_q$ and sets $\mathcal{A}_i = g^{a_i}$;
   (b) P_i randomly picks a polynomial $p^{(i)}$ of degree $t-1$ such that $p^{(i)}(0) = a_i$.
   (c) P_i computes $[\mathtt{KGC}_i, \mathtt{KGD}_i] = \mathtt{Com}(\mathcal{A}_i)$;
   (d) P_i computes (as per Section 3) $\beta_{i,j} = p^{(i)}(\alpha^j)$, $\gamma_{i,j} = z^{(i)}(\alpha^j)$, $C_{0,i,k}$;
   (e) P_i computes $[\mathtt{KCC}_{i,j}, \mathtt{KCD}_{i,j}] = \mathtt{Com}(g^{\beta_{i,j}})$ then publishes the commitments $\mathtt{KGC}_i$, $C_{0,i,k}$, $\mathtt{KCC}_{i,j}$;
2. Shares verification and private key computation:
   (a) once all the commitments have been published, P_i publishes the decommitments $\mathtt{KGD}_i$, $\mathtt{KCD}_{i,j}$;
   (b) P_i gets $\mathcal{A}_j$, $g^{\beta_{j,\ell}}$ for $1 \leq j \leq t, i \neq j$, $1 \leq \ell < t$ and checks their consistency by interpolating at the exponent;

   (c) P_i proves in ZK the knowledge of a_i using Schnorr's protocol (this proof can either be interactive or non interactive. In the second case the proofs are checked as soon as they are received. In either cases, if a party fails the ZKP the protocol aborts).
   (d) P_i sends $(\beta_{i,j}, \gamma_{i,j})$ to player P_j;
   (e) P_i checks the integrity and consistency of the shards $\beta_{j,i}$ as in Section 3 and also with the values $g^{\beta_{j,i}}$;
3. P_i computes its private key $\beta_i = \sum_{j=1}^{t} \beta_{j,i}$.
4. The public key is $\mathcal{A} = \prod_{i=1}^{t} \mathcal{A}_i$. Implicitly we set $\sum_{i=1}^{t} a_i = a$.

The public values $\mathcal{A}$, $C_{0,i,k}$ for $i \in \{1, \ldots, t\}, k \in \{0, \ldots, t-1\}$ are saved in a public register `pub_reg`.

4.2 Signature Algorithm

This algorithm is used when a set J of at least t players agrees to sign a message M. The protocol proceeds as follows.

1. Generation of r:
   (a) P_i randomly chooses $k_i \in \mathbb{Z}_q$ and computes $r_i = g^{k_i}$;
   (b) P_i computes $[\mathtt{NGC}_i, \mathtt{NGD}_i] = \mathtt{Com}(r_i)$ and sends $\mathtt{NGC}_i$;
   (c) once every $\mathtt{NGC}_j$ for $j \in J$ has been received, P_i sends $\mathtt{NGD}_i$;
   (d) P_i computes $r_j = \mathtt{Ver}([\mathtt{NGC}_j, \mathtt{NGD}_j])$ for each $j \in J$;
   (e) P_i computes $r = \prod_{j \in J} r_j$.
2. Generation of s:
   (a) P_i compute ω_i such that $\sum_{j \in J} \omega_j = a$, as in Equation (4);
   (b) P_i computes $e = H(r||M)$ and $s_i = k_i - \omega_i e$;
   (c) P_i computes $[\mathtt{SGC}_i, \mathtt{SGD}_i] = \mathtt{Com}(s_i)$ and sends $\mathtt{SGC}_i$;
   (d) once every $\mathtt{SGC}_j$ for $j \in J$ has been received, P_i sends $\mathtt{SGD}_i$;
   (e) P_i computes $s_j = \mathtt{Ver}([\mathtt{SGC}_j, \mathtt{SGD}_j])$ for each $j \in J$;
   (f) P_i computes $s = \sum_{j \in J} s_j$.
3. P_i computes $r_v = g^s \mathcal{A}^e$ and checks that $H(r_v||M) = e$.

The output signature is (s, e). If a check fails, the protocol aborts.

4.3 Participant Addition

This protocol allows any set J of at least t users to add a new user P_m to the protocol. After the protocol P_m will have the same powers (i.e. can sign and add new users) as the other users. The protocol works as follows:

1. the players in J share all the public data with P_m;
2. the players in J participate in the Participant Addition Protocol of Section 3.3, publishing $C_{m,J,k,\ell}$, for $k, \ell \in \{1, \ldots, t\}$;
3. P_{j_ℓ} sends to P_m $b_{m,J,\ell}$, $z_{m,J,\ell}$;
4. using the homomorphic commitments received in Steps 1 and 2, P_m performs all the checks described in Section 3.3 and computes its share $\beta_m \sum_{\ell=1}^{t} b_{m,J,\ell}$.

4.4 Verification

The verification protocol is the same of the centralized one. In particular, to verify a signature (s, e) for a received message M, the receiver computes $r_v = g^s \mathcal{A}^e$ and checks that $H(r_v||M) = e$. If the checks fails the receiver reject the signature, otherwise it accepts.

5 Security Proof

It is possible to prove that the described protocol is unforgeable. Namely:

Definition 4 (Unforgeability)**.** *A (t, n)-threshold signature scheme is unforgeable if no malicious adversary who corrupts at most $t - 1$ players can produce the signature on a new message m with non negligible probability, given the view of the threshold sign on input messages $m_1, \ldots, m_Q$ (adaptively chosen by the adversary), as well as the signatures on those messages.*

The unforgeability of our protocol is formally stated in the following theorem:

Theorem 4. *Assuming that the Schnorr signature scheme instantiated on the group $\mathbb{G}$ of prime order q with the hash function H is unforgeable,* `Com`, `Ver` *is a non-malleable commitment scheme, and that the Decisional Diffie-Hellman Assumption holds, then our threshold protocol is unforgeable.*

6 Comparison with Concurrent Works

Concurrently with our work, Crites, Komlo and Maller proposed Sparkle [6], a new $(t, n)-$threshold Shnorr Signature. The two protocols are very similar and have almost the exact structure, in particular we have the following correspondences with our Signature Algorithm of Section 4.2:

- the `Sign` algorithm of Sparkle correspond to our steps from 1a to 1b;
- the `Sign`$'$ algorithm of Sparkle corresponds to our steps from 1c to 1e;
- the `Sign`$''$ algorithm of Sparkle correspond to our steps from 2a to 2d;
- the `Combine` algorithm of Sparkle correspond to our steps from 2e to 2f.

In [6] there is a deep security analysis focused on adaptive corruption of parties after the key generation. However, a key difference between our and their proofs is that Sparkle's security proof does not allow the adversary to participate in the key generation phase, and thus the adversary is not able to choose its secret key freely. One may see our analysis as covering adversaries that participate in the key generation and Sparkle's analysis as covering adversaries that corrupt parties afterwards, thus the two somewhat complete each other.

Acknowledgements

This work was created with the co-financing of the European Union FSE-REACT-EU, PON Research and Innovation 2014-2020 DM1062/2021 and has been partially supported by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU. Michele Battagliola acknowledges support from TIM S.p.A. through the PhD scholarship. Alessio Meneghetti acknowledges support from Ripple's University Blockchain Research Initiative. The authors are members of the INdAM Research Group GNSAGA. The core of this work was partially presented as a poster at the conference CANS 2022, held in Abu Dhabi from 13 to 16 November 2022.

References

[1] M. Battagliola, A. Galli, R. Longo, A. Meneghetti, et al. "A Provably-Unforgeable Threshold Schnorr Signature With an Offline Recovery Party". In: *DLT@ITASEC, CEUR Workshop Proceedings*. Vol. 3166. 2022, pp. 60–76. URL: https://api.semanticscholar.org/CorpusID:251026965.

[2] M. Battagliola, R. Longo, A. Meneghetti, and M. Sala. "Threshold ECDSA with an Offline Recovery Party". In: *Mediterranean Journal of Mathematics* 19.1 (2022), pp. 1–29.

[3] D. Boneh, M. Drijvers, and G. Neven. "Compact Multi-signatures for Smaller Blockchains". In: *Advances in Cryptology – ASIACRYPT 2018*. Ed. by T. Peyrin and S. Galbraith. Cham: Springer International Publishing, 2018, pp. 435–464. ISBN: 978-3-030-03329-3.

[4] G. Brassard, D. Chaum, and C. Crépeau. "Minimum disclosure proofs of knowledge". In: *Journal of computer and system sciences* 37.2 (1988), pp. 156–189.

[5] R. Canetti, R. Gennaro, S. Goldfeder, N. Makriyannis, and U. Peled. "UC Non-Interactive, Proactive, Threshold ECDSA with Identifiable Aborts". In: *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*. CCS '20. Virtual Event, USA: Association for Computing Machinery, 2020, pp. 1769–1787. ISBN: 9781450370899. DOI: 10.1145/3372297.3423367. URL: https://doi.org/10.1145/3372297.3423367.

[6] E. Crites, C. Komlo, and M. Maller. "Fully Adaptive Schnorr Threshold Signatures". In: *Advances in Cryptology – CRYPTO 2023*. Ed. by H. Handschuh and A. Lysyanskaya. Cham: Springer Nature Switzerland, 2023, pp. 678–709. ISBN: 978-3-031-38557-5.

[7] R. Gennaro and S. Goldfeder. "Fast Multiparty Threshold ECDSA with Fast Trustless Setup". In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1179–1194. ISBN: 9781450356930. DOI: 10.1145/3243734.3243859. URL: https://doi.org/10.1145/3243734.3243859.

[8] C. Komlo and I. Goldberg. "FROST: Flexible Round-Optimized Schnorr Threshold Signatures". In: *Selected Areas in Cryptography*. Ed. by O. Dunkelman, M. J. Jacobson Jr., and C. O'Flynn. Cham: Springer International Publishing, 2021, pp. 34–65. ISBN: 978-3-030-81652-0.

[9] G. Maxwell, A. Poelstra, Y. Seurin, and P. Wuille. "Simple Schnorr multi-signatures with applications to Bitcoin". In: *Designs, Codes and Cryptography* 87 (Sept. 2019). DOI: 10.1007/s10623-019-00608-x.

[10] A. Nicolosi, M. N. Krohn, Y. Dodis, and D. Mazières. "Proactive Two-Party Signatures for User Authentication". In: *Proceedings of the Network and Distributed System Security Symposium, NDSS 2003, San Diego, California, USA*. The Internet Society, 2003. URL: https://www.ndss-symposium.org/ndss2003/proactive-two-party-signatures-user-authentication/.

[11] T. P. Pedersen. "Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing". In: *Advances in Cryptology — CRYPTO '91*. Ed. by J. Feigenbaum. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 129–140.

On Functions of $\mathbb{F}_{2^{2t}}$ mapping Cosets of $\mathbb{F}_{2^t}^*$ to Cosets of $\mathbb{F}_{2^t}^*$

Jules Baudrin, Anne Canteaut, and Léo Perrin

Inria, Paris, France
{jules.baudrin, anne.canteaut, leo.perrin}@inria.fr

Abstract. The punctured finite field $\mathbb{F}_{2^{2t}}^*$ can be partitioned into multiplicative cosets of $\mathbb{F}_{2^t}^*$. In this paper, we investigate functions mapping these cosets either to multiplicative cosets of $\mathbb{F}_{2^t}^*$ or to additive cosets of $\mathbb{F}_{2^t}^*$. First, we analyze the "subspace property", a well-known feature of the so-called Kim mapping, that corresponds to functions preserving the partition into multiplicative cosets. As a particular case, we study functions whose restriction to each coset coincides with a power permutation. Those correspond to a particular case of (generalized) cyclotomic mappings. We also identify conditions for such functions to be quadratic and APN, and we study the structure of their Walsh zeroes. There is only one function in the literature mapping the partition into multiplicative cosets to that into additive cosets: the S-box of the last symmetric primitives standardized in Russia. We shed some more light on its design process, and we derive, for any t, permutations of $\mathbb{F}_{2^{2t}}$ with a similar structure, and a linearity smaller than 2^{t+2}.

Keywords: Subspace property · Cyclotomic Mappings · APN · S-box

1 Introduction

Finding functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$ with optimal cryptographic properties is a major open problem when n is even. In this case, even the optimal values for the differential uniformity and for the linearity of a bijective mapping are unknown. For instance, the existence of APN permutations on $\mathbb{F}_{2^n}$, when n is even and greater than 6, has been open for 15 years, since an APN permutation for $n = 6$ was exhibited by Dillon *et al.* in 2009 [6]. For constructing such functions, a natural idea is to decompose $\mathbb{F}_{2^n}$, $n = 2t$, into the $(2^t + 1)$ multiplicative cosets of $\mathbb{F}_{2^t}$. This spread is actually used in several constructions, including the bent functions of *partial-spread type* introduced much earlier by Dillon [16].

Our work then focuses on two types of functions: functions F mapping each multiplicative coset $\gamma\mathbb{F}_{2^t}^*$ to a multiplicative coset $F(\gamma)\mathbb{F}_{2^t}^*$, and bijections F mapping each multiplicative coset $\gamma\mathbb{F}_{2^t}^*$, except $\mathbb{F}_{2^t}^*$, to an additive coset $\alpha + \mathbb{F}_{2^t}^*$. The first type, studied in Section 2, corresponds to the functions satisfying the *subspace property* defined by Dillon *et al.* [6], and includes the so-called Kim mapping for $n = 6$. The second type, studied in Section 3, includes the Sbox of the recent Russian standard primitives Streebog and Kuznyechik [26]. We leave

aside bijections mapping the additive cosets of $\mathbb{F}_{2^t}$ into additive cosets, as those have been studied by several authors, e.g. to design backdoored primitives [25,1].

1.1 Preliminaries

For two sets A and B, we denote by $A \sqcup B$ the *disjoint union* of A and B. Regarding integers, if $u, v \in \mathbb{Z}$ with $u \leq v$, the set $\{u, u+1, \cdots, v\}$ is denoted by $[\![u, v]\!]$. The Hamming weight is denoted by $\mathrm{wt}(.)$.

Let K_0 be a field of size q and K_1 be an extension of K_0 of degree d. The trace function from K_1 to K_0 is defined as $\mathrm{Tr}_{K_1/K_0} : x \mapsto \sum_{i=0}^{d-1} x^{q^i}$. Its standard properties can be found for instance in [23, Theorems 2.23 & 2.24 p.56].

Though famous, the Walsh transform has many variants. We use the following one.

Definition 1 (Walsh coefficients and linearity). *Let K be a field over $\mathbb{F}_2$. Let $F\colon K \to K$ and $\alpha, \beta \in K$. The (α, β)-Walsh coefficient of F is defined as:*

$$W_{K,F}(\alpha, \beta) := \sum_{x \in K} (-1)^{\mathrm{Tr}_{K/\mathbb{F}_2}(\alpha x + \beta F(x))} ;$$

and the linearity *of F is the maximum $\mathcal{L}(F) := \max_{\alpha \neq 0, \beta \in K} |W_{K,F}(\alpha, \beta)|$.*

Let $t \in \mathbb{N} \setminus \{0\}$, $n = 2t$. In the following, we consider only the finite field $\mathbb{F}_{2^n}$ and its subfields. When there is no ambiguity on the cardinality, the field with 2^n elements and its subfield with 2^t elements are denoted by $\mathbb{L} := \mathbb{F}_{2^n}$, and $\mathbb{F} := \mathbb{F}_{2^t} \subset \mathbb{L}$. We reserve the letter λ (resp. φ) for elements of $\mathbb{L}$ (resp. $\mathbb{F}$).

As $\mathbb{F}^*$ is a subgroup of $\mathbb{L}^*$, we can naturally partition $\mathbb{L}^*$ as an union of cosets that is, $\mathbb{L}^* = \bigsqcup_{\gamma \in \Gamma} \gamma\mathbb{F}^*$, where Γ is any complete system of distinct representatives. In the following, we shorten complete system of distinct representatives into *system of directions* where *direction* refers to the geometrical interpretation of such a partition. We also reserve the notation Γ for systems of directions and γ for elements of Γ. By construction, $\Gamma \cap \mathbb{F}^*$ contains a single element, which is denoted by γ°, that is, $\{\gamma^\circ\} := \Gamma \cap \mathbb{F}^*$. We also define Γ° as $\Gamma^\circ := \Gamma \setminus \{\gamma^\circ\}$.

Because $2^t - 1$ and $2^t + 1$ are two coprime divisors of $2^{2t} - 1$, the Chinese Remainder Theorem suggests the subgroup $\mathbb{G}$ of order $2^t + 1$ as a natural system of directions. This choice, known as the *polar coordinate system* or *polar representation* (see for instance [11, page 191]), has especially been studied to design Boolean functions with notable properties [12,13,24,31]. More generally, any *set* S with $2^t + 1$ elements is a system of directions if and only if it satisfies $\Theta(S) = \mathbb{G}$ where $\Theta\colon \mathbb{L} \to \mathbb{L}, \quad x \mapsto x^{2^t - 1}$. In [19], Göloğlu introduces the trace-0/trace-1 representation which corresponds to $\Gamma = \{1\} \cup \mathrm{Tr}^{-1}_{\mathbb{L}/\mathbb{F}}(\{1\})$.

2 Subspace Property and the Kim Mapping

The so-called *Kim mapping* [6] is a quadratic APN function of 6 variables. A remarkable property is that it is CCZ-equivalent to a permutation. The Kim

mapping is defined by $\kappa\colon \mathbb{F}_{64} \to \mathbb{F}_{64} \quad x \mapsto x^3 + x^{10} + ux^{24}$, where u is a root of the primitive polynomial $X^6 + X^4 + X^3 + X + 1$.

As first introduced and studied in [6], κ is particularly structured and interacts with the multiplicative decomposition of $\mathbb{F}_{64}$. This is highlighted by the so-called subspace property defined by Dillon *et al.* as follows.

Definition 2 (Subspace property [6]). *$F : \mathbb{L} \to \mathbb{L}$ satisfies the subspace property if, for all $\lambda \in \mathbb{L}$, $F(\lambda\mathbb{F}) = F(\lambda)\mathbb{F}$.*

The subspace property exactly corresponds to function mapping each multiplicative coset onto another one. Indeed if $F(\lambda) \in F(\lambda\mathbb{F}) = \gamma\mathbb{F}$, so $\gamma\mathbb{F} = F(\lambda)\mathbb{F}$. The image of any function satisfying the subspace property is then a partial spread since it is a union of multiplicative cosets. When F satisfies the subspace property we introduce, for any $\lambda \neq 0$, the function $F_\lambda\colon \mathbb{F} \to \mathbb{F}, \varphi \mapsto \frac{F(\lambda\varphi)}{F(\lambda)}$. The subspace property is a purely set-theoretic property. Indeed, F satisfies the subspace property if and only if, for any $\lambda \neq 0$, F_λ is well-defined and bijective. In the following, we focus on the particular case where any F_λ is a power mapping.

2.1 Cyclotomic mappings satisfying the subspace property

In previous works, functions whose restriction to each multiplicative coset is a power mapping are called (generalized) cyclotomic mapping.

Definition 3 ((Generalized) cyclotomic mapping [5]). *Let $\ell \mid 2^n - 1$ and $\mathbb{G}$ be the subgroup of order $|\mathbb{G}| = \frac{2^n-1}{\ell}$* (i.e. *of index ℓ). A generalized cyclotomic mapping of index ℓ of $\mathbb{L}$ is a function $F\colon \mathbb{L} \to \mathbb{L}$ which satisfies:*

$$\forall \lambda \in \mathbb{L}, \exists\ d_\lambda \in \mathbb{N},\ \forall\ x \in \mathbb{G}, F(\lambda x) = F(\lambda)x^{d_\lambda}.$$

When $\mathbb{G}$ is clear from context, we rather call it generalized cyclotomic mapping over $\mathbb{G}$. *If there exists d such that $d_\lambda = d$ for all λ, such mapping is called* cyclotomic mapping [29] of order d over $\mathbb{G}$ *(or of index ℓ).*

Note that if $F(\lambda) \neq 0$, the value of $d_\lambda \bmod |\mathbb{G}|$ only depends on the coset of λ. Generalized cyclotomic mappings that satisfy the subspace property can therefore be easily characterized.

Proposition 1. *A generalized cyclotomic mapping $F : \mathbb{L} \to \mathbb{L}$ over $\mathbb{F}^*$ satisfies the subspace property if and only if for any $\lambda \in \mathbb{L}$, $\gcd(d_\lambda, 2^t - 1) = 1$ (where d_λ are defined as in Definition 3).*

Corollary 1. *Let $F : \mathbb{L} \to \mathbb{L}$ be a generalized cyclotomic mapping that satisfies the subspace property. If F is APN, then all exponents d_λ, $\lambda \neq 0$ defined as in as Definition 3 are such that $x \mapsto x^{d_\lambda}$ is APN on $\mathbb{F}$. Most notably, no APN generalized cyclotomic mapping that satisfies the subspace property exist when n is a multiple of 4.*

The first statement remains true for any function satisfying the subspace property. The second one is an immediate corollary of the inexistence of APN bijective power mappings over $\mathbb{F}_{2^k}$ for even k.

Not only does the Kim mapping satisfies the subspace property: it is, above all, a cyclotomic mapping of order 3 over $\mathbb{F}^*$. Indeed, for $\lambda \in \mathbb{L}^*$ and $\varphi \in \mathbb{F}^*$, $\kappa(\lambda\varphi) = \varphi^3(\lambda^3 + \lambda^{10} + u\lambda^{24}) = \kappa(\lambda)\varphi^3$. Because $\gcd(3, |\mathbb{F}^*|) = \gcd(3, 7) = 1$, the subspace property is a consequence of its cyclotomicity.

Cyclotomic mapping over $\mathbb{F}^*$ are of the form $x^d P(x^{2^t-1})$ for some polynomial P, and as such, they are a particular type of Wan-Lidl polynomials. The latter ones have been studied by several authors but mainly in the bijective case, e.g. [4,15,21,28,30]. Most notably, Chen and Coulter recently investigated their differential uniformity but their results do not provide any relevant information for our parameters in characteristic 2.

Obviously, the polynomial forms of these functions can be easily characterized.

Proposition 2. *[19, p.264], [29, Lemma 1] Let $\mathbb{G}$ be a subgroup of $\mathbb{L}^*$. Let $F\colon \mathbb{L} \to \mathbb{L}, x \mapsto \sum_{i=0}^{2^{2t}-1} a_i x^i$, where $a_i \in \mathbb{L}$ for any i. Then F is a cyclotomic mapping of order d over $\mathbb{G}$ if and only if $a_i \neq 0 \implies i \equiv d \mod |G|$.*

Proposition 2 can then be refined to quadratic cyclotomic functions of order d over $\mathbb{F}^*$. In that case $\mathrm{wt}(d) \leq 2$, but the case $\mathrm{wt}(d) = 1$ is of low interest since such a cyclotomic mapping cannot be APN. When $\mathrm{wt}(d) = 2$, any such $x^d P(x^{2^t-1})$ is linearly-equivalent to $x^{2^i+1} Q(x^{2^t-1})$ for some i and Q. From now on, we therefore only focus on Gold exponents $d = 2^i+1$. This family is described in the following corollary.

Corollary 2. *[19, p.264] Let $F\colon \mathbb{L} \to \mathbb{L}$ be a quadratic cyclotomic mapping of order $2^i + 1$ and index $2^t + 1$ where $i < t$. Then its univariate representation contains at most 4 monomials, with exponents $2^i+1, 2^i+2^t, 2^{t+i}+1$ and $2^{t+i}+2^t$.*

The family described in Corollary 2 has already attracted attention. For instance, it includes the APN trinomials exhibited by Göloğlu for $n \equiv 0 \bmod 4$ in [19], which have been latter proved affine equivalent to the Gold power mapping $x^{2^{t-i}+1}$ [7, Section 4]. Despite a reference to the subspace property in [19], Corollary 1 implies that none of those trinomials satisfy the original subspace property defined in Definition 2. The subspace property as defined in [19] actually corresponds to the definition of cyclotomic mapping with a Gold exponent (that is not necessarily bijective).

Carlet [10, § 3.7] also focuses on a sub-family of the polynomials in Corollary 2, namely those with $\gcd(i, n) = 1$ and such that the coefficient of x^{2^i+1} is nonzero. When $i = 1$ (i.e. $d = 3$) and $t > 3$, those functions are called *Kim-type functions* in [14]. APN Kim-type functions have been completely characterized in [22] and proved CCZ-equivalent to a Gold function, x^3 or $x^{2^{n/2-1}+1}$ over $\mathbb{L}$ in [14]. But, it is worth noting that the more general case of APN quadratic cyclotomic mappings of order $d = 2^i + 1$, for any i with $\gcd(i, t) = 1$ is still

open. Again, even if the different terminologies used in these works may appear confusing[1], the functions from [10, § 3.7], and in particular Kim-type functions, only satisfy the original subspace property when t is odd.

According to the authors of [6], the cyclotomicity explains some of the "simplicity" of the Kim function and part of its CCZ-equivalence to a permutation. On the other hand, all power mappings, including the APN Gold function $x \mapsto x^3$ over $\mathbb{F}_{64}$, are obviously cyclotomic mappings (because $(\lambda\varphi)^3 = \lambda^3\varphi^3$) but do not share all the properties of κ. In the following, we thus continue investigating cyclotomicity, the subspace property and the functions satisfying them.

2.2 Properties of the functions satisfying the subspace property

Proposition 3. *Let Γ be a system of directions. Let $F\colon \mathbb{L} \to \mathbb{L}$ be a function satisfying the subspace property. Let $\alpha, \beta \in \mathbb{L}$, then:*

$$W_{\mathbb{L},F}(\alpha,\beta) = -2^t + \sum_{\gamma\in\Gamma} W_{\mathbb{F},F_\gamma}\left(\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\alpha\gamma), \mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\beta F(\gamma))\right).$$

The proof is a direct computation.

For any $\lambda \in \mathbb{L}$, the size of the preimage of $\lambda\mathbb{F}$ (divided by 2^t) is denoted by $N_{F,\lambda} := \frac{|F^{-1}(\lambda\mathbb{F})|}{2^t}$. Given any system of directions Γ, this quantity is also equal to $N_{F,\lambda} = |\{\gamma \in \Gamma \text{ s.t. } F(\gamma) \in \lambda\mathbb{F}\}|$.

If there exists a nonzero $\gamma_0 \in \Gamma$ such that $F(\gamma_0) = 0$, then γ_0 belongs to $\{\gamma \in \Gamma \text{ s.t. } F(\gamma) \in \lambda\mathbb{F}\}$ for all $\lambda \in \mathbb{L}$. It follows that $\sum_{\lambda\in\Gamma} N_{F,\lambda} = 2^t + 1 + 2^t N_{F,0}$. However, the case $N_{F,0} > 0$ is of low interest in our context since any such function has differential uniformity at least 2^t, because all $(\gamma_0\varphi), \varphi \in \mathbb{F}$ satisfy $F(\gamma_0\varphi + \gamma_0) + F(\gamma_0\varphi) = 0$. This situation occurs for instance when F is defined by the same polynomial as the Kim mapping, $x \mapsto x^3 + x^{10} + u'x^{24}$, but when u' is a root of the primitive polynomial $X^6 + X^5 + X^2 + X + 1$. Indeed, $F(u'^2) = 0$.

Proposition 4. *If $F : \mathbb{L} \to \mathbb{L}$ is a function satisfying the subspace property and $\beta \in \mathbb{L}^*$, then $W_{\mathbb{L},F}(0,\beta) = 2^t(N_{F,\beta^{-1}} - 1)$.*
Most notably, $\mathcal{L}(F) \geq 2^t\left(\max_{\lambda\in\mathbb{L}^*}(N_{F,\lambda}) - 1\right)$.

Proposition 5 (Symmetries of the Walsh coefficients). *Let $F\colon \mathbb{L} \to \mathbb{L}$. Let $G\colon \mathbb{F} \to \mathbb{F}$. F satisfies the subspace property with $F_\lambda = G$ for any λ if and only if it verifies :*

$$\forall\alpha,\beta \in \mathbb{L},\ \forall\varphi \in \mathbb{F}^*, \quad W_{\mathbb{L},F}(\alpha,\beta G(\varphi)) = W_{\mathbb{L},F}(\alpha\varphi^{-1},\beta).$$

Again, the proof is a straight-forward computation. Propositions 4 and 5 respectively explains the first row of Fig. 1, and its the "structured-square" pattern.

[1] If they are APN, the functions from [10, § 3.7] are named *generalized Kim* functions.

2.3 Quadratic functions satisfying the subspace property

Proposition 4 becomes particularly interesting when looking at quadratic functions. Indeed, when F is quadratic, all its components are plateaued, i.e., there exists $\ell_\beta \in \mathbb{N}$ such that for any $\alpha \in \mathbb{L}$, $W_{\mathbb{L},F}(\alpha,\beta) \in \{0, \pm 2^{\ell_\beta}\}$. Proposition 4 then provides the linearity of all components F_β except those such that $N_{F,\beta^{-1}} = 1$. Therefore, it can be refined as follows.

Corollary 3. *Let $F : \mathbb{L} \to \mathbb{L}$ be a quadratic function satisfying the subspace property, and Γ be a system of directions. Then,*

$$\mathcal{L}(F) = \max\left(2^t\left(\max_{\lambda\in\Gamma}(N_{F,\lambda}) - 1\right); \max\{\mathcal{L}(F_\beta), \beta \in \Gamma \text{ s.t. } N_{F,\beta^{-1}} = 1\}\right) .$$

More importantly, we can derive a simple necessary condition for a quadratic function satisfying the subspace property to be APN.

Theorem 1. *Let $F : \mathbb{L} \to \mathbb{L}$ be a quadratic function satisfying the subspace property. Let $\mathcal{N}_i = |\{\gamma \in \Gamma : N_{F,\gamma} = i\}|$ where Γ is a system of directions. If F is APN, then $\mathcal{N}_0 + \mathcal{N}_2 \geq 2(2^t+1)/3$. Moreover, if $\mathcal{L}(f) = 2^{t+1}$, then F is APN if and only if $\mathcal{N}_0 + \mathcal{N}_2 = 2(2^t+1)/3$, which can only occur when t is odd.*

Proof. As F is quadratic, all its components are plateaued. Moreover, the β-component is bent if and only if $W_{\mathbb{L},F}(0,\beta) = \pm 2^t$, which equivalently means that $N_{F,\beta^{-1}} \in \{0, 2\}$. The result is then deduced from the fact that, if a function F of $2t$ variables with plateaued components is APN, then it has at least $2(2^t+1)/3$ bent components. Conversely, if F has $2(2^t+1)/3$ bent components and $\mathcal{L}(F) = 2^{t+1}$, then F is APN [2, Coro. 3]. □

Most notably, Theorem 1 provides an easy way to check that a given F is not APN by evaluating it at (2^t+1) points only. For instance, if we consider the same polynomial expression as the Kim mapping, $x \mapsto x^3 + x^{10} + u'x^{24}$, with u' a root of any primitive polynomial different from $X^6 + X^4 + X^3 + X + 1$, then we can easily check that $\mathcal{N}_0 + \mathcal{N}_2 \leq 4$, implying that the function is not APN. On the other hand, it is currently not known how to assess in an effective way whether F is a quadratic function satisfying the subspace property. Because of Corollary 2, this technique is however well-suited for quadratic cyclotomic mappings.

For $t = 3$, we computationally exhausted all quadratic cyclotomic mappings of order 3 and kept only those for which $N_{F,0} > 0$. Among them, all functions satisfying $\mathcal{N}_0 + \mathcal{N}_2 \geq 6$ are APN. All the obtained APN functions are CCZ-equivalent either to κ or to $x \mapsto x^3$.

2.4 Cyclotomic mappings and 3-to-1-ness

A simple particular case is the situation where $\mathcal{N}_3 = (2^t+1)/3$ and $\mathcal{N}_0 = 2(2^t+1)/3$ like in the case of Gold functions $F(x) = x^{2^\ell+1}$ with $\gcd(\ell, t) = 1$. Indeed, we can directly deduce from Theorem 1 that F is APN. This case has already been studied in a more general context [8,20] since the corresponding functions F are *almost 3-to-1*, i.e. F is *3-to-1* on $\mathbb{L}^*$. Most notably, the following proposition is a direct consequence of [20].

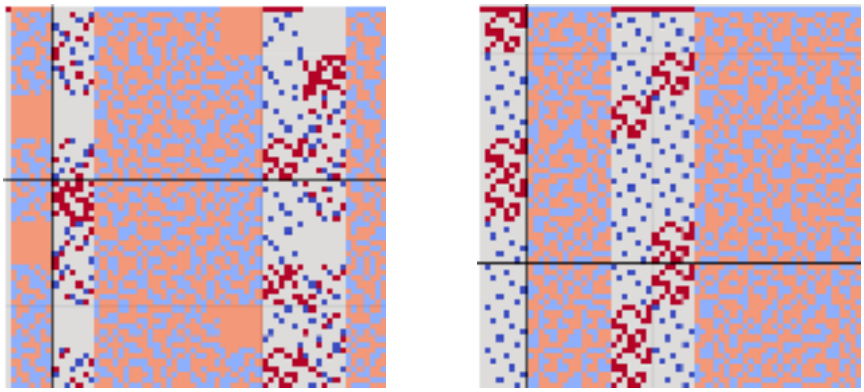


Fig. 1: *The Walsh coefficients (LAT) of κ (left) and $x \mapsto x^3$ (right). Dark blue is a low negative value, dark red a high positive value.*

Proposition 6. *Let F be a 3-divisible cyclotomic mapping of order d over $\mathbb{F}^*$ with plateaued components. The following properties are equivalent:*

- *F is APN*
- *F is almost 3-to-1 with $F^{-1}(0) = \{0\}$*
- *F has the Gold-like Walsh spectrum.*

2.5 Walsh zeroes of functions satisfying the subspace property

A remarkable property of the Kim mapping is that its Walsh zeroes [9], i.e., the set of all (α, β) such that $W_\kappa(\alpha, \beta) = 0$, include two n-dimensional subspaces of $\mathbb{F}_2^{2n}$ in direct sum. One of the reasons is the existence of some (α, β) such that $W_\kappa(\alpha\varphi_1, \beta\varphi_2) = 0$ for all $\varphi_1, \varphi_2 \in \mathbb{F}$. Fig. 1 contains a representation of the Walsh coefficients of the Kim mapping where the masks are sorted by cosets. The grey squares on Fig. 1 then correspond to such spaces. The following theorem provides a necessary and sufficient condition for this phenomenon, for the special case of cyclotomic mappings.

Theorem 2. *Let $F : \mathbb{L} \to \mathbb{L}$ be a cyclotomic mapping of order d over $\mathbb{F}^*$ with $\gcd(d, 2^t - 1) = 1$, Γ be a system of directions, $\mathcal{L}(F) \neq 2^{2t}$, and $\alpha, \beta \in \mathbb{L}^*$. Then,*

$$W_{\mathbb{L},F}(\alpha\varphi_1, \beta\varphi_2) = 0, \ \forall \varphi_1, \varphi_2 \in \mathbb{F}^*$$

if and only if β is such that $N_{F,\beta^{-1}} = 1$ and α satisfies $\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\alpha\gamma_0) \neq 0$ and

$$\left\{ \frac{\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\alpha\gamma)}{\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\beta F(\gamma))^e}, \gamma \in \Gamma \setminus \{\gamma_0\} \right\} = \mathbb{F}$$

where γ_0 is the unique element in Γ such that $\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\beta F(\gamma_0)) = 0$ and $x \mapsto x^e$ is the inverse of $x \mapsto x^d$ over $\mathbb{F}$.

The proof is omitted due to space limitations. Theorem 2 characterizes these peculiar subspaces among the Walsh zeros. Determining whether other cyclotomic mappings satisfying the subspace property share this property with κ now sums up to determining whether the bijectivity of $\frac{\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\alpha\gamma)}{\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\beta F(\gamma))^e}$ is sporadic or not.

3 Additive Subspace Property for Permutations and Streebog Sbox

So far, we only considered multiplicative decompositions of $\mathbb{L}^*$. But as $\mathbb{F}$ is an additive subgroup of $\mathbb{L}$, we can also partition $\mathbb{L}$ as $\mathbb{L} = \bigsqcup_{\lambda\in\mathcal{O}} \lambda + \mathbb{F}$ given any system of representatives $\mathcal{O}$. We refer to any $\lambda + \mathbb{F}$ as an *affine line* and to such λ as its *origin*. Any *set* $\mathcal{O}$ with 2^t elements such that $\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\mathcal{O}) = \mathbb{F}$ is a system of origins; and in particular the subspaces $\lambda\mathbb{F}$ with $\lambda \in \mathbb{L} \setminus \mathbb{F}$.

3.1 Additive Subspace Property for Permutations

We now define a family of permutations Π, which map any multiplicative coset $\gamma\mathbb{F}^*$, $\gamma \neq \gamma^\circ$ onto the (punctured) additive coset $G(\gamma) + \mathbb{F}^*$. Moreover, the restriction of Π to all cosets are the same, up to the additive offset, i.e., all the lines are shuffled the same way. Therefore, this property can be seen as "an additive variant" of the subspace property. However, for $\gamma = \gamma^\circ$, $\mathbb{F}^*$ must be mapped onto $\mathcal{O}$, if we want to construct a permutation of $\mathbb{L}^*$. As we will show later, the functions satisfying this property include, for $n = 8$, the Sbox used in the Russian standard primitives Streebog and Kuznyechik [17,18].

Definition 4 (ASPP). *A function $\Pi\colon \mathbb{L} \to \mathbb{L}$ with $\Pi(0) = 0$ is said to have the* additive subspace property for permutations (ASPP) *if there exist*

- *Γ a system of directions and $\mathcal{O}$ a system of origins, and*
- *two bijective maps: $G\colon \Gamma^\circ \to \mathcal{O}$, $F\colon \mathbb{F} \to \mathbb{F}$ with $F(0) = 0$, such that:*

$$\Pi|_{\mathbb{L}\setminus\mathbb{F}}\colon \gamma\varphi \mapsto G(\gamma) + F(\varphi) \text{ and } \Pi(\mathbb{F}^*) = \mathcal{O}. \tag{1}$$

Such a tuple $(\Gamma^\circ, \mathcal{O}, F, G)$ is called an ASPP-decomposition of Π.

Since only Γ° matters in Definition 4 and not Γ, γ° can always be chosen freely.

Proposition 7. *Any function satisfying the ASPP is bijective.*

For any such function, the decomposition is almost unique.

Definition 5 (Trivially-equivalent decompositions). *Let Π be a function satisfying ASPP and $(\Gamma^\circ, \mathcal{O}, F, G)$ be a decomposition of Π. Let $\varphi \in \mathbb{F}^*$. Let $\widetilde{\Gamma^\circ} := \varphi\Gamma^\circ$, and $\widetilde{F}\colon \mathbb{F} \to \mathbb{F}$, $\widetilde{G}\colon \widetilde{\Gamma^\circ} \to \mathcal{O}$ be defined as:*

$$\widetilde{F} = F \circ \mathrm{M}_\varphi \quad \text{and} \quad \widetilde{G} = G \circ \mathrm{M}_{\varphi^{-1}},$$

where M_φ (resp. $\mathrm{M}_{\varphi^{-1}}$) denotes the multiplication-by-φ (resp by φ^{-1}) mapping. Then, the tuple $(\widetilde{\Gamma^\circ}, \mathcal{O}, \widetilde{F}, \widetilde{G})$ is an ASPP-decomposition of Π. $(\Gamma^\circ, \mathcal{O}, F, G)$ and $(\widetilde{\Gamma^\circ}, \mathcal{O}, \widetilde{F}, \widetilde{G})$ are called trivially-equivalent.

Proposition 8 (Uniqueness of the decomposition). *Let Π be a function satisfying the ASPP. Then all decompositions of Π are trivially equivalent.*

It is worth noting that, while the subspace property was independent of the choice of Γ, this is not the case for the ASPP.

3.2 Streebog/Kuznyechik S-box

The S-box used by the last two Russian standards, Streebog and Kuznyechik, is defined for $n = 8$, i.e., $\mathbb{L} = \mathbb{F}_{256}$ and $\mathbb{F} = \mathbb{F}_{16}$. It is specified [17,18] as a look-up table of integers ranging from 0 to 255. In order to rather study a function $F\colon \mathbb{L} \to \mathbb{L}$, identification between $\mathbb{L}$ and $[\![0, 255]\!]$ must be specified, as two different identifications gives two functions over $\mathbb{L}$ with *a priori* different properties.

The latest study [26] on the S-box points out a representation of $\mathbb{F}_{256}$ that we will also be using: $\mathbb{F}_{256} \simeq \mathbb{F}_2[X]/(X^8+X^4+X^3+X^2+1)$. Identification between 8-bit words, integers of $[\![0, 255]\!]$ and polynomials of degree at most 7 is done canonically: $\sum_{i=0}^{7} b_i 2^i \simeq \sum_{i=0}^{7} b_i X^i \simeq (b_7, \cdots, b_0)$. Finally, we denote $\Lambda\colon \mathbb{L} \to [\![0, 255]\!]$ the isomorphism built from these relations. Through Λ, $\mathbb{L}$ inherits from the ordering of integers.

In the following, we consider the bijective S-box $\pi\colon \mathbb{L} \to \mathbb{L}$ defined as $\pi = \Lambda^{-1} \circ \mathrm{LUT} \circ \Lambda$ where LUT is the look-up table given as specifications [17,18]. The previous studies [26,27,3] of π show how much it interacts with both additive and multiplicative decompositions. We continue along this line of work by studying its normalized form $\pi_0 := \pi + \pi(0)$, which acts as a permutation of $\mathbb{L}^*$.

The *TK-log* decomposition of π [26] can be restated by the fact that π_0 satisfies the ASPP. The decomposition of π is therefore unique and partially exhibited in [26]. Indeed, Perrin shows that the multiplicative coordinates correspond to the decomposition $\Gamma \times \mathbb{F}^*$ where $\Gamma = \{a^i, i \in [\![0, 16]\!]\}$ and a is a well-chosen root of the polynomial defining $\mathbb{L}$, namely $a := \Lambda(2)$ as $2 \simeq X$. In other words, a is the class of X. Instead of Γ, we observe that $a^{17} \{a^{-i}, i \in [\![1, 16]\!]\} = \Gamma^\circ$ and use the associated trivially-equivalent decomposition to describe some properties of π_0 that can be easily verified.

Proposition 9. *Let $b = a^{-1}$. Let $(\Gamma^\circ, \mathcal{O}, F, G)$ be the ASPP-decomposition of π_0 with $\Gamma = \{b^i, i \in [\![0, 16]\!]\} = \{1\} \cup \Gamma^\circ$. Let $G^\circ := \pi_0|_{\mathbb{F}}$.*

1. *Let $\lambda \in \mathbb{L}$, and $\mathcal{O} \cap (\lambda + \mathbb{L}) = \{o_\lambda\}$. Then $\forall \varphi \in \mathbb{F}$, $\Lambda(o_\lambda) \leq \Lambda(\lambda + \varphi)$, meaning that the origin chosen for each affine line is the smallest possible one.*
2. *Let $\mathbb{F}$ be enumerated in increasing order as $\mathbb{F} = \{\varphi_0, \varphi_1, \cdots, \varphi_{15}\}$. Let $i \in [\![0, 16]\!], j \in [\![0, 14]\!]$. Then:*

$$\mathrm{Tr}_{\mathbb{L}/\mathbb{F}} \circ \pi_0(b^{i+17j}) = \begin{cases} \mathrm{Tr}_{\mathbb{L}/\mathbb{F}} \circ G(b^i) = \varphi_{i-1} & \text{if } i \neq 0 \\ \mathrm{Tr}_{\mathbb{L}/\mathbb{F}} \circ G^\circ(b^{17j}) = \varphi_{j+1} & \text{if } i = 0 \end{cases}, \tag{2}$$

and $\mathrm{Tr}_{\mathbb{L}/\mathbb{F}} \circ G^\circ(0) = 0 = \varphi_0$. In other words, enumerating both coordinates of preimages by increasing powers results in enumerating the origins of the images by increasing traces.

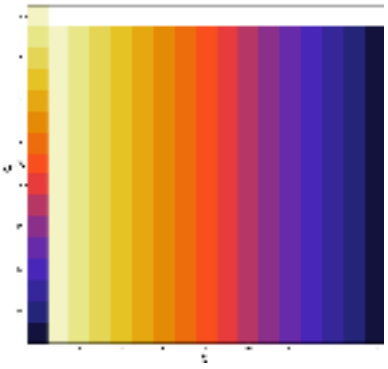


Fig. 2: *Graphical representation of the values of* $\mathrm{Tr}_{\mathbb{L}/\mathbb{F}} \circ \pi_0(b^{i+17j})$. *The first column corresponds to* $i = 0$, *i.e., to* $\mathrm{Tr}_{\mathbb{L}/\mathbb{F}} \circ \pi_0(\mathbb{F})$.

As pointed out in [26], $\mathcal{O}$ is even more structured as it is an $\mathbb{F}_2$-vectorial subspace of dimension 4. This structure is in line with Proposition 9. Indeed, a natural way to obtain a system of origins is to complete any basis of $\mathbb{F}$, $\mathcal{B}_0$, into a basis of $\mathbb{L}$, $\mathcal{B}_0 \cup \mathcal{B}_1$. Then, $\mathcal{O} = \langle \mathcal{B}_1 \rangle$ is a system of origins that is also an $\mathbb{F}_2$-subspace. The most natural algorithm to do so is to exhaust $\mathbb{L} \setminus \mathbb{F}$ while keeping the first four vectors which make the rank grow. This procedure leads to the described system of origins. Regarding π_0, F is the least understood building-block. It remains an open question to determine whether a simple and natural description of F exists or not.

3.3 Walsh coefficients of the functions satisfying the ASPP

We now focus on the Walsh coefficients of the functions satisfying the ASPP. The next proposition expresses them in terms of the functions involved in their ASPP-decomposition. Its proof and that of the following corollary are omitted due to space constraints.

Proposition 10. *Let* Π *be a function satisfying the ASPP,* $(\Gamma^\circ, \mathcal{O}, F, G)$ *an ASPP-decomposition of* Π, *and* $G^\circ = \Pi_{|\mathbb{F}}$. *Let* $\alpha \in \mathbb{L}$ *and* $\beta \in \mathbb{L}^*$ *be decomposed as* $\beta = \gamma_\beta \varphi_\beta$ *with* $\gamma_\beta \in \Gamma^\circ \cup \{1\}, \varphi_\beta \in \mathbb{F}^*$. *Finally, let* $G^\circ_{\gamma_\beta}$ *be the function from* $\mathbb{F}$ *to* $\mathbb{F}$ *defined by* $x \mapsto \mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\gamma_\beta G^\circ(x))$. *Then* $W_{\mathbb{L},\Pi}(\alpha, \beta) = S_{G^\circ} + S_F$ *where*

$$S_{G^\circ} = W_{\mathbb{F},G^\circ_{\gamma_\beta}}\left(\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\alpha), \varphi_\beta\right) - W_{\mathbb{F},G^\circ_{\gamma_\beta}}(0, \varphi_\beta)$$

$$\text{and } S_F = \sum_{\gamma \in \Gamma^\circ} (-1)^{\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\beta G(\gamma))} W_{\mathbb{F},F}\left(\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\alpha\gamma), \mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\beta)\right).$$

Corollary 4 (Output mask in $\mathbb{F}^*$). *Let* Π *be a function satisfying the ASPP,* $(\Gamma^\circ, \mathcal{O}, F, G)$ *be an ASPP-decomposition of* Π, *and* $G^\circ = \Pi_{|\mathbb{F}}$. *Let* $\alpha \in \mathbb{L}^*$, *and* α^{-1} *be decomposed as* $\alpha^{-1} = \varphi_{\alpha^{-1}}\gamma_{\alpha^{-1}}$, $\gamma_{\alpha^{-1}} \in \Gamma, \varphi_{\alpha^{-1}} \in \mathbb{F}^*$. *Let* $\beta \in \mathbb{F}^*$. *Then*

$$W_{\mathbb{L},\Pi}(\alpha, \beta) = \begin{cases} 0 & , \text{ if } \alpha \in \mathbb{F}^* \\ W_{\mathbb{F},G^\circ_1}\left(\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(\alpha), \beta\right) + 2^t (-1)^{\mathrm{Tr}_{\mathbb{L}/\mathbb{F}}\left(\beta G(\gamma_{\alpha^{-1}})\right)} & , \text{ if } \alpha \neq \mathbb{F}, \end{cases}$$

where $G_1^\circ : x \mapsto \mathrm{Tr}_{\mathbb{L}/\mathbb{F}}(G^\circ(x))$.

As an interesting particular case, we can construct functions satisfying the ASPP with remarkable properties such as a linearity which is at most twice the smallest known linearity for an S-box on $\mathbb{L}$.

Proposition 11. *Let Π be a function satisfying the ASPP, with $(\Gamma^\circ, \mathcal{O}, F, G)$ such that $\Gamma^\circ = \mathbb{G} \setminus \{1\}$ and $F = \mathrm{Id}$. Then $\mathcal{L}(F) \leq 2^{t+2}$.*

We can even build such permutations with $\mathcal{L}(F) \leq 3 \times 2^t$: for $n = 8, 10, 12, 14$ and 16 bits, we obtained linearities equal respectively to $44, 80, 156, 300$ and 568.

4 Conclusions

Our results show that the functions of $\mathbb{F}_{2^{2t}}$ obtained by mapping the multiplicative cosets of $\mathbb{F}_{2^t}^*$ into multiplicative or additive cosets of $\mathbb{F}_{2^t}^*$ have interesting properties, like a low differential uniformity and a low linearity. Most notably, these properties capture some functions with remarkable properties, like the Kim mapping and the Streebog S-box. Our work enlightens for instance the role played by this structure on the fact that the Kim mapping is CCZ-equivalent to a permutation.

Acknowledgments. Jules Baudrin is funded by ANR (French National Research Agency) grant ANR-20-CE48-0017 (SELECT). The work of Anne Canteaut is partially supported by ANR grant ANR-21-CE39-0012. The work of Léo Perrin is supported by the European Research Council (ERC, grant agreement no. 101041545 "ReSCALE").

References

1. Bannier, A., Bodin, N., Filiol, E.: Partition-based trapdoor ciphers. Cryptology ePrint Archive, Report 2016/493 (2016), https://eprint.iacr.org/2016/493
2. Berger, T.P., Canteaut, A., Charpin, P., Laigle-Chapuy, Y.: On almost perfect nonlinear functions over $\mathbb{F}_2^n$. IEEE Transactions on Information Theory **52**(9), 4160–4170 (2006)
3. Biryukov, A., Perrin, L., Udovenko, A.: Reverse-engineering the S-box of Streebog, Kuznyechik and STRIBOBr1. In: Fischlin, M., Coron, J.S. (eds.) EUROCRYPT 2016, Part I. LNCS, vol. 9665, pp. 372–402. Springer, Heidelberg, Germany, Vienna, Austria (May 8–12, 2016). https://doi.org/10.1007/978-3-662-49890-3_15
4. Bors, A., Panario, D., Wang, Q.: Functional graphs of generalized cyclotomic mappings of finite fields. arXiv 1108.1873 (2023). https://doi.org/10.48550/arXiv.2304.00181
5. Bors, A., Wang, Q.: Generalized cyclotomic mappings: Switching between polynomial, cyclotomic, and wreath product form. Communications in Mathematical Research **38**(2), 246–318 (2022). https://doi.org/https://doi.org/10.4208/cmr.2021-0029

6. Browning, K.A., Dillon, J., McQuistan, M.T., Wolfe, A.J.: An APN permutation in dimension six. In: Post-proceedings of the 9-th International Conference on Finite Fields and Their Applications. vol. 518, pp. 33–42. American Mathematical Society (2010)
7. Budaghyan, L., Helleseth, T., Li, N., Sun, B.: Some results on the known classes of quadratic APN functions. In: Hajji, S.E., Nitaj, A., Souidi, E.M. (eds.) Codes, Cryptology and Information Security - C2SI 2017. Lecture Notes in Computer Science, vol. 10194, pp. 3–16. Springer (2017). https://doi.org/10.1007/978-3-319-55589-8_1
8. Budaghyan, L., Ivkovic, I., Kaleyski, N.: Triplicate functions. Cryptography and Communications **15**(1), 35–83 (2023)
9. Canteaut, A., Perrin, L.: On CCZ-equivalence, extended-affine equivalence, and function twisting. Finite Fields and Their Applications **56**, 209–246 (2019). https://doi.org/https://doi.org/10.1016/j.ffa.2018.11.008
10. Carlet, C.: Open questions on nonlinearity and on APN functions. In: Koç, Ç.K., Mesnager, S., Savaş, E. (eds.) Arithmetic of Finite Fields. pp. 83–107. Springer International Publishing, Cham (2015)
11. Carlet, C.: Boolean Functions for Cryptography and Coding Theory. Cambridge University Press (2021). https://doi.org/10.1017/9781108606806
12. Carlet, C., Feng, K.: An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity. In: Pieprzyk, J. (ed.) ASIACRYPT 2008. LNCS, vol. 5350, pp. 425–440. Springer, Heidelberg, Germany, Melbourne, Australia (Dec 7–11, 2008). https://doi.org/10.1007/978-3-540-89255-7_26
13. Carlet, C., Mesnager, S.: A note on semi-bent boolean functions. Cryptology ePrint Archive, Paper 2010/486 (2010), https://eprint.iacr.org/2010/486, https://eprint.iacr.org/2010/486
14. Chase, B., Lisonek, P.: Kim-type APN functions are affine equivalent to Gold functions. Cryptogr. Commun. **13**(6), 981–993 (2021). https://doi.org/10.1007/S12095-021-00490-2
15. Chen, L., Coulter, R.S.: Bounds on the differential uniformity of the Wan-Lidl polynomials. Cryptogr. Commun. **15**(6), 1069–1085 (2023). https://doi.org/10.1007/S12095-023-00634-6
16. Dillon, J.F.: Elementary Hadamard difference-sets. Ph.D. thesis, University of Maryland, USA (1974)
17. Federal Agency on Technical Regulation and Metrology: Information technology – data security: Hash function. English version available at http://wwwold.tc26.ru/en/standard/gost/GOST_R_34_11-2012_eng.pdf (2012)
18. Federal Agency on Technical Regulation and Metrology: Information technology – data security: Block ciphers. English version available at http://wwwold.tc26.ru/en/standard/gost/GOST_R_34_12_2015_ENG.pdf (2015)
19. Göloğlu, F.: Almost perfect nonlinear trinomials and hexanomials. Finite Fields and Their Applications **33**, 258–282 (2015)
20. Kölsch, L., Kriepke, B., Kyureghyan, G.M.M.: Image sets of perfectly nonlinear maps. Des. Codes Cryptogr. **91**(1), 1–27 (2023). https://doi.org/10.1007/S10623-022-01094-4
21. Laigle-Chapuy, Y.: Permutation polynomials and applications to coding theory. Finite Fields and Their Applications **13**(1), 58–70 (2007). https://doi.org/10.1016/J.FFA.2005.08.003

22. Li, K., Li, C., Helleseth, T., Qu, L.: A complete characterization of the APN property of a class of quadrinomials. IEEE Trans. Inf. Theory **67**(11), 7535–7549 (2021). https://doi.org/10.1109/TIT.2021.3102872
23. Lidl, R., Niederreiter, H.: Finite fields, Encyclopedia of Mathematics and its Applications, vol. 20. Cambridge University Press (1997)
24. Lou, Y., Han, H., Tang, C., Xu, M.: Constructing vectorial boolean functions with high algebraic immunity based on group decomposition. Cryptology ePrint Archive, Paper 2012/335 (2012), https://eprint.iacr.org/2012/335, https://eprint.iacr.org/2012/335
25. Paterson, K.G.: Imprimitive permutation groups and trapdoors in iterated block ciphers. In: Knudsen, L.R. (ed.) FSE'99. LNCS, vol. 1636, pp. 201–214. Springer, Heidelberg, Germany, Rome, Italy (Mar 24–26, 1999). https://doi.org/10.1007/3-540-48519-8_15
26. Perrin, L.: Partitions in the S-box of Streebog and Kuznyechik. IACR Trans. Symm. Cryptol. **2019**(1), 302–329 (2019). https://doi.org/10.13154/tosc.v2019.i1.302-329
27. Perrin, L., Udovenko, A.: Exponential s-boxes: a link between the s-boxes of BelT and Kuznyechik/Streebog. IACR Trans. Symm. Cryptol. **2016**(2), 99–124 (2016). https://doi.org/10.13154/tosc.v2016.i2.99-124, https://tosc.iacr.org/index.php/ToSC/article/view/567
28. Wan, D., Lidl, R.: Permutation polynomials of the form $x^r f(x^{(q-1)/d})$ and their group structure. Monatshefte für Mathematik **112**, 149–163 (1991)
29. Wang, Q.: Cyclotomic mapping permutation polynomials over finite fields. In: Golomb, S.W., Gong, G., Helleseth, T., Song, H.Y. (eds.) Sequences, Subsequences, and Consequences. pp. 119–128. Springer Berlin Heidelberg, Berlin, Heidelberg (2007)
30. Wang, Q.: A note on inverses of cyclotomic mapping permutation polynomials over finite fields. Finite Fields and Their Applications **45**, 422–427 (2017). https://doi.org/https://doi.org/10.1016/j.ffa.2017.01.006
31. Zheng, J., Wu, B., Chen, Y., Liu, Z.: Constructing $2m$-variable boolean functions with optimal algebraic immunity based on polar decomposition of $\mathbb{F}_{2^{2m}}^*$. arXiv 1304.2946 (2013)

How to Lose Some Weight – A Practical Template Syndrome Decoding Attack

Sebastian Bitzer[1], Jeroen Delvaux[2], Elena Kirshanova[2], Sebastian Maaßen[3], Alexander May[3], and Antonia Wachter-Zeh[1]

[1] Technical University of Munich, Munich, Germany
[2] Technology Innovation Institute, Abu Dhabi, UAE
[3] Ruhr University Bochum, Bochum, Germany

Abstract. We study the hardness of the Syndrome Decoding problem, the base of most code-based cryptographic schemes, such as Classic McEliece, in the presence of side-channel information. We use ChipWhisperer equipment to perform a template attack on Classic McEliece running on an ARM Cortex-M4, and accurately classify the Hamming weights of consecutive 32-bit blocks of the secret error vector $\mathbf{e} \in \mathbb{F}_2^n$. With these weights at hand, we optimize Information Set Decoding algorithms. Technically, we show how to speed up information set decoding via a dimension reduction, additional parity-check equations, and an improved information set search, all derived from the Hamming weight information.
Consequently, using our template attack, we can practically recover an error vector $\mathbf{e} \in \mathbb{F}_2^n$ in dimension $n = 2197$ in a matter of seconds. Without side-channel information, such an instance has a complexity of around 88 bit. We also estimate how our template attack affects the security of the proposed McEliece parameter sets. Roughly speaking, even an error-prone leak of our Hamming weight information leads for $n = 3488$ to a security drop of 89 bits.

1 Introduction

Hardness of Syndrome Decoding. Central to all code-based schemes that advanced to the 4th Round of the NIST Post-Quantum Standardization Process [ARBC+20,ABB+23,MAB+23] lies the *Syndrome Decoding* (SD) *problem*: given a parity-check matrix $H \in \mathbb{F}_2^{(n-k)\times n}$, where $\mathbb{F}_2$ denotes the binary finite field, a syndrome $\mathbf{s} \in \mathbb{F}_2^{n-k}$, and an integer $w < n$, find the error vector $\mathbf{e}$ such that $H\mathbf{e} = \mathbf{s}$ and $|\mathbf{e}| < w$, where $|\cdot|$ denotes the Hamming weight.

An algorithm for solving this problem for a uniformly random H leads to a message or key recovery attack for the aforementioned schemes. Therefore, the syndrome decoding problem has received a significant amount of attention, resulting in various methods to solve it: Information Set Decoding (ISD) [Pra62,Ste89,MMT11], Statistical Decoding [Al 01,CDMHT22], and, recently, Sieving-style algorithms [GJN23,DEEK23]. Despite this extensive theoretical effort, the problem remains tractable for relatively small dimensions.

Concretely, in the setting of Classic McEliece (e.g., $w \approx \frac{n}{5 \log_2 n}$ and $k \approx 0.8n$), the largest solved instance reported today [ALL19] is for $n = 1470$, and it already requires an optimized GPU implementation of an advanced information set decoding algorithm [NFK23], together with significant computational resources.[4]

Side-Channel Attacks. For the practical security of code-based schemes, it is important that the syndrome decoding problem also offers sufficient robustness against realistic side-channel attacks using leaks of the secret error vector $\mathbf{e} \in \mathbb{F}_2^n$. Compared to the comprehensive study of the syndrome decoding problem's classical security, its side-channel resistance has received much less attention.

Some initial theoretical work of Horlemann et al. [HPR+22] classifies different leakages and shows how to incorporate them into ISD algorithms to solve the syndrome decoding problem faster. One of the leakages considered in [HPR+22, Section 4] is *known Hamming weights of error blocks*.

In this leakage setting, one knows $\{|\mathbf{e}_i|\}_{i \le t}$, where $\mathbf{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_t)$ and all $\mathbf{e}_i$'s (except, may be the last $\mathbf{e}_t$) are of the same length, i.e., the *word size* of the Central Processing Unit (CPU). For example, for an ARM Cortex-M4, each word $\mathbf{e}_i$ consists of 32 bits. Typical target instructions are *loads*, which move 32-bit words from SRAM to CPU registers, and *stores*, which move 32-bit words from CPU registers to SRAM. When executing such instructions, the power consumption is slightly different for each possible weight $|\mathbf{e}_i|$, and these unique characteristics can be condensed into a so-called *template* [CRR03]. We call the respective modified syndrome decoding problem, which additionally receives $\{|\mathbf{e}_i|\}_{i \le t}$, the *template syndrome decoding* (template SD) problem.

While Horlemann et al. [HPR+22] describe a potential template syndrome decoding attack, their attack remains purely theoretical. Neither do the authors realize concrete power trace leaks, nor do they provide an improved information set decoding implementation. Thus, the practical implications of code-based template attacks remain unclear.

Contribution. In this work, we perform for the first time an explicit template attack on a Classic McEliece implementation. To this end, we realize a concrete power trace leak, from which we derive with high accuracy (but still error-prone) the desired Hamming weight information $\{|\mathbf{e}_i|\}_{i \le t}$.

We then improve information set decoding by using and enhancing the techniques of Horlemann et al. [HPR+22]. Building on information set decoding software from Esser, May, and Zweydinger [EMZ22], we provide a concrete implementation of these improvements.

With our (erroneous but easily correctable) leakage, we run our template information set decoding and retrieve the secret $\mathbf{e} \in \mathbb{F}_2^n$. Concretely, we are able to solve the template syndrome decoding problem for Classic McEliece in dimension $n = 2197$ in a matter of seconds. Without template, such an instance has complexity around 88 bits. In more detail, our results are as follows.

[4] See also https://groups.google.com/a/list.nist.gov/g/pqc-forum/c/WzgqEmAfnk8 for the discussion on hardness predictions for this instance.

1. We use ChipWhisperer equipment to measure the power consumption of an open-source implementation [CC21] of Classic McEliece running on an ARM Cortex-M4, or at least a decapsulation subroutine that checks whether $|\mathbf{e}| = w$. Using $48k$ traces for template building, and $12k$ for matching, the weights $\{|\mathbf{e}_i|\}$ we recover are correct with a probability of around 97%. We show how to deal with this measurement noise in the full version.
2. We modify the ISD algorithms of Prange [Pra62] and Dumer [Dum91] by incorporating the template. Specifically, we show how to encode the knowledge of weights of error blocks into the parity-check matrix H. Then, using such modified H, we show how to decrease the expected running time of the above ISD algorithms, again exploiting the leakage.
3. We provide efficient and parallelized implementations of the modified ISD algorithms. With our software we are able to solve the $n = 2197$ instance from [ALL19] in a matter of 10 seconds on AMD EPYC 7742 using 200 threads. Based on our implementation, we estimate the hardness of larger McEliece instances under this template attack.

Related work. Closely related to the template syndrome decoding is *regular syndrome decoding* introduced in [AFS05]. In regular SD, for each block $\mathbf{e}_i$ of $\mathbf{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_t)$ it holds that $|\mathbf{e}_i| = 1$. Note that regular SD is a special case of Template SD. Recent work of Esser and Santini [ES23] studies the hardness of regular SD, and some of their ideas apply to our setting, e.g., the construction of new parity-check equations, see also [EMZ22].

Another template attack on Classic McEliece was presented by Grosso et al. in [GCCD23]. The authors of [GCCD23] aim at the same leakage, namely, $\{|\mathbf{e}_i|\}_{i \le t}$ but they retrieve it from the matrix-vector multiplication $H \cdot \mathbf{e}$ that computes the syndrome. In our template attack, similar to [GCCD23], we discard the columns of H that correspond to the zero-weight blocks in the template. Contrary to [GCCD23], in our work we show how to make use of the *non-zero* weight blocks to speed-up ISD algorithms, and we implement our ISD algorithms in order to actually retrieve the secret.

Another side-channel attack exploiting failures of the decoding procedure in McEliece decryption is studied in [LNPS20]. The authors show how to learn the positions of 1's in the secret vector by querying the decoder with modified syndromes. Similar to our work, the authors combine the obtained information with ISD algorithms and estimate their attack performance. In contrast, we implement our (modified) ISD routines, report on concrete runtimes for feasible instance and then give estimates for large dimensions.

In summary, in contrast to [GCCD23] and [LNPS20] we do not only *estimate* the effects on ISD, but we retrieve Hamming weight side-channel information, correct errors, provide improved ISDs via dimension reduction and additional parity check equations, and *practically solve* an $n = 2197$-dimensional template SD instance in a matter of seconds.

Artifacts. Our software for Template ISD algorithms as well as scripts to generate the figures are available in [Tem24].

2 Template ISD

Notations. Let $|\mathbf{x}|$ denote the Hamming weight of $\mathbf{x}$ and by $[i, j)$ the interval of consecutive integers $\{i, i+1, \ldots, j-1\}$. By S_n we denote the group of all permutations on sets of size n. By $\mathbb{I}_n$ we denote the identity matrix of rank n.

Problem definitions. In the Classic McEliece KEM [ARBC+20], the decryption process receives as input a syndrome $\mathbf{s} \in \mathbb{F}_2^{n-k}$ and recovers the secret message $\mathbf{e}$ by calling an efficient syndrome decoder using the McEliece secret key. Once $\mathbf{e}$ is retrieved, the decryption checks if $|\mathbf{e}| = w$, where w is the decoding capacity of the syndrome decoder. The parameter w is a fixed public parameter. Classic McEliece decryption only returns $\mathbf{e}$, if $\mathbf{e}$ passes the check $|\mathbf{e}| = w$.

Without knowledge of the secret key, message recovery attacks on Classic McEliece require solving the Syndrome Decoding (SD) problem.

Definition 1 (Syndrome Decoding (SD)). *Let $H \in \mathbb{F}_2^{(n-k)\times n}$ be a random-looking parity-check matrix, $\mathbf{e}$ an error vector of Hamming weight w, and $\mathbf{s} = H\mathbf{e} \in \mathbb{F}_2^{n-k}$ the corresponding syndrome. SD asks to find the unique weight-w $\mathbf{e} \in \mathbb{F}_2^n$ satisfying $H\mathbf{e} = \mathbf{s}$.*

The side-channel attack we consider in this work creates a template for the function that computes $|\mathbf{e}|$. In the ideal scenario, such a template allows the attacker to learn the blockwise weight of $\mathbf{e}$. We call the SD problem that in addition receives the blockwise weight *template Syndrome Decoding.*

Definition 2 (Template SD). *Let $H \in \mathbb{F}_2^{(n-k)\times n}$ be a parity-check matrix of a random code and $\mathbf{s} = H\mathbf{e} \in \mathbb{F}_2^{n-k}$, for some $\mathbf{e}$ of Hamming weight w. Let further $\mathbf{e} = (\mathbf{e}_1, \ldots, \mathbf{e}_t)$ with $\mathbf{e}_i \in \mathbb{F}_2^b$ for $i = [1, t)$, $\mathbf{e}_t \in \mathbb{F}_2^{n-b\cdot(t-1)}$, and $w_i = |\mathbf{e}_i|$. Template ISD asks to find $\mathbf{e}$ given $H, \mathbf{s}$, and $\{w_i\}_{i\le t}$.*

Definition 3 (Guess). *We call any vector $\{\hat{w}_i\}_{i\le t} \in \mathbb{N}_0^t$ a* guess. *The* accuracy *of a guess is the percentage of correctly identified weights, i.e. $\frac{|\{i\in[1,t]|\hat{w}_i=w_i\}|}{t}$. A guess is* error-free *if it has accuracy 1, otherwise it is* error-prone. *Notice that in general error-prone guesses do not satisfy $\sum_{i=1}^{t} \hat{w}_i = w$.*

The block size b, and, therefore, also the template's length depends on the target architecture's specifications. Our template attack targets ARM Cortex-M4 processor that operates on words of 32 bits. Hence, our guesses will be of length $t = \lceil n/32 \rceil$.

Running Example $n = 2197$. Our running example uses the parameters $n = 2197$, $k = 1758$, and $w = 37$. Therefore, a guess is a string of length $t = 69$, its i-th entry indicating the weight of the i-th 32-bit block of $\mathbf{e}$.

3 Algorithms for Template ISD

3.1 Permutation-based Template ISD – Improving Prange

Let us start with the fundamental information set decoding algorithm due to Prange [Pra62]. Prange's algorithm permutes the columns of H, which is equivalent to permuting the positions of 1's in $\mathbf{e}$.

Let $\pi \in S_n$ be a permutation and let $\pi(H) = (Q \mid \cdot)$ be the result of applying the permutation π to H such that $Q \in \mathbb{F}_2^{(n-k)\times(n-k)}$ is invertible (this event occurs with constant probability). Multiplying by Q^{-1} from the left both $\pi(H)$ and $\mathbf{s}$ leads to an equivalent SD instance written in systematic form:

$$H'\pi(\mathbf{e}) = \mathbf{s}', \quad \text{where} \quad H' = Q^{-1}H = (\mathbb{I}_{n-k} \mid \cdot), \text{ and } \mathbf{s}' = Q^{-1}\mathbf{s}.$$

If $\pi(\mathbf{e})$ has weight 0 on the last k coordinates, then $|\mathbf{s}'| = w$. This means that the first $(n-k)$ coordinates of $\pi(\mathbf{e})$ are given by $\mathbf{s}'$ and $\mathbf{e}$ can be reconstructed.

Dimension reduction. As already noticed in the work of Grosso et al. [GCCD23], any weight-0 block with $w_i = 0$ does not contribute to the solution $\mathbf{e}$. Let m_0 denote the number of error-free blocks. Then $b \cdot m_0$ columns of H do not contribute and can be eliminated, leading to a modified parity-check matrix $\bar{H} \in F_2^{(n-k)\times\bar{n}}$ with $\bar{n} = n - b \cdot m_0$ columns. This in turn reduces the dimension of the solution $\mathbf{e}$ from n to $\bar{n} = n - b \cdot m_0$ leaving its weight w unchanged.

Improved permutation. The idea of the permutation π in Prange's algorithm is to move all w 1-entries of $\mathbf{e}$ upfront to the first $n-k$ coordinates. Having weight w_i for the i-th block, we permute a number proportional to w_i upfront. Concretely, in Algorithm 2, we use the following *template-optimized permutation.*

Let P be a permutation matrix and $v_i \in \mathbb{Z}$ with $0 \le v_i \le b$ and $\sum_{i=1}^{t} v_i = n-k$. Further, denote the permuted error vector as

$$P\mathbf{e} = (\mathbf{e}', \mathbf{e}'') = (\mathbf{e}'_1, \ldots, \mathbf{e}'_t, \mathbf{e}''_1, \ldots, \mathbf{e}''_t)$$

with $\mathbf{e}'_i \in \mathbb{F}_2^{v_i}$ and $\mathbf{e}''_i \in \mathbb{F}_2^{b-v_i}$. Then, P is a template permutation if $\mathbf{e}'_i$ and $\mathbf{e}''_i$ originate from $\mathbf{e}_i$ for all i. The success probability $P(\sum_i |\mathbf{e}''_i| = 0)$ is determined by the v_i. In [HPR+22], a greedy algorithm for optimizing v_i is given. We observe that this optimal choice corresponds to the setting $v_i \approx \frac{w_i}{w} \cdot (n-k)$, i.e., the number of columns is chosen *proportional* to the weight of the block. This proportional assignment of columns generalizes the puncturing of [GCCD23]: columns of H with $w_i = 0$ are implicitly ignored by taking 0 columns from the blocks with $w_i = 0$. In Algorithm 2, the procedure that samples a template permutation as described above is called `TemplatePerm`.

In practice, $v_i = \frac{w_i}{w} \cdot (n-k)$ cannot be used directly due to rounding issues and the restriction $v_i \le b$. In our implementation, we minimize $|v_i - \frac{w_i}{w} \cdot (n-k)|$.

Additional Parity Check Equations. Note that $|\mathbf{e}_i| = w_i$ implies that the sum of the entries of $\mathbf{e}_i$ is $w_i \bmod 2$, see also [EMZ22]. Hence, for $w_i > 0$, one can extend the parity-check matrix by appending a row of the shape $(0, \ldots, 0, 1, \ldots, 1, 0, \ldots 0)$, where the all-1 block is at the positions $[i \cdot b, (i+1) \cdot b)$. The syndrome $\mathbf{s}$ is extended by appending $w_i \bmod 2$. Each appended check increases the co-dimension

of the code by one to eventually $n-k+t-m_0$. This makes it simpler for ISD to find a permutation that puts all weight to the first co-dimension many positions.

Algorithm 1: Prange

Input : H, $\mathbf{s}$, w
Output: $\mathbf{e}$

repeat
  Sample $P \in S_n$
  Let $H' = Q^{-1}HP$ be the systematic form of HP
  $\mathbf{s}' = Q^{-1}\mathbf{s}$
until $|\mathbf{s}'| = w$
return $P^{-1} \cdot (\mathbf{s}', 0^k)$.

Algorithm 2: Template Prange

Input : $H \in \mathbb{F}_2^{(n-k)\times n}$, $\mathbf{s}$, $\{w_i\}_{i\le t}$; $\sum_i w_i = w$
Output: $\mathbf{e}$

Let $m_0 := |\{i \le t \mid w_i = 0\}|$, $\bar{n} = n - m_0 b$, $\bar{k} = k + m_0 - t$.
Obtain $\bar{H} \in \mathbb{F}_2^{(n-k)\times \bar{n}}$ by removing zero blocks.
Obtain $\bar{H} \in \mathbb{F}_2^{(n-\bar{k})\times \bar{n}}$, $\bar{\mathbf{s}} \in \mathbb{F}_2^{n-\bar{k}}$ by adding checks.
repeat
  $P \leftarrow \texttt{TemplatePerm}(w)$
  Let $H' = Q^{-1}\bar{H}P$ be the systematic form of $\bar{H}P$
  $\mathbf{e}' = Q^{-1}\bar{\mathbf{s}}$.
until $|\mathbf{e}'| = w$
return $P^{-1} \cdot (\mathbf{e}', 0^k)$.

We summarize all modifications to the parity-check matrix and the optimized permutations in Figure 1.

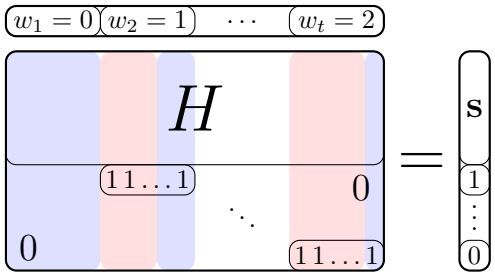


Fig. 1: Illustration of our improved Template ISD method. Columns in blocks with error weight $w_i = 0$ are punctured. For $w_i \neq 0$, an additional check is appended to the parity-check matrix and the syndrome. For each block, the number of columns chosen for permutation upfront (colored red) is set proportionally to the error weight.

Theorem 1. *Let $\{w_i\}_{i\le t}$ be a an error-free guess with m_0 many zeros. The expected number of permutations of Algorithm 2 for solving Template SD is*

$$\prod_{i=1}^{t} \binom{b}{w_i} \binom{\lfloor \frac{w_i}{w}(n-k+t-m_0) \rceil}{w_i}^{-1}.$$

Proof. Our Algorithm 2 finds a good permutation if for all t blocks of length b, all w_i-many 1's from the i-th block will be moved upfront to the first $n-\bar{k} = n-k+t-m_0$ coordinates. As from each block we take $\lfloor \frac{w_i}{w}(n-k+t-m_0) \rceil$ many positions, the expected number of required permutation follows. □

Running example $n = 2197$. According to [EVZB23], the concrete complexity of Algorithm 1 for $n = 2197$, $k = 1758$, $w = 37$ is estimated as 116 bits. Dimension reduction by weight-0 blocks reduces the complexity of this instance to 71 bits. With improved permutation and additional parity check equations from Algorithm 2, the complexity further decreases to 62 bits. Figures for larger McEliece instances are available in [Tem24].

3.2 Enumeration-based Template ISD – Improving Dumer

Recall from Section 3.1 that for a parity-check matrix $H \in \mathbb{F}_2^{(n-k)\times n}$ Prange's algorithm finds a permutation that shifts all w 1-entries of $\mathbf{e}$ upfront to the first $n-k$ entries. That is why Prange is called a permutation-based ISD.

Instead, enumeration based ISD algorithms like [Dum91,FS09,MMT11] choose small parameters p, ℓ and permute $\mathbf{e}$ such that weight $w-2p$ lands on the first $n-k-\ell$ coordinates, and the remaining weight $2p$ lands on the last $k+\ell$ coordinates. On the one hand, such a permutation is way more likely to find the secret. On the other hand, we now have to enumerate a search space of size $\binom{k+\ell}{2p}$, in Dumer's algorithm in a meet-in-the-middle fashion. For usual McEliece such a tradeoff pays off, i.e., the benefit of faster finding a suitable permutation outweighs the drawback of enumeration.

In this work, we chose to adapt Dumer's algorithm to the Template ISD setting. In the parameter range that we practically solve, Dumer's algorithm is known to perform best, whereas more advanced algorithm like [MMT11] are taking over for large n of cryptographic size [EMZ22].

Although we now choose with Dumer an enumeration-based ISD algorithm, the benefits from the Template ISD still contribute to a large extent to the search for a suitable permutation. Namely, analogous to Section 3.1, we obtain the template version of Dumer using the following modifications and improvements:

Dimension reduction: 0-weight blocks from the guess $\{w_i\}_{i\le t}$ are removed, let m_0 be their number. Such a dimension reduction helps to significantly speed up permutation search, and it also decreases the search space for enumeration by $m_0 b$.

Additional parity checks: Encoding all $w_i \ge 1$ into additional check equations increases the co-dimension from $n-k$ to $n-k+t-m_0$. This speeds up permutation search further, and slightly reduces the enumeration cost.

Improved permutation: Similar to Section 3.1, we permute upfront proportionally to the weights w_i to improve the permutation. For this, we set $v_i \approx \frac{w_i \cdot c}{w-2p}(n-\bar{k}-\ell)$, where $c = \frac{w-2p}{w}$ is a re-scaling factor. We do not exploit non-zero weights for enumeration.

The resulting algorithm Template Dumer is given in Algorithm 3.

Theorem 2 (Template Dumer). *Let $k' \coloneqq \bar{n} - (n-\bar{k}) + \ell$ with $\bar{n}$ and $\bar{k}$ as in Algorithm 3. Then, the number of iterations that Template Dumer ISD performs on average is the inverse of the success probability*

$$\binom{k'/2}{p}^2 \binom{k'}{2p}^{-1} \sum_{p_1+\ldots+p_t=2p} \prod_{i=1}^{t} \binom{\lfloor \frac{w_i}{w}(n-\bar{k}-\ell)\rceil}{w_i-p_i} \binom{b}{w_i}^{-1}, \tag{1}$$

where each iteration has a meet-in-the-middle cost of $2\binom{k'/2}{p} + \binom{k'/2}{p}^2 \cdot 2^{-\ell}$.

Proof. Since $\lfloor \frac{w_i}{w}(n-\bar{k}-\ell)\rceil$ positions of the i-th block are moved upfront, it contributes p_i errors to the last k' positions with probability $\binom{\lfloor \frac{w_i}{w}(n-\bar{k}-\ell)\rceil}{w_i-p_i}\binom{b}{w_i}^{-1}$.

Algorithm 3: Template Dumer

Input : $H \in \mathbb{F}_2^{(n-k)\times n}$, $\mathbf{s}$, $\{w_i\}_{i\le t}$; $\sum_i w_i = w$, b, p, ℓ
Output: $\mathbf{e}$

Let $m_0 := |\{i \le t \mid w_i = 0\}|$, $\bar{n} = n - m_0 b$, $\bar{k} = k + m_0 - t$, $k' = \bar{n} - (n - \bar{k}) + \ell$.
Obtain $\bar{H} \in \mathbb{F}_2^{(n-k)\times \bar{n}}$ by removing zero blocks.
Obtain $\bar{H} \in \mathbb{F}_2^{(n-\bar{k})\times \bar{n}}$, $\bar{\mathbf{s}} \in \mathbb{F}_2^{n-\bar{k}}$ by adding checks.
repeat
  $P \leftarrow$ `TemplatePerm`$(w - 2p)$
  Let $H' = Q^{-1}\bar{H}P$ be the quasi-systematic form of $\bar{H}P$,
    $(\mathbf{s}', \mathbf{s}'') = Q^{-1}\bar{\mathbf{s}} \in \mathbb{F}_2^{n-\bar{k}-\ell} \times \mathbb{F}_2^{\ell}$.
  for *all collisions* $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{F}_2^{k'/2}$ *with weight* p **do**
    Compute $\mathbf{e}' := H_1\mathbf{e}_1 + H_2\mathbf{e}_2 + \mathbf{s}'$. ▷ via Meet-in-the-Middle
until $|\mathbf{e}'| = w - 2p$
return $P^{-1} \cdot (\mathbf{e}', \mathbf{e}_1, \mathbf{e}_2)$.

Similar to [HPR+22], the probability of $2p$ errors in the last k' positions is obtained by summing over all possibilities $p_1 + \ldots + p_t = 2p$. Further, the error needs to split evenly in the last k' positions. Randomizing the order of these coordinates, this probability is $\binom{k'/2}{p}^2 \binom{k'}{2p}^{-1}$. The meet-in-the-middle step requires enumerating $2\binom{k'/2}{p}$ vectors $\mathbf{e}_1$, $\mathbf{e}_2$, leading to $\binom{k'/2}{p}^2 2^{-\ell}$ collisions on average.

Running example $n = 2197$. For $n = 2197$, $k = 1758$, $w = 37$, we pick $\ell = 16$ and $p = 2$. The performance differs between guesses. On average, $2^{19.9}$ iterations are sufficient, each with a Meet-in-the-Middle cost of processing $2^{14.7}$ vectors.

3.3 Dealing with Noisy Guesses

The full version provides an algorithm that deals with noisy guesses. In particular, we show that Template Prange and Template Dumer are robust to single (or very few) misclassifications.

4 Side-Channel Experiments

4.1 Measurement Setup

We target an open-source C implementation of McEliece, which is made by Chen and Chou [CC21], optimized for the ARM Cortex-M4, and unprotected against side-channel attacks. The targeted function is a Hamming-weight computation in the decryption, as specified in Listing 1.1 [CC22]. To accelerate our measurements, we do not run the entire decapsulation, and instead communicate via UART with a custom wrapper around function `weight_3488`. Likewise, although solving $n = 3488$ is computationally feasible, we reduce n to 2197 for faster

results. The code is compiled by arm-none-eabi-gcc using O3 optimization. Although the right-shift of the 32-bit word $\mathbf{v}$ in Listing 1.1 might leak bit-level information, we only aim to recover word-level information, i.e., weights $|\mathbf{v}|$.

```
static int weight_3488(uint32_t *v)
{
  int i, w = 0;
  for (i = 0; i < 3488; i++)
    w += (v[i>>5] >> (i&31)) & 1;
  return w;
}
```

Listing 1.1: Targeted C function [CC22].

The power consumption is measured using ChipWhisperer equipment: a CW308 UFO board, an STM32F405RGT6 target that contains an ARM Cortex-M4, and a Husky oscilloscope. The clock frequency is set to 16 MHz and the sampling frequency is set to 128 MHz, i.e., there are 8 samples per clock period. To synchronize traces, the wrapper raises a trigger signal right before function weight_3488 is executed. To capture the entire operation, 201559 samples suffice. As the Husky has a buffer of 131070 samples, we stitch together 2 traces by varying the offset from the trigger. Traces for template building and template matching are taken from the same STM chip, which is fair: to build templates, the attacker can perform unlimited encapsulations to obtain known pairs $(\mathbf{c}, \mathbf{e})$.

4.2 Template Building

Given that error $\mathbf{e}$ spans 69 words, each having weight $W \in \{0, 1, 2, 3\}$ with overwhelming probability, $276 = 69 \times 4$ templates are built. For this purpose, we randomly generate $48k$ error vectors $\mathbf{e}$ and measure one trace for each $\mathbf{e}$. To ensure that the templates have similar qualities, we impose $P(W = 0) = P(W = 1) = P(W = 2) = P(W = 3) = 1/4$. This deviation from the McEliece distribution is optional and is only realistic for a 2-device attack. All choose-W-out-of-32 selections are equally likely. For example, words 0x80020040 and 0x01400002 are equally likely in the case of $W = 3$. For each out of 69 words, we only retain the 100 samples that matter most. All other samples primarily generate classification noise, so it's beneficial to discard them. To make a selection, we use an extension of Welch's t-test specified below, where M_W is the *sample mean*, V_W is the *sample variance*, and N_W is the number of traces for each weight W.

$$T = \frac{1}{3}\left(\frac{M_0 - M_1}{\sqrt{\frac{V_0}{N_0} + \frac{V_1}{N_1}}} + \frac{M_1 - M_2}{\sqrt{\frac{V_1}{N_1} + \frac{V_2}{N_2}}} + \frac{M_2 - M_3}{\sqrt{\frac{V_2}{N_2} + \frac{V_3}{N_3}}}\right)$$

4.3 Template Matching

For each error $\mathbf{e}$ we aim to recover, we collect $12k$ traces, and average them into a single trace X. Now, the weights W are non-uniform and follow the McEliece distribution. For each out of 69 words, the distinguisher $D_W = \sum_{i=0}^{99} |T_i| \cdot |M_W - X_i| \in \mathbb{R}^+$ is computed. The weight W for which D_W is the smallest is the best guess. The probability that this guess is correct is around 97%.

5 Practical Template SD Solving with Our Algorithms

We implemented Algorithm 2 and Algorithm 3. The source-code of our implementation can be found in [Tem24]. We ran our experiments on the parity-check matrices of Classic McEliece instances with parameters provided by [ALL19], where we generated the solution vectors $\mathbf{e}$ ourselves. We fully recovered the secret error vector for all instances $n \leq 2197$.

In the experiments, we always worked with an error-free guess. Indeed, for our running example $n = 2197$, the actual side-channel attacks gave guesses with 97% accuracy resulting in a single mispredicted block: we observed a guess $\hat{\mathbf{w}}$ with $\sum_i \hat{w}_i = w - 1$, which can be corrected with low overhead.

To accurately estimate the running time of Algorithms 1 (Original Prange), 2 (Template Prange), and 3 (Template Dumer) for all dimensions n, we generated random error vectors and measured the runtime *per permutation*. To obtain the overall runtime, we multiply by the expected number of permutations, which is computed as in Theorems 1 and 2 by averaging over different error vectors. The resulting estimates are presented in Figure 2.

Running Example $n = 2197$. Our implementation of Algorithm 3 (with $p = 2$, $\ell = 16$) on 2x AMD EPYC 7742 CPUs recovers the secret $\mathbf{e}$ for $n = 2197$ in 1019 seconds with 1134185 iterations required (the predicted number of iterations for this instance is $1.6 \cdot 10^6$). The implementation is parallelized over the choice of permutation, and with 200 threads outputs the secret in 10 seconds using only 334 MB of RAM.

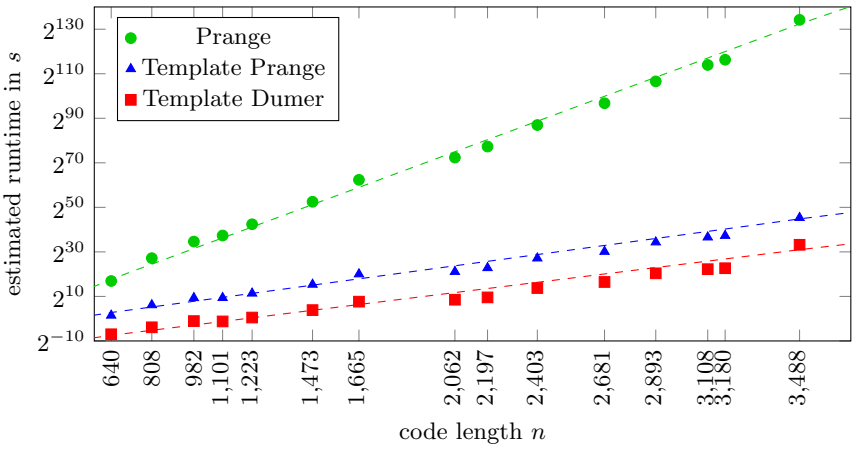


Fig. 2: Single-threaded performance of our implementation on AMD EPYC 7742

References

ABB+23. Nicolas Aragon, Pailo L. Barreto, Slim Bettaieb, Loıc Bidoux, Olivier Blazy, Jean-Christophe Deneuville, Philippe Gaborit, Santosh Ghosh, Shay Gueron, Tim Guneysu, Carlos Aguilar Melchor, Rafael Misoczki, Edoardo Persichetti, Jan Richter-Brockmann, Nicolar Sendrier, Jean-Pierre Tillich, Vasseur Valentin, and Gilles Zémor. BIKE - Bit Flipping Key Encapsulation. https://bikesuite.org/, 2023.

AFS05. Daniel Augot, Matthieu Finiasz, and Nicolas Sendrier. A family of fast syndrome based cryptographic hash functions. In *Progress in Cryptology – Mycrypt 2005*, pages 64–83, 2005.

Al 01. A. Kh. Al Jabri. A statistical decoding algorithm for general linear block codes. In Bahram Honary, editor, *8th IMA International Conference on Cryptography and Coding*, volume 2260 of *LNCS*, pages 1–8. Springer, Heidelberg, December 2001.

ALL19. Nicolas Aragon, Julien Lavauzelle, and Matthieu Lequesne. decodingchallenge.org, 2019.

ARBC+20. Martin Albrecht R., Daniel J. Bernstein, Tung Chou, Carlos Cid, Jan Gilcher, Tanja Lange, Varun Maram, Ingo von Maurich, Rafael Misoczki, Ruben Niederhagen, Kenneth G. Paterson, Persichetti Edoardo, Christiane Peters, Peter Schwabe, Nicolas Sendrier, Jakub Szefer, Cen Jung Tjhai, Martin Tomlinson, and Wen Wang. Classic McEliece: conservative code-based cryptography. https://classic.mceliece.org/nist/mceliece-20201010.pdf, 2020.

CC21. Ming-Shing Chen and Tung Chou. Classic McEliece on the ARM cortex-M4. *IACR TCHES*, 2021(3):125–148, 2021. https://tches.iacr.org/index.php/TCHES/article/view/8970.

CC22. Ming-Shing Chen and Tung Chou. Classic McEliece implementation for ARM-Cortex M4. https://github.com/pqcryptotw/mceliece-arm-m4/blob/main/pqm4-projects/crypto_kem/mceliece348864/ches2021/decrypt_n3488_t64.c, commit f2a699dd480f9f91d566eb4b910fd4e51e3bdc91, January 2022.

CDMHT22. Kevin Carrier, Thomas Debris-Alazard, Charles Meyer-Hilfiger, and Jean-Pierre Tillich. Statistical decoding 2.0: Reducing decoding to LPN. In Shweta Agrawal and Dongdai Lin, editors, *ASIACRYPT 2022, Part IV*, volume 13794 of *LNCS*, pages 477–507. Springer, Heidelberg, December 2022.

CRR03. Suresh Chari, Josyula R. Rao, and Pankaj Rohatgi. Template attacks. In Burton S. Kaliski Jr., Çetin Kaya Koç, and Christof Paar, editors, *CHES 2002*, volume 2523 of *LNCS*, pages 13–28. Springer, Heidelberg, August 2003.

DEEK23. Léo Ducas, Andre Esser, Simona Etinski, and Elena Kirshanova. Asymptotics and improvements of sieving for codes. *Cryptology ePrint Archive*, 2023.

Dum91. Ilya Dumer. On minimum distance decoding of linear codes. In *Proc. 5th Joint Soviet-Swedish Int. Workshop Inform. Theory*, pages 50–52. Moscow, 1991.

EMZ22. Andre Esser, Alexander May, and Floyd Zweydinger. McEliece needs a break - solving McEliece-1284 and quasi-cyclic-2918 with modern ISD. In Orr Dunkelman and Stefan Dziembowski, editors, *EUROCRYPT 2022,*

Part III, volume 13277 of *LNCS*, pages 433–457. Springer, Heidelberg, May / June 2022.

ES23. Andre Esser and Paolo Santini. Not just regular decoding: Asymptotics and improvements of regular syndrome decoding attacks. *Cryptology ePrint Archive*, 2023.

EVZB23. Andre Esser, Javier A. Verbel, Floyd Zweydinger, and Emanuele Bellini. {CryptographicEstimators}: a software library for cryptographic hardness estimation. {*IACR*} *Cryptol. ePrint Arch.*, page 589, 2023.

FS09. Matthieu Finiasz and Nicolas Sendrier. Security bounds for the design of code-based cryptosystems. In *Advances in Cryptology–ASIACRYPT 2009: 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, December 6-10, 2009. Proceedings 15*, pages 88–105. Springer, 2009.

GCCD23. Vincent Grosso, Pierre-Louis Cayrel, Brice Colombier, and Vlad-Florin Drăgoi. Punctured syndrome decoding problem: Efficient side-channel attacks against classic McEliece. In *International Workshop on Constructive Side-Channel Analysis and Secure Design*, pages 170–192. Springer, 2023.

GJN23. Qian Guo, Thomas Johansson, and Vu Nguyen. A new sieving-style information-set decoding algorithm. *Cryptology ePrint Archive*, 2023.

HPR+22. Anna-Lena Horlemann, Sven Puchinger, Julian Renner, Thomas Schamberger, and Antonia Wachter-Zeh. Information-set decoding with hints. In *Code-Based Cryptography*, pages 60–83, 2022.

LNPS20. Norman Lahr, Ruben Niederhagen, Richard Petri, and Simona Samardjiska. Side channel information set decoding using iterative chunking - plaintext recovery from the "classic McEliece" hardware reference implementation. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part I*, volume 12491 of *LNCS*, pages 881–910. Springer, Heidelberg, December 2020.

MAB+23. Carlos Aguilar Melchor, Nicolas Aragon, Slim Bettaieb, Loıc Bidoux, Olivier Blazy, Jurjen Bos, Jean-Christophe Deneuville, Arnaud Dion, Philippe Gaborit, Jérôme Lacan, Edoardo Persichetti, Jean-Marc Robert, Pascal Véron, and Gilles Zémor. Hamming quasi-cyclic (HQC). https://pqc-hqc.org/, 2023.

MMT11. Alexander May, Alexander Meurer, and Enrico Thomae. Decoding random linear codes in $\tilde{\mathcal{O}}(2^{0.054n})$. In Dong Hoon Lee and Xiaoyun Wang, editors, *ASIACRYPT 2011*, volume 7073 of *LNCS*, pages 107–124. Springer, Heidelberg, December 2011.

NFK23. Shintaro Narisada, Kazuhide Fukushima, and Shinsaku Kiyomoto. Multiparallel MMT: faster ISD algorithm solving high-dimensional syndrome decoding problem. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 106(3):241–252, 2023.

Pra62. Eugene Prange. The use of information sets in decoding cyclic codes. *IRE Trans. Inf. Theory*, 8(5):5–9, 1962.

Ste89. Jacques Stern. A method for finding codewords of small weight. In *Coding Theory and Applications*, pages 106–113, 1989.

Tem24. A practical template syndrome decoding attack. https://github.com/ChuTriel/TemplateISD, 2024.

The geometry of covering codes in the sum-rank metric

Matteo Bonini[1], Martino Borello[2], and Eimear Byrne[3]

[1] Aalborg University, Department of Mathematical Sciences, Aalborg, Denmark mabo@math.aau.dk
[2] Université Paris 8, Laboratoire de Géométrie, Analyse et Applications, LAGA, Université Sorbonne Paris Nord, CNRS, UMR 7539, France martino.borello@univ-paris8.fr
[3] School of Mathematics and Statistics, University College Dublin, Ireland ebyrne@ucd.ie

Abstract. We introduce the concept of a sum-rank saturating system and outline its correspondence to a sum-rank metric code with a given covering radius. We consider the problem of determining the shortest ρ-saturating systems for a fixed dimension, which is equivalent to the covering problem in the sum-rank metric. We obtain upper and lower bounds on this quantity. We give constructions of saturating systems arising from geometrical structures.

Keywords. Linear sets, saturating sets, sum-rank metric codes, covering radius
MSC2020. 05B40, 11T71, 51E20, 52C17, 94B75

Introduction

Researchers have extensively explored the connections between linear codes and sets of points in finite geometries, as evidenced by previous works as [1,10,12,13,15,17]. The construction of a generator matrix or parity check matrix for a linear code can be accomplished through a multiset of projective points, with the supports of codewords corresponding to complements of hyperplanes in a fixed projective set. The interconnection between these two domains facilitates the application of methods from one field to the other. Notably, this approach has been employed in constructing codes with a bounded *covering radius*, associated with *saturating sets* in projective space. Recent investigations into the geometry of rank-metric codes codes [2, 23] reveal their correspondence to q-systems and linear sets. A similar correspondence holds for sum-rank metric codes [21, 24].

The covering radius of a code is the smallest positive integer ρ such that the union of the spheres of radius ρ about each codeword equals the entire ambient space. The covering radius serves as an indicator of combinatorial properties, such as *maximality*, and is an invariant of code equivalence. It also provides insight into error-correcting capabilities by determining the maximal weight of a correctable error. This essential coding theoretical parameter has been extensively studied for codes in the context of the Hamming metric [7,11–16]. However, only a few papers in the literature on rank-metric codes and sum-rank metric address this parameter [5,9,19,22]. Recently, in [5] a purely geometrical approach based on saturating system was proposed the study the covering radius in the rank metric. This approach allowed to provide new bounds, and interesting examples of covering codes in the rank metric, see [3,5].

In this paper, we extend these ideas to the sum-rank metric by introducing the concept of a sum-rank saturating system, aligning it with a sum-rank metric covering code. We also provide new bounds for covering codes in the sum-rank metric, as well as examples arising from cutting systems.

1 Preliminaries

Let q be a prime power and $\mathbb{F}_q$ be the finite field with q elements. Let t be a positive integer and $\mathbf{n} = (n_1, \dots, n_t), \mathbf{m} = (m_1, \dots, m_t) \in \mathbb{N}^t$ be ordered tuples with $n_1 \leq n_2 \leq \cdots \leq n_t$ and $m_1 \leq m_2, \leq \cdots \leq m_t$, and we set $N := n_1 + \cdots + n_t$. Throughout the paper, we will use the following notations for the direct sums of vector spaces $\mathbb{F}_q^{\mathbf{n}} := \bigoplus_{i=1}^t \mathbb{F}_q^{n_i}$ and for direct sums of matrix spaces $\mathrm{Mat}_{\mathbf{n}\times\mathbf{m}}(\mathbb{F}_q) := \bigoplus_{i=1}^t \mathbb{F}_q^{n_i \times m_i}$.

Definition 1. *Given a pair of nonnegative integers N and M, the* q-binomial *or* Gaussian coefficient *counts the number of M-dimensional subspaces of an N-dimensional subspace over $\mathbb{F}_q$ and is given by:*

$$\begin{bmatrix} N \\ M \end{bmatrix}_q := \prod_{i=0}^{M-1} \frac{q^N - q^i}{q^M - q^i}.$$

We write $\mathcal{N}$ to denote the poset $\{(a_1,\dots,a_t) : 0 \leq a_i \leq n_i\}$ endowed with the partial order $\leq$ defined by

$$(a_1,\dots,a_t) \leq (b_1,\dots,b_t) \iff a_i \leq b_i \text{ for all } i \in [t].$$

We will adopt the following notation: for $\mathbf{u} = (u_1,\dots,u_t), \mathbf{v} = (v_1,\cdots,v_t)$ we define,

$$|\mathbf{u}| := \sum_{j=1}^{t} u_j, \qquad \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}_q := \prod_{j=1}^{t} \begin{bmatrix} u_j \\ v_j \end{bmatrix}_q, \qquad q^{\binom{\mathbf{u}}{2}} := \prod_{j=1}^{t} q^{\binom{u_j}{2}}.$$

Definition 2. *Let $X := (X_1,\dots,X_t) \in Mat(\mathbf{n},\mathbf{m},\mathbb{F}_q)$. The sum-rank support of X is defined as the space*

$$\operatorname{supp}(X) := (\operatorname{colsp}(X_1), \operatorname{colsp}(X_2), \dots, \operatorname{colsp}(X_t)) \subseteq \mathbb{F}_q^{\mathbf{n}},$$

where $\operatorname{colsp}(X_i)$ is the $\mathbb{F}_q$-span of the columns of X_i. The rank-list *of X is defined as*

$$\operatorname{rkl}(X) := (\operatorname{rk}(X_1), \dots, \operatorname{rk}(X_t)) \in \mathbb{N}^t.$$

Finally, the sum-rank weight of X is the quantity

$$\mathrm{w}_{\mathrm{srk}}(X) := \dim_{\mathbb{F}_q}(\operatorname{supp}(X)) := \sum_{i=1}^{t} \operatorname{rk}(X_i).$$

Definition 3. *A sum-rank metric code $\mathcal{C}$ is an $\mathbb{F}_q$-linear subspace of $\mathrm{Mat}_{\mathbf{n}\times\mathbf{m}}(\mathbb{F}_q)$ endowed with the sum-rank distance*

$$\begin{aligned} d_{\mathrm{srk}} :& \mathrm{Mat}_{\mathbf{n}\times\mathbf{m}}(\mathbb{F}_q) \times \mathrm{Mat}_{\mathbf{n}\times\mathbf{m}}(\mathbb{F}_q) \longrightarrow \mathbb{N} \\ & (X,Y) \mapsto \mathrm{w}_{\mathrm{srk}}(X-Y). \end{aligned}$$

The minimum sum-rank distance of a sum-rank code $\mathcal{C}$ is defined to be:

$$d_{\mathrm{srk}}(\mathcal{C}) := \min\{w_{\mathrm{srk}}(X) \ : \ X \in \mathcal{C}, X \neq \mathbf{0}\}.$$

The sum-rank support of the code C is the $\mathbb{F}_q$-span of the supports of all the codewords of C, that is

$$\operatorname{supp}(C) := \sum_{X \in C} \operatorname{supp}(X) \subseteq \mathbb{F}_q^{\mathbf{n}}.$$

$\mathcal{C}$ is said to be sum-rank non-degenerate *if $\operatorname{supp}(\mathcal{C}) = \mathbb{F}_q^{\mathbf{n}}$. For each $i \in [t]$, we write $C_i := \Pi_i(C)$, where $\Pi_i : \mathrm{Mat}_{\mathbf{n}\times\mathbf{m}}(\mathbb{F}_q) \longrightarrow \mathbb{F}_q^{n_i \times m_i}$ denotes the canonical projection map.*

Definition 4. *The* dual *of a code $C \leq \mathrm{Mat}_{\mathbf{n}\times\mathbf{m}}(\mathbb{F}_q)$ is defined as*

$$C^{\perp} := \left\{ (Y_1,...,Y_t) \in \mathrm{Mat}_{\mathbf{n}\times\mathbf{m}}(\mathbb{F}_q) \;\middle|\; \sum_{i=1}^{t} \operatorname{Tr}(X_i Y_i^T) = 0 \text{ for all } (X_1,...,X_t) \in C \right\} \leq \mathrm{Mat}_{\mathbf{n}\times\mathbf{m}}(\mathbb{F}_q).$$

Definition 5. *Let $\mathcal{U}$ be an $\mathbb{F}_q$-subspace of dimension n in $\mathbb{F}_{q^m}^k$. The* $\mathbb{F}_q$-linear set *in $\mathrm{PG}(k-1,q^m)$ of rank n associated to $\mathcal{U}$ is the set*

$$L_{\mathcal{U}} := \{\langle u \rangle_{\mathbb{F}_{q^m}} : u \in \mathcal{U} \setminus \{0\}\},$$

where $\langle u \rangle_{\mathbb{F}_{q^m}}$ denotes the projective point corresponding to u.

2 Sum-rank saturating systems

Definition 6. *For each $i \in \{1,\dots,t\}$, let $\mathcal{U}_i$ be an $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^m}^k$ of dimension n_i. If the ordered t-tuple $\mathcal{U} = (\mathcal{U}_1,\dots,\mathcal{U}_t)$ satisfies $\langle \mathcal{U}_1,\dots,\mathcal{U}_t \rangle_{\mathbb{F}_{q^m}} = \mathbb{F}_{q^m}^k$ then $\mathcal{U}$ is called an $[\mathbf{n},k]_{q^m/q}$* system. *We say that $\mathcal{U}$ has dimension $\mathbf{n}$. A* generator matrix *for $\mathcal{U}$ is a $k \times \sum_{j=1}^{t} n_j$ matrix over $\mathbb{F}_{q^m}$ the form $G = [G_1|\cdots|G_t]$, where for each i, G_i is a generator matrix for the $[n_i,k]_{q^m/q}$ system $\mathcal{U}_i$, that is such that the $\mathbb{F}_q$-span of the columns of each G_i is $\mathcal{U}_i$.*

Definition 7. *Two sum-rank systems* $(\mathcal{U}_1, \ldots, \mathcal{U}_t)$ *and* $(\mathcal{V}_1, \ldots, \mathcal{V}_t)$ *are equivalent if there exists an isomorphism* $\varphi \in \mathrm{GL}(k, \mathbb{F}_{q^m})$*, an element* $\mathbf{a} = (a_1, \ldots, a_t) \in (\mathbb{F}_{q^m}^*)^t$ *and a permutation* $\sigma \in \mathcal{S}_t$*, such that for every* $i \in \{1, \ldots, t\}$

$$\varphi(\mathcal{U}_i) = a_i \mathcal{V}_{\sigma(i)}.$$

We recall the definition of a ρ-saturating set.

Definition 8. *Let* $\mathcal{S} \subseteq \mathrm{PG}(k-1, q^m)$.

(a) *A point* $Q \in \mathrm{PG}(k-1, q^m)$ *is said to be* ρ-saturated *by* $\mathcal{S}$ *if there exist* $\rho+1$ *points* $P_1, \ldots, P_{\rho+1} \in \mathcal{S}$ *such that* $Q \in \langle P_1, \ldots, P_{\rho+1} \rangle_{\mathbb{F}_{q^m}}$*. We also say that* $\mathcal{S}$ ρ-saturates Q.
(b) *The set* $\mathcal{S}$ *is called a* ρ-saturating *set of* $\mathrm{PG}(k-1, q^m)$ *if every point* $Q \in \mathrm{PG}(k-1, q^m)$ *is* ρ*-saturated by* $\mathcal{S}$ *and* ρ *is the smallest value with this property.*

Definition 9. $\mathcal{U}$ *is sum-rank* ρ*-saturating if* $L_{\mathcal{U}_1} \cup \cdots \cup L_{\mathcal{U}_t}$ *is* $(\rho-1)$*-saturating.*

Theorem 1. *Let* $\mathcal{U}$ *be an* $[\mathbf{n}, k]_{q^m/q}$ *system and let* G *be any generator matrix of* $\mathcal{U}$*. The following are equivalent:*

(a) $\mathcal{U}$ *is sum-rank* ρ*-saturating.*
(b) *For each vector* $v \in \mathbb{F}_{q^m}^k$ *there exists* $\lambda = (\lambda_1, \ldots, \lambda_t) \in \mathbb{F}_{q^m}^{1 \times n_1} \times \ldots \times \mathbb{F}_{q^m}^{1 \times n_t}$ *with* $\mathrm{wt}_{\mathrm{srk}}(\lambda) \leq \rho$ *such that*

$$v = G(\lambda_1, \ldots, \lambda_t)^T,$$

and ρ *is the smallest value with this property.*
(c) *We have*

$$\mathbb{F}_{q^m}^k = \bigcup_{\substack{(\mathcal{S}_i : i \in [t]) \colon \mathcal{S}_i \leq_{\mathbb{F}_q} \mathcal{U}_i, \\ \sum_{i=1}^t \dim_{\mathbb{F}_q} \mathcal{S}_i \leq \rho}} \left(\bigcup_{i=1}^t \langle \mathcal{S}_i \rangle_{\mathbb{F}_{q^m}} \right)$$

and ρ *is the smallest integer with this property.*

Definition 10. *Let* $\mathcal{U}$ *be an* $[\mathbf{n}, k]_{q^m/q}$ *system. For each positive integer* ρ*, we define*

$$\mathbb{S}_\rho(\mathcal{U}) := \bigcup_{\substack{(\mathcal{S}_i : i \in [t]) \colon \mathcal{S}_i \leq_{\mathbb{F}_q} \mathcal{U}_i, \\ \sum_{i=1}^t \dim_{\mathbb{F}_q} \mathcal{S}_i \leq \rho}} \left(\bigcup_{i=1}^t (\mathcal{S}_i \otimes \mathbb{F}_{q^m}) \right).$$

It is immediate from Theorem 1 that $\mathcal{U}$ is sum-rank ρ-saturating if ρ is the least integer satisfying $\mathbb{F}_{q^m}^k = \mathbb{S}_\rho(\mathcal{U})$.

The following statement is the sum-rank analogue of [5, Theorem 2.5].

Theorem 2. *Let* $\mathcal{U}$ *be an* $[\mathbf{n}, k]_{q^m/q}$ *system associated to a code* $\mathcal{C}$*. The following are equivalent.*

(a) $\mathcal{U}$ *is sum-rank* ρ*-saturating.*
(b) $\rho_{srk}(\mathcal{C}^\perp) = \rho$.

Definition 11. *For* $i = 1, 2$*, let* $\mathcal{U}_i$ *be a sum-rank* ρ_i*-saturating* $[\mathbf{n}_i, k_i]_{q^m/q}$ *system that is associated with a code* $\mathcal{C}_i$ *that has generator matrix* G_i*. We define the* direct sum *of* $\mathcal{U}_1$ *and* $\mathcal{U}_2$*, which we denote by* $\mathcal{U}_1 \oplus \mathcal{U}_2$*, to be the* $[(\mathbf{n}_1, \mathbf{n}_2), k_1 + k_2]_{q^m/q}$ *system associated with the direct sum of* $\mathcal{C}_1$ *and* $\mathcal{C}_2$*, i.e. the code whose generator matrix is*

$$G_1 \oplus G_2 := \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}.$$

It is straightforward to establish the following (c.f. [5]).

Theorem 3. *For* $i \in [t]$*, let* $\mathcal{U}_i$ *be a sum-rank* ρ_i*-saturating* $[\mathbf{n}_i, k_i]_{q^m/q}$ *system. Then* $\mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_t$ *is an* $[(\mathbf{n}_1, \ldots, \mathbf{n}_t), k_1 + \cdots + k_t]_{q^m/q}$ *system and is sum-rank* ρ*-saturating, where* $\rho \leq \rho_1 + \cdots + \rho_t$.

Definition 12. *A sum-rank* ρ*-saturating system is* reducible *if there exists* $i \in \{1, \ldots, t\}$ *such that the system without the block* $\mathcal{U}_i$ *is sum-rank* ρ*-saturating. Otherwise, the system is* irreducible.

3 Bounds on the dimension of sum-rank saturating systems

As in the classical cases, it is interesting to know how short can a sum-rank metric code of a given dimension and covering radius ρ be, or equivalently how small can the rank of a sum-rank ρ-saturating system be, in a given vector space. However, the situation here is much more complicated than in the Hamming and rank-metric cases.

We start with a bound which follows from the geometric characterisation of our systems. In the proof, we will use the following well-known estimates:

$$\begin{bmatrix} a \\ b \end{bmatrix}_q < f(q)\, q^{b(a-b)}, \qquad \text{for } a, b \in \mathbb{N}, \tag{1}$$

$$q^{e_1} + \ldots + q^{e_r} < \frac{q}{q-1} q^{e_r}, \qquad \text{for } e_i \in \mathbb{Z},\ 0 \le e_1 < \ldots < e_r. \tag{2}$$

where $f(q) = \prod_{i=1}^{+\infty} (1 - q^{-i})^{-1}$.

Theorem 4. *Let $\mathcal{U}$ be a sum-rank ρ-saturating $[\mathbf{n}, k]_{q^m/q}$ system. Then*

$$q^{m\rho} \sum_{\mathbf{s} \in \mathcal{N}, |\mathbf{s}| = \rho} \begin{bmatrix} \mathbf{n} \\ \mathbf{s} \end{bmatrix}_q \ge q^{mk}.$$

In particular,

$$\frac{1}{4t} \cdot \sum_{1 \le i < j \le t} (n_j - n_i)^2 + \frac{\rho(|n| - \rho)}{t} + 2t \ge m(k - \rho). \tag{3}$$

Remark 1. We have that $f(q) \longrightarrow 1$ as $q \longrightarrow \infty$, and so asymptotically $\frac{q f(q)^t}{q-1} \longrightarrow 1$ as $q \longrightarrow \infty$. For this reason, as q grows, we may replace (3) with

$$\frac{1}{4t} \cdot \sum_{1 \le i < j \le t} (n_j - n_i)^2 + \frac{\rho(|n| - \rho)}{t} \ge m(k - \rho),$$

for sufficiently large q. Indeed, even for relatively small values of q, $\frac{q f(q)^t}{q-1}$ takes values much smaller than q for t not exceeding q. For example, for $q = 211, t = 20$, we have $\frac{q f(q)^t}{q-1} \approx 1.105407$; for $q = 111, t = 111$ we have $\frac{q f(q)^t}{q-1} \approx 2.780617$.

Remark 2. For $t = 1$ (rank-metric), the bound coincides asymptotically with the one obtained in [5] (while for small q in [5] we could avoid the rough estimate). When $n_1 = \ldots = n_t = n$,

$$N = tn \ge \frac{tm}{\rho}(k - \rho) + \rho - \frac{2t^2}{\rho}. \tag{4}$$

Lemma 1. *Fix t and N. Let $[n_1, \ldots, n_t]$ and $[n'_1, \ldots, n'_t]$ be such that $n_1 \ge \ldots \ge n_t$, $n'_1 \ge \ldots \ge n'_t$ and $N = n_1 + \ldots + n_t = n'_1 + \ldots + n'_t$. Then*

$$\sum_{1 \le i < j \le t} (n_j - n_i)^2 \le \sum_{1 \le i < j \le t} \left(n'_j - n'_i\right)^2$$

if and only if $[n_1, \ldots, n_t] \preceq [n'_1, \ldots, n'_t]$ in the lexicographic ordering.

This means that, ρ, t, N being fixed, the left hand-side of (3) gets its minimum and maximum values when $n_1 = \ldots = n_t$ and when $n_2 = \ldots = n_t = 1$ respectively. This gives sense to the following definition.

Definition 13. *Let t be a positive integer. We define the* shortest length

$$s_{q^m/q}(k, \rho, t) := \min \left\{ \sum_{i=1}^{t} \dim(\mathcal{U}_i) : \mathcal{U}_i \le_{\mathbb{F}_q} \mathbb{F}_{q^m}^k, (\mathcal{U}_1, \ldots, \mathcal{U}_t) \text{ is sum-rank-}\rho \text{ saturating} \right\},$$

i.e. it is the minimal sum of the $\mathbb{F}_q$-dimensions of the $\mathcal{U}_i$, $i \in \{1, \ldots, t\}$, of a sum-rank ρ-saturating system $\mathcal{U} = (\mathcal{U}_1, \ldots, \mathcal{U}_t)$ in $\mathbb{F}_{q^m}^k$.
We define the homogeneous shortest length

$$s^{\hom}_{q^m/q}(k, \rho, t) := \min \left\{ tn : \mathcal{U}_i \le_{\mathbb{F}_q} \mathbb{F}_{q^m}^k, \dim(\mathcal{U}_i) = n, (\mathcal{U}_1, \ldots, \mathcal{U}_t) \text{ is sum-rank-}\rho \text{ saturating} \right\},$$

i.e. it is the minimal sum of the $\mathbb{F}_q$-dimensions of the $\mathcal{U}_i$, $i \in \{1, \ldots, t\}$, of a sum-rank ρ-saturating system $\mathcal{U} = (\mathcal{U}_1, \ldots, \mathcal{U}_t)$ in $\mathbb{F}_{q^m}^k$, with the additional hypothesis that they all have equal dimension.

Remark 3. Notice that given a sum-rank saturating system $\mathcal{U} = (\mathcal{U}_1, \ldots, \mathcal{U}_t)$ having generator matrix $G = [G_1|\cdots|G_t]$. We can always consider the system $\mathcal{U}' = (\mathcal{U}_1, \ldots, \mathcal{U}_{t-2}, \mathcal{U}'_{t-1})$ having generator matrix $G = [G_1|\cdots|G_{t-2}|G'_{t-1}]$, where G'_{t-1} is the matrix having as columns the union of $\mathbb{F}_q$-bases of $\mathcal{U}_{t-1}$ and $\mathcal{U}_t$. Since $\mathcal{U}_{t-1} + \mathcal{U}_t = \mathcal{U}'_{t-1}$ we have $\dim_{\mathbb{F}_q}(\mathcal{U}'_{t-1}) \leq \dim_{\mathbb{F}_q}(\mathcal{U}_{t-1}) + \dim_{\mathbb{F}_q}(\mathcal{U}_t)$, while $\rho(\mathcal{U}') \leq \rho(\mathcal{U})$.

This remark shows that reducing the number of the blocks usually provides an yields the following result, that provides us the motivation to fix a given value for t before starting the investigation on the minimal dimension of a system $\mathcal{U}$.

Proposition 1 (Monotonicity in t). *We have that $s_{q^m/q}(k, \rho, t) \leq s_{q^m/q}(k, \rho, t+1)$*

Lemma 2. *Let $\mathcal{U} = (\mathcal{U}_1, \ldots, \mathcal{U}_t)$ be a sum-rank ρ-saturating $[\mathbf{n}, k]_{q^m/q}$ system. Suppose for some $i \in [t]$, $\mathcal{U}_i$ is not scattered. Then $\mathcal{U}' = (\mathcal{U}_1, \ldots, \mathcal{U}'_i, \ldots, \mathcal{U}_t)$ is a sum-rank-ρ'-saturating $[\mathbf{n}', k]_{q^m/q}$ system satisfying $\rho' \leq \rho + 1$ and $\mathbf{n}' = (n_1, \ldots, n_i - 1, \ldots, n_t)$.*

Proof. The statement follows as a direct consequence of [5, Lemma 4.5]

Lemma 3. *Let $\mathcal{U} = (\mathcal{U}_1, \ldots, \mathcal{U}_t)$ be a sum-rank ρ-saturating $[\mathbf{n}, k]_{q^m/q}$ system. Suppose that for each $i \in [t]$, $\mathcal{U}_i$ has an $\mathbb{F}_q$-basis $\{u_1^{(i)}, \ldots, u_{n_i}^{(i)}\}$ such that*

$$u_{n_t}^{(t)} = \lambda \sum_{i \in S} \sum_{\substack{j=1, \\ j \neq n_t}}^{n_i} a_j^{(i)} u_j^{(i)},$$

for some $a_j^{(i)} \in \mathbb{F}_q$ and $S \subseteq [t]$. Then $\mathcal{U}' = (\mathcal{U}_1, \ldots, \mathcal{U}_{t-1}, \mathcal{U}'_t)$ is a sum-rank-ρ'-saturating $[\mathbf{n}', k]_{q^m/q}$ system satisfying $\rho' \leq \rho + |S|$ and $\mathbf{n}' = (n_1, \ldots, n_{t-1}, n_t - 1)$.

In particular, Lemma 2 follows as a special case of Lemma 3.

We have the following observations of the monotonicity of $s_{q^m/q}(k, \rho, t)$. The proofs are similar to those of [5, Theorem 4.6].

Theorem 5 (Monotonicity in ρ). *Let $|\mathbf{n}| > k$. The following hold.*

1. $s_{q^m/q}(k, \rho, t) \leq s_{q^m/q}(k, \rho + 1, t)$.
2. $s_{q^m/q}(k, \rho, t) \leq s_{q^m/q}(k + 1, \rho, t) - 1$.
3. $s_{q^m/q}(k + 1, \rho + 1, t) \leq s_{q^m/q}(k, \rho + 1, t) + 1$.

Definition 14. *For each $i \in \{1, 2\}$, let $\mathcal{U}^{(i)}$ be an $[\mathbf{n^{(i)}}, k_i]_{q^m/q}$ system, associated with an $[\mathbf{n^{(i)}}, k_i]_{q^m/q}$ sum-rank-metric code $\mathcal{C}_i$. Let $f : \mathbb{F}_{q^m}^{\mathbf{n^{(1)}}} \longrightarrow \mathbb{F}_{q^m}^{\mathbf{n^{(2)}}}$ be an $\mathbb{F}_{q^m}$-linear map. The code*

$$\mathcal{C} := \{(u, f(u) + v) : u \in \mathcal{C}_1, v \in \mathcal{C}_2\}$$

is an $[(\mathbf{n^{(1)}}, \mathbf{n^{(2)}}), k_1 + k_2]_{q^m/q}$, which we call the f-sum of $\mathcal{C}_1$ and $\mathcal{C}_2$. Its associated $[(\mathbf{n^{(1)}}, \mathbf{n^{(2)}}), k_1 + k_2]_{q^m/q}$ system is called the f-sum of $\mathcal{U}^{(1)}$ and $\mathcal{U}^{(2)}$, which we denote by $\mathcal{U}^{(1)} \oplus_f \mathcal{U}^{(2)}$. If f is the zero map, we write $\mathcal{U}^{(1)} \oplus \mathcal{U}^{(2)}$, and call it the direct sum; if f is the identity map, we we write $\mathcal{U}^{(1)} \oplus_\iota \mathcal{U}^{(2)}$ and call it the Plotkin-sum of $\mathcal{U}^{(1)}$ and $\mathcal{U}^{(2)}$.

Theorem 6. *For each $i \in \{1, 2\}$, let $\mathbf{n}^{(i)} = (\mathbf{n}_1^{(i)}, \ldots, \mathbf{n}_{t_i}^{(i)})$, and let $\mathcal{U}^{(i)}$ be an $[\mathbf{n^{(i)}}, k_i]_{q^m/q}$ sum-rank-ρ_i-saturating system, associated with an $[n_i, k_i]_{q^m/q}$ code $\mathcal{C}_i$. Let $f : \mathbb{F}_{q^m}^{\mathbf{n}^{(1)}} \longrightarrow \mathbb{F}_{q^m}^{\mathbf{n}^{(2)}}$ be an $\mathbb{F}_{q^m}$-linear map. Then $\mathcal{U}^{(1)} \oplus_f \mathcal{U}^{(2)}$ is an $[(\mathbf{n^{(1)}}, \mathbf{n^{(2)}}), k_1 + k_2]_{q^m/q}$ system that is sum-rank-ρ-saturating, where $\rho \leq \rho_1 + \rho_2$. In particular, if $\rho_1 + \rho_2 \leq \min\{k_1 + k_2, m\}$, then*

$$s_{q^m/q}(k_1 + k_2, \rho_1 + \rho_2, t_1 + t_2) \leq s_{q^m/q}(k_1, \rho_1, t_1) + s_{q^m/q}(k_2, \rho_2, t_2).$$

Theorem 7. *Let $\mathbb{F}_{q^m} = \mathbb{F}_q[\alpha]$, $r \geq 1$, $h \geq r$ and*

$$A_{h,r} := \left[\begin{array}{c|c|c|c|c} I_r & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & I_{h-r} & \alpha I_{h-r} & \cdots & \alpha^{m-1} I_{h-r} \end{array}\right]$$

Then

$$G_t := \underbrace{\left[\begin{array}{c|c|c|c} A_{h,r} & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & A_{h,r} & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & A_{h,r} \end{array}\right]}_{t \text{ times}}$$

generates an homogeneous sum-rank rt-saturating system. So

$$s^{\mathrm{hom}}_{q^m/q}(th, tr, t) \leq t(m(h-r)+r).$$

Remark 4. Since

$$t\left(\frac{m}{r}(h-r)+r\right) \leq s^{\mathrm{hom}}_{q^m/q}(th, tr, t) \leq t(m(h-r)+r),$$

we see immediately that when $r = 1$ the lower and the upper bounds coincide, so that

$$s^{\mathrm{hom}}_{q^m/q}(th, t, t) = t(m(h-1)+1).$$

4 Constructions of short sum-rank saturating systems

4.1 Sum-rank saturating systems from partitions of the projective space

A partition of the vector space $\mathbb{F}_{q^m}^k$ yields a partition of $\mathrm{PG}(k-1, q^m)$ into subspaces. In [8], some necessary conditions and constructions of partitions are presented. We know that $\mathcal{U}$ is sum-rank ρ-saturating if $L_{\mathcal{U}_1} \cup \cdots \cup L_{\mathcal{U}_t}$ is $(\rho-1)$-saturating.

Proposition 2. *Let $\mathcal{P} = \{\mathcal{P}_i\}_{i\in\{1,\ldots,t\}}$ a partition of $\mathrm{PG}(k-1, q^m)$ into subspaces. Let k_i be a positive integer such that $\mathcal{P}_i \simeq \mathrm{PG}(k_i - 1, q^m)$. If $\mathcal{U}$ is such that each $\mathcal{U}_i$ is rank ρ-saturating in $\mathcal{P}_i$, then $\mathcal{U}$ is sum-rank ρ'-saturating with $\rho' \leq \rho$.*

In [20, Theorem 4.28] we get that, if $(m, k) = 1$, there exists a partition of $\mathrm{PG}(k-1, q^m)$ into

$$t = \frac{(q^{mk}-1)(q-1)}{(q^m-1)(q^k-1)}$$

subgeometries $\mathrm{PG}(k-1, q)$. This gives us an homogeneuous 1-saturating system of length

$$k \cdot \frac{(q^{mk}-1)(q-1)}{(q^m-1)(q^k-1)}.$$

4.2 Sum-rank $(k-1)$-saturating systems from cutting designs

In this section we introduce the notion of sum-rank metric minimal codes and we investigate their parameters. The geometry of minimal codes have been important in order to construct and give bounds in both Hamming and rank metric, via the so called *strong blocking sets*. These, introduced first in [13] in relation to saturating sets, are sets of points in the projective space such that the intersection with every hyperplane spans the hyperplane. In [18] strong blocking sets are referred to as generator sets and they are constructed as union of disjoint lines. They have gained very recently a renovated interest in coding theory, since [4], in which they are named *cutting blocking sets* and they are used to construct minimal codes. Quite surprisingly, they have been shown to be the geometric counterparts of minimal codes.

Definition 15. *Let $\mathcal{C}$ be an $[\mathbf{n}, k]_{q^m/q}$ sum-rank metric code. A codeword $c \in \mathcal{C}$ is said* minimal *if for every $c' \in \mathcal{C}$ such that $\mathrm{supp}_{\mathbf{n}}(c') \subseteq \mathrm{supp}_{\mathbf{n}}(c)$ then $c' = \lambda c$ for some $\lambda \in \mathbb{F}_{q^m}$. We say that $\mathcal{C}$ is* minimal *if all of its codewords are minimal.*

Definition 16. *A system $\mathcal{U} = (\mathcal{U}_1, \ldots, \mathcal{U}_t) \subset \mathbb{F}_{q^m}^k$ is* cutting *if $L_{\mathcal{U}_1} \cup \ldots \cup L_{\mathcal{U}_t}$ is a* strong blocking set *in $\mathrm{PG}(k-1, q^m)$, that is if*

$$\langle (L_{\mathcal{U}_1} \cup \ldots \cup L_{\mathcal{U}_t}) \cap \mathcal{H} \rangle_{\mathbb{F}_{q^m}} = \mathcal{H},$$

for every hyperplane $\mathcal{H}$ in $\mathrm{PG}(k-1, q^m)$.

The following is a generalization of the geometric characterization of minimal codes in the Hamming and in the rank metric.

Theorem 8 ([24, Corollary 10.25]). *A sum-rank metric code is minimal if and only if an associated system is cutting.*

Theorem 9. *If $\mathcal{U}$ is a cutting system in $\mathbb{F}_{q^m}^k$, then $\mathcal{U}$ is a sum-rank $(k-1)$-saturating system in $\mathbb{F}_{q^{m(k-1)}}^k$.*

Remark 5. In [6], the authors provide interesting bounds and examples about sum-rank minimal codes. Combining them with Theorem 9, it is possible to obtain more examples of saturating systems in the sum-rank metric.

References

1. G. N. Alfarano, M. Borello, and A. Neri. A geometric characterization of minimal codes and their asymptotic performance. *Adv. in Math. Commun.*, 2020.
2. G. N. Alfarano, M. Borello, A. Neri, and A. Ravagnani. Linear cutting blocking sets and minimal codes in the rank metric. *Journal of Combinatorial Theory, Series A*, 2022.
3. D. Bartoli, M. Borello, and G. Marino. Saturating linear sets of minimal rank. *arXiv preprint arXiv:2306.17081*, 2023.
4. M. Bonini and M. Borello. Minimal linear codes arising from blocking sets. *Journal of Algebraic Combinatorics*, 53:327–341, 2021.
5. M. Bonini, M. Borello, and E. Byrne. Saturating systems and the rank-metric covering radius. *Journal of Algebraic Combinatorics*, 58:1173–1202, 2023.
6. M. Borello and F. Zullo. Geometric dual and sum-rank minimal codes. *arXiv preprint arXiv:2303.07288*, 2023.
7. R. Brualdi, V. Pless, and R. Wilson. Short codes with a given covering radius. *IEEE Transactions on Information Theory*, 35(1):99–109, 1989.
8. T. Bu. Partitions of a vector space. *Discrete Mathematics*, 31(1):79–83, 1980.
9. E. Byrne and A. Ravagnani. Covering radius of matrix codes endowed with the rank metric. *SIAM Journal on Discrete Mathematics*, 31(2):927–944, 2017.
10. R. Calderbank and W. M. Kantor. The geometry of two-weight codes. *Bulletin of the London Mathematical Society*, 18(2):97–122, 1986.
11. G. D. Cohen, I. Honkala, S. Litsyn, and A. Lobstein. *Covering Codes*. North-Holland Mathematical Library, 1997.
12. A. A. Davydov. Constructions and families of covering codes and saturated sets of points in projective geometry. *IEEE Transactions on Information Theory*, 41(6):2071–2080, 1995.
13. A. A. Davydov, M. Giulietti, S. Marcugini, and F. Pambianco. Linear nonbinary covering codes and saturating sets in projective spaces. *Advances in Mathematics of Communications*, 5(1):119–147, 2011.
14. A. A. Davydov, S. Marcugini, and F. Pambianco. On saturating sets in projective spaces. *Journal of Combinatorial Theory, Series A*, 103(1):1–15, 2003.
15. A. A. Davydov and P. R. Östergård. On saturating sets in small projective geometries. *European Journal of Combinatorics*, 21(5):563–570, 2000.
16. L. Denaux. Constructing saturating sets in projective spaces using subgeometries. *Designs, Codes and Cryptography*, pages 1–32, 2021.
17. S. Dodunekov and J. Simonis. Codes and projective multisets. *The Electronic Journal of Combinatorics*, 5(1):R37, 1998.
18. S. Fancsali and P. Sziklai. Lines in higgledy-piggledy arrangement. *the electronic journal of combinatorics*, 21, 2014.
19. M. Gadouleau. *Algebraic codes for random linear network coding*. Lehigh University, 2009.
20. J. Hirschfeld. *Projective geometries over finite fields. Oxford Mathematical Monographs*. Oxford University Press New York, 1998.
21. A. Neri, P. Santonastaso, and F. Zullo. The geometry of one-weight codes in the sum-rank metric. *Journal of Combinatorial Theory, Series A*, 194:105703, 2023.
22. C. Ott, H. Liu, and A. Wachter-Zeh. Covering properties of sum-rank metric codes. In *2022 58th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1–7, 2022.
23. T. H. Randrianarisoa. A geometric approach to rank metric codes and a classification of constant weight codes. *Designs, Codes and Cryptography*, 88:1331–1348, 2020.
24. P. Santonastaso and F. Zullo. On subspace designs. *EMS Surveys in Mathematical Sciences*, 2023.

Linear programming lower bounds for energy of weighted spherical codes

S. Borodachov[1], P. G. Boyvalenkov[2], P. D. Dragnev[3], D. P. Hardin[4], E. B. Saff[4], and M. M. Stoyanova[5]

[1] Department of Mathematics, Towson University, 7800 York Rd, Towson, MD, 21252, USA, sborodachov@towson.edu;
[2] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 8 G Bonchev Str., 1113 Sofia, Bulgaria, peter@math.bas.bg;
[3] Department of Mathematical Sciences, Purdue University, Fort Wayne, IN 46805, USA, dragnevp@pfw.edu;
[4] Center for Constructive Approximation, Department of Mathematics, Vanderbilt University, Nashville, TN 37240, USA, doug.hardin@vanderbilt.edu, edward.b.saff@vanderbilt.edu;
[5] Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohridski", 5 James Bourchier Blvd., 1164 Sofia, Bulgaria, stoyanova@fmi.uni-sofia.bg

Abstract. Universal lower bounds for potential energy of weighted spherical codes are obtained by linear programming. The universality is in the sense of Cohn-Kumar – every attaining code (if any) is optimal with respect to a large class of potential functions, in the sense of Levenshtein – there is a bound for every weighted code, and in the sense of parameters (nodes and weights) which do not depend on the potential function.

Keywords: Discrete potentials · linear programming · universal bounds

1 Introduction

A collection (C, W) of distinct points $C = \{x_1, x_2, \ldots, x_N\} \subset \mathbb{S}^{n-1}$, where $\mathbb{S}^{n-1}$ is the unit sphere in $\mathbb{R}^n$, and corresponding weights $W = (w_1, w_2, \ldots, w_N)$, where $w_i > 0$ corresponds to x_i and $w_1 + w_2 + \cdots + w_N = 1$, is called *a weighted spherical code*.

For a continuous function $h : [-1, 1) \to \mathbb{R}$ we consider the *weighted h-energy of C*

$$E_h(C, W) := \sum_{i \neq j} w_i w_j h(x_i \cdot x_j),$$

where $x \cdot y$ is the usual inner product in $\mathbb{R}^n$. Let

$$\mathcal{E}^h(N, W) := \inf_{|C|=N} E_h(C, W)$$

be the *minimum h-energy* among all codes (C, W) with fixed cardinality $|C| = N \geq 2$ and weights set W. Optimization problems for the h-energy arise, for

example, in the electrostatics when it is necessary to distribute $N = |W|$ positive charges (not necessarily equal) on the unit sphere.

We will use the version of the Gegenbauer polynomials $P_i^{(n)}$, $i = 0, 1, \ldots$, orthogonal with respect to the measure $d\mu(t) := \gamma_n(1-t^2)^{(n-3)/2}\,dt$, $t \in [-1, 1]$, where $\gamma_n := \Gamma(\frac{n}{2})/\sqrt{\pi}\Gamma(\frac{n-1}{2})$ is a normalizing constant that makes μ a probability measure, and $P_i^{(n)}(1) = 1$ for normalization. Note that $P_i^{(n)}(t)$ is exactly the Jacobi polynomial $P_i^{(\alpha,\beta)}(t)$ with parameters $\alpha = \beta = (n-3)/2$ and the corresponding normalization.

Given a weighted code (C, W), we consider its (weighted) moments

$$\mathcal{M}_\ell(C, W) := \sum_{i,j=1}^{N} w_i w_j P_\ell^{(n)}(x_i \cdot x_j), \quad \ell \geq 1.$$

It follows from the positive definiteness of the Gegenbauer polynomials that $M_\ell(C, W) \geq 0$ for every positive integer ℓ. The case of equality for some ℓ is especially interesting.

Definition 1. *A weighted spherical code (C, W) is called a* weighted spherical design of strength τ *(or* a weighted spherical τ-design*) if its first τ weighted moments are zero; i.e.,*

$$\mathcal{M}_\ell(C, W) = 0 \text{ for } 1 \leq \ell \leq \tau.$$

In the equi-weighted case $w_1 = \cdots = w_N = 1/N$ one obtains the classical spherical designs introduced in the seminal paper of Delsarte, Goethals, and Seidel [8] from 1977. The weighted case can be traced back to 1960's and 70's when cubature formulas for approximate calculation of multiple integrals on $\mathbb{S}^{n-1}$ were investigated [14, 16, 15, 13, 10].

Utilizing linear programming, we shall obtain lower bounds for the weighted h-energy $\mathcal{E}^h(N, W)$ for all *absolutely monotone potentials* h, that is $h^{(k)}(t) \geq 0$ for every $k \geq 0$. Our bounds are universal in the sense of Levenshtein (there is a bound for *every* weighted code), and in the sense of defining parameters (nodes and weights) which are *independent* of the potential function. Also, assuming existence of attaining codes, the bounds are universal in the sense of Cohn-Kumar (every attaining code is optimal with respect to *all* absolutely monotone potentials). We present examples, where our bounds are very close to the actual weighted energy of certain weighted spherical designs.

Our bounds are derived as certain solutions of linear programs which arise naturally as generalizations of the equi-weighted frameworks from [5]. We present some examples for weighted codes which have attracted attention previously for their high degree of precision as cubature formulas by Sobolev [14], Goethals and Seidel [10] and Waldron [17].

2 A general linear programming lower bound for weighted codes

Given a potential function h, we consider the set of polynomials

$$L_h := \{f(t) = \sum_{i=0}^{\deg(f)} f_i P_i^{(n)}(t) : f(t) \le h(t), t \in [-1,1), f_i \ge 0, i = 1, \dots, \deg(f)\},$$

where $(P_i^{(n)})_{i=0}^{\infty}$ are the Gegenbauer polynomials as defined in the Introduction. The set L_h will be the feasible domain for linear programming bounds for $\mathcal{E}^h(N, W)$.

Theorem 1. *If $f(t) = \sum_{i=0}^{\deg(f)} f_i P_i^{(n)}(t) \in L_h$, then for every weighted (C, W) code on $\mathbb{S}^{n-1}$ with cardinality N*

$$E_h(C, W) \ge E_f(C, W) \ge f_0 - f(1) \sum_{i=1}^{N} w_i^2. \tag{1}$$

Consequently,

$$\mathcal{E}^h(N, W) \ge \sup_{f \in L_h} \left(f_0 - f(1) \sum_{i=1}^{N} w_i^2 \right) =: \mathrm{ULB}(W, h). \tag{2}$$

If the equality is attained in (1) *for some (C, W) and f, then $f(x_i \cdot x_j) = h(x_i \cdot x_j)$ for every $i \ne j$ and $f_\ell M_\ell(C, W) = 0$ for every $\ell \in \{1, 2, \dots, \deg(f)\}$.*

Proof. The first inequality in (1) follows obviously from $f \le h$ in $[-1, 1)$. For the second, we estimate $E_h(C, W)$ from below as follows:

$$\begin{aligned}
E_h(C, W) &= \sum_{i \ne j} w_i w_j h(x_i \cdot x_j) \ge \sum_{i \ne j} w_i w_j f(x_i \cdot x_j) \\
&= \sum_{i,j} w_i w_j f(x_i \cdot x_j) - f(1) \sum_{i=1}^{N} w_i^2 \\
&= \sum_{\ell=0}^{\deg(f)} f_\ell \sum_{i,j} w_i w_j P_\ell^{(n)}(x_i \cdot x_j) - f(1) \sum_{i=1}^{N} w_i^2 \\
&= f_0 + \sum_{\ell=1}^{\deg(f)} f_\ell M_\ell(C, W) - f(1) \sum_{i=1}^{N} w_i^2 \\
&\ge f_0 - f(1) \sum_{i=1}^{N} w_i^2.
\end{aligned}$$

We used that the coefficient in front of f_0 is $\sum_{i,j=1}^{N} w_i w_j = \left(\sum_{i=1}^{N} w_i \right)^2 = 1$, the inequalities $f_\ell \ge 0$ for $i \ge 1$, and $\sum_{i,j} w_i w_j P_\ell^{(n)}(x_i \cdot x_j) \ge 0$ because of the positive definiteness of the Gegenbauer polynomials. The conditions for equality follow immediately from the above. □

Of particular importance is the case when the supremum in (2) is taken over the class of polynomials $L_h \cap \mathcal{P}_\tau$, where $\mathcal{P}_\tau$ denotes the polynomials of degree at most τ. This yields the linear program

$$\begin{aligned} \text{maximize} \quad & f_0 - f(1)\sum_{i=1}^{N} w_i^2 \\ \text{subject to} \quad & f \in L_h \cap \mathcal{P}_\tau. \end{aligned} \tag{3}$$

For particular parameters h, W, and τ we shall obtain explicit solutions of this linear program. We shall denote the maximized objective function by $\mathrm{ULB}_\tau(W, h)$. Note that

$$S(W) := \sum_{i=1}^{N} w_i^2 \geq \frac{1}{N} \tag{4}$$

with equality if and only if $w_1 = w_2 = \cdots = w_N = 1/N$ (i.e., in the classical case of equi-weighted code). Then it follows that

$$f_0 - f(1)\sum_{i=1}^{N} w_i^2 \leq f_0 - \frac{f(1)}{N},$$

where the right-hand side coincides exactly with the quantity which appears in the linear programming for the equi-weighted codes. This means that the bounds from Theorem 1 will be always less than the bounds for the corresponding equi-weighted case. Anyway, it is important to see that the quantity

$$N_W := \frac{1}{S(W)} = 1/\sum_{i=1}^{N} w_i^2$$

has to play an important role since it is going to determine the parameters (nodes and weights) of the universal lower bound in the same way as the cardinality N does in [5]. Clearly, as the weights w_i get closer in value to one another, as measured by the *variance* $\operatorname{var} W := (1/N)S(W) - 1/N^2$ the quantity N_W approaches N from below. The inequality (4), written as $N \geq N_W$, means that N_W is always less than or equal to the cardinality N (with equality only for equal weights) and serves to replace N in the framework from [5].

We introduce the necessary parameters as follows. Assume that

$$D(n, \tau) < N_W \leq D(n, \tau + 1), \tag{5}$$

where $\tau = 2k - 1 + \varepsilon$, $\varepsilon \in \{0, 1\}$ shows the parity of τ, and

$$D(n, \tau) := \binom{n + k - 2 + \varepsilon}{n - 1} + \binom{n + k - 2}{n - 1}$$

is the Delsarte-Goethals-Seidel bound [8]. The numbers $D(n, \tau)$, $\tau = 1, 2, 3, \ldots$, define a partition of the positive integers into consecutive intervals. It is not necessary to have N_W and N in the same interval $(D(n, \tau), D(n, \tau + 1)]$ but examples below will suggest that better bounds are obtained for closer values of N_W and N.

3 Universal lower bound for weighted codes

Let the parameters $(\alpha_i, \rho_i)_{i=0}^{k-i+\varepsilon}$ be determined (see the explanations in the next paragraph) by the equation

$$L_\tau(n, s) = N_W,\ \tau = 2k - 1 + \varepsilon,\ \varepsilon \in \{0, 1\}, \tag{6}$$

where $L_\tau(n, s)$ is the Levenshtein bound (see [12, Section 6]). The Levenshtein bound $L_\tau(n, s)$ is valid (and optimal in a sense) in the interval $s \in \left[t_{k-1+\varepsilon}^{1,1-\varepsilon}, t_k^{1,\varepsilon}\right]$, where $t_k^{a,b}$ is the largest zero of the Jacobi polynomial $P_k^{(a+(n-3)/2, b+(n-3)/2)}(t)$, $i \geq 1$, $t_0^{1,1} = -1$ by definition.

The numbers $(\alpha_i)_{i=0}^{k-i+\varepsilon}$ are the roots of the equation (6) w.r.t. s taking into account that the largest root $\alpha_{k-1+\varepsilon}$ equals s and $\alpha_0 = -1$ whenever $\varepsilon = 1$ (this is for even $\tau = 2k$). Then the weights $(\rho_i)_{i=0}^{k-i+\varepsilon}$ are computed by plugging in the quadrature formula (7) (see the next paragraph) the Lagrange basis polynomials $\ell_i(t) = \prod_{j \neq i}(t - \alpha_j)$ for $i = 0, 1, \ldots, k - 1 + \varepsilon$. Explicit formulas for $(\rho_i)_{i=0}^{k-i+\varepsilon}$ in the case $\varepsilon = 0$ (this is for odd $\tau = 2k - 1$) were found in [4, Appendix A4]. Note also the identity $\sum_{i=0}^{k-1+\varepsilon} \rho_i = 1 - 1/N_W$ which is obtained via plugging $f(t) = 1$ in (7).

It is instrumental for our approach that (see [12, Theorem 5.39]) the quadrature formula (it is a $1/N_W$-quadrature rule in the framework from [5])

$$f_0 = \frac{f(1)}{L_\tau(n, s)} + \sum_{i=0}^{k-1+\varepsilon} \rho_i f(\alpha_i) = \frac{f(1)}{N_W} + \sum_{i=0}^{k-1+\varepsilon} \rho_i f(\alpha_i) \tag{7}$$

holds true for every polynomial $f(t) = f_0 + \sum_{i=1}^{\deg(f)} f_i P_i^{(n)}(t)$ of degree at most $2k - 1 + \varepsilon$.

Like in the equi-weighted case (see [5, 6] and [12]) we will need two facts from the theory of orthogonal polynomials. Namely, the Gegenbauer expansions of the polynomials $P_i^{(n)}(t)P_j^{(n)}(t)$ and $(t + 1)P_i^{(n+2)}(t)P_j^{(n+2)}(t)$ have nonnegative coefficients for every i, j. These properties are called Krein conditions and strengthened Krein conditions, respectively.

We are now in a position to solve the linear program (3).

Theorem 2. *(ULB for weighted codes) Let N and W be such that (5) is satisfied. Let h be absolutely monotone. Then*

$$\mathcal{E}^h(N, W) \geq \mathrm{ULB}_\tau(W, h) := \sum_{i=0}^{k-1+\varepsilon} \rho_i h(\alpha_i),$$

where the parameters $(\alpha_i, \rho_i)_{i=0}^{k-1+\varepsilon}$ are defined as above. This bound can not be improved by any polynomial from $L_h \cap \mathcal{P}_\tau$.

Proof. Let f be the unique Hermite interpolant to h at the nodes $(\alpha_i)_{i=0}^{k-i+\varepsilon}$ counted twice except for the case $\alpha_0 = -1$ (equivalent to $\tau = 2k$) which is

counted once. Then $\deg(f) \le \tau$, so (7) along with the interpolation conditions $f(\alpha_i) = h(\alpha_i)$ yields

$$f_0 - f(1)\sum_{i=1}^{N} w_i^2 = \sum_{i=0}^{k-1+\varepsilon} \rho_i f(\alpha_i) = \sum_{i=0}^{k-1+\varepsilon} \rho_i h(\alpha_i).$$

Moreover, it follows from the Rolle's Theorem (or from the error formula for the Hermite interpolation) that $f(t) \le h(t)$ for every $t \in [-1, 1)$.

Let $\varepsilon = 0$. Order the multiset of nodes as

$$(\alpha_0, \alpha_0, \alpha_1, \alpha_1, \ldots, \alpha_{k-1}, \alpha_{k-1}) = (t_1, t_2, \ldots, t_{2k-1}, t_{2k})$$

(i.e., $t_{2i+1} = t_{2i+2} = \alpha_i$ for $i = 0, 1, \ldots, k-1$; we need to make difference between the first and the second α_0, etc.). Then the Newton interpolation formula

$$f(t) = h(t_1) + \sum_{r=1}^{2k-1} h[t_1, \ldots, t_{r+1}] \prod_{j=1}^{r} (t - t_j)$$

(see, for example, [3]) implies that the polynomial f is a nonnegative linear combination of the constant 1 (obtained when $m = 0$) and the partial products

$$\prod_{j=1}^{m} (t - t_j), \ m = 1, 2, \ldots, 2k-1. \tag{8}$$

It follows from [7, Theorem 3.1] that all polynomials $(t-\alpha_0)(t-\alpha_1)\ldots(t-\alpha_i)$, $i = 0, 1, \ldots, k-2$, expand in the system $\{P_i^{((n-1)/2,(n-3)/2)}(t)\}$ with nonnegative coefficients. Since every polynomial $P_i^{((n-1)/2,(n-3)/2)}(t)$ is positive definite (this follows directly from the Christoffel-Darboux formula which relates this polynomial to the Gegenbauer polynomials), the Krein condition implies that all partial products (8) with $m \le 2k-2$ are positive definite. The only remaining partial product (with $m = 2k-1$ in (8)) is exactly the Levenshtein polynomial $f_{2k-1}^{(n,s)}(t)$ which is positive definite as well (see, for example, [12, Theorem 5.42]). Therefore f is positive definite.

The case $\varepsilon = 1$ is dealt similarly by using the strengthened Krein condition.

If $g(t) = \sum_{i=0}^{\deg(g)} g_i P_i^{(n)}(t)$ is a polynomial from $L_h \cap \mathcal{P}_\tau$, then (7) can be applied to see that the bound of g is

$$g_0 - g(1)\sum_{i=1}^{N} w_i^2 = \sum_{i=0}^{k-1+\varepsilon} \rho_i g(\alpha_i) \le \sum_{i=0}^{k-1+\varepsilon} \rho_i h(\alpha_i) = f_0 - f(1)\sum_{i=1}^{N} w_i^2 = \mathrm{ULB}_\tau(W, h),$$

which completes the proof. □

We next establish the monotonicity of $\mathrm{ULB}_\tau(W, h)$ in N_W.

Theorem 3. *Let $V = \{v_1, \ldots, v_N\}$ and $W = \{w_1, \ldots, w_N\}$ be two sets of positive weights such that $\sum_{i=1}^{N} v_i = \sum_{i=1}^{N} w_i = 1$, and suppose that $N_V < N_W$*

(equivalent to $S(V) > S(W)$). Let η and τ be the positive integers associated with V and W, respectively, via (5). Then $\tau \geq \eta$ and $\mathrm{ULB}_\tau(W,h) > \mathrm{ULB}_\eta(V,h)$. If $\tau = \eta$, then the nodes $(\alpha_i)_{i=0}^{k-1+\varepsilon}$ for N_W are strictly greater than the corresponding nodes for N_V.

Proof. The inequality $N_V < N_W$ implies via (5) that $\tau \geq \eta$. If the equality holds, then (6) and the monotonicity of the Levenshtein function $L(n,s)$ imply the monotonicity of the nodes α_i; i.e. these are increasing with $s = \alpha_{k-1+\varepsilon}$ which is increasing with N_W (see [4]).

Let f and g be the (unique) polynomial solutions of (3) associated with W and V, respectively. Then, as $g \in L_h \cap \mathcal{P}_\eta \subset L_h \cap \mathcal{P}_\tau$ the optimality of f over $L_h \cap \mathcal{P}_\tau$ yields

$$\mathrm{ULB}_\tau(W,h) = f_0 - \frac{f(1)}{N_W} \geq g_0 - \frac{g(1)}{N_W} > g_0 - \frac{g(1)}{N_V} = \mathrm{ULB}_\eta(V,h).$$

Note that g has positive Gegenbauer coefficients and $g(1) = g_0 + \cdots + g_\eta > 0$. □

The potential $h(t) = -\sqrt{2(1-t)}$ fits in the above scheme because $2 + h(t)$ is absolutely monotone. This potential corresponds to the Fejes Tóth problem[6] [9] and it has been studied by many authors (see, for example, [2, 1] and references therein). The degrees 1-3 ULB for weighted codes and this particular potential (and their asymptotic consequences) can be extracted from [1] simply by replacing N by N_W and dividing by N_W^2.

4 Examples

In contrast to difficulties for derivation of more explicit analytic expressions of ULB_τ for $\tau \geq 3$, the numerical calculations of bounds for given n, N, and W can be easily programmed. In this subsection we present examples, where the ULB and the actual weighted energy are computed.

Example 1. Let $C_{32} \subset \mathbb{S}^2$ consist of the 12 vertices of an icosahedron, each of weight $w_I = 20/(21 \cdot 32) = 5/168$, and the 20 vertices of a dodecahedron, each of weight $w_D = 36/(35 \cdot 32) = 9/280$. The vertices of the icosahedron are the centers of the spherical caps defined by the twelve faces of the dodecahedron. In geometry, this is called pentakis dodecahedron or kisdodecahedron. Note that (C_{32}, W) is a weighted spherical 9-design (see [10, Section 5], [11, Example 3.6]).

We proceed with computations of the actual weighted energy of (C_{32}, W) and the corresponding $\mathrm{ULB}_9(W,h)$ for the potential function $h(t) = 1/\sqrt{2(1-t)}$.

The weighted energy of (C_{32}, W) is computed from the information about its structure from Table 1. There are two types of points – I and D, respectively, according to whether they belong to the icosahedron or the dodecahedron, which define the two different distance distributions (the last two rows of Table 1). We set $a := \sqrt{1 - 2/\sqrt{5}}/\sqrt{3}$ and $b := \sqrt{1 + 2/\sqrt{5}}/\sqrt{3}$ to shorten the notation.

[6] Other famous absolutely monotone potentials are named after Riesz, Newton, Gauss, etc.

Table 1. Structure of (C, W).

	Inner products					
	-1	$\pm 1/\sqrt{5}$	$\pm a$	$\pm b$	$\pm 1/3$	$\pm\sqrt{5}/3$
Type	Number of points					
I	1	5	5	5	0	0
D	1	0	3	3	6	3

Therefore,

$$\begin{aligned} E_h(C_{32}, W) = \sum_{i \neq j} w_i w_j h(x_i \cdot x_j) &= 12 w_I^2 \left(h(-1) + 5h(-1/\sqrt{5}) + 5h(1/\sqrt{5}) \right) \\ &+ 120 w_I w_D \left(h(a) + h(-a) + h(b) + h(-b) \right) \\ &+ 20 w_D^2 \left(h(-1) + 6h(-1/3) + 6h(1/3) + 3h(-\sqrt{5}/3) + 3h(\sqrt{5}/3) \right) \\ &\approx 0.8050318. \end{aligned}$$

We have $N_W = 1/\sum_{i=1}^{32} w_i^2 \approx 31.9565217$ which is close to the cardinality 32 of C_{32}. We compute the ULB for $n = 3$, N_W, and $h(t) = 1/\sqrt{2(1-t)}$. Since both $N = 32$ and N_W belong to the interval $(D(3,9), D(3,10)] = (30, 36]$, we have $\tau = 9$ and solve the equation $L_9(3, s) = N_W$ to derive the parameters $(\alpha_i, \rho_i)_{i=0}^4$ as shown approximately in Table 2.

Table 2. Parameters $(\alpha_i, \rho_i)_{i=0}^4$ for $(n, N, N_W) = (3, 32, \approx 31.9565)$.

i	0	1	2	3	4
α_i	-0.941	-0.674	-0.2109	0.328	0.779
ρ_i	0.077	0.1889	0.2636	0.261	0.17777

Therefore, $\mathcal{E}^h(32, W) \geq \mathrm{ULB}_9(W, h) = \sum_{i=0}^4 \rho_i h(\alpha_i) \approx 0.804786$, which is very close to the actual h-energy ≈ 0.8050318 of (C_{32}, W).

Example 2. We consider a weighted union C_{cp} of a cube and a cross-polytope on $\mathbb{S}^{n-1}$ defined by their duality, i.e. each pair of antipodal vertices of the cross-polytope defines an symmetry axis of two opposite faces of the cube. Each point of the cross-polytope has weight $w_p := 1/(2n + n^2)$ and each point of the cube has weight $w_c := n^2/2^n(2n + n^2)$. It is easy to see that the sum of weights of the union is 1 and, furthermore, C_{cp} is a weighted spherical 5-design on $\mathbb{S}^{n-1}$.

In small dimensions, the codes C_{cp} look as follows. On $\mathbb{S}^2$, each point of the cross-polytope will have weight $1/15$ and each point of the cube will have weight $3/40$, giving a weighted spherical 5-design of 14 points; on $\mathbb{S}^3$, each point of the cross-polytope will have weight $1/24$ and each point of the cube will also have weight $1/24$ (we get a 24-cell, a equi-weighted spherical 5-design).

For any h, the actual h-energy of C_{cp} is

$$E_h(C_{cp},W)=2nw_p^2\left(h(-1)+(2n-2)h(0)\right)+2^{n+1}nw_pw_c\left(h\left(\frac{1}{\sqrt{n}}\right)+h\left(-\frac{1}{\sqrt{n}}\right)\right)$$
$$+2^nw_c^2\sum_{k=0}^{n-1}\binom{n}{k}h\left(-1+\frac{2k}{n}\right).$$

Table 3. Approximate parameters and ULB for $(n,N,N_W)=(n,2n+2^n,N_W)$, $2\le n\le 7$, and $h(t)=(2(1-t))^{-(n-2)/2}$ (the Newton potential).

n	N_W	N	(α_i)	(ρ_i)	ULB	Energy of (C_{qp},W)
2	8	8	-1 $-\sqrt{2}/2$ 0 $\sqrt{2}/2$	1/8 1/4 1/4 1/4	0.875	0.875
3	13.95	14	−0.8580 −0.2701 0.5225	0.1832 0.3832 0.3618	0.7058	0.7070
4	24	24	−0.8173 −0.2575 0.4749	0.1384 0.4339 0.3858	0.5781	0.5798
5	41.48	42	−0.7428 −0.1910 0.4684	0.1424 0.4680 0.3653	0.4825	0.4901
6	71.44	76	−0.6753 −0.1327 0.4705	0.1540 0.4996 0.3323	0.4074	0.4314
7	121.16	142	-1 −0.5936 −0.0772 0.4748	0.0022 0.1785 0.5165 0.2944	0.3462	0.3993

The $\mathrm{ULB}_5(W,h)$ for corresponding parameters $(n,|C_{cp}|=2n+2^n,N_W)$, where

$$N_W=\frac{1}{\sum_{i=1}^{2n+2^n}w_i^2}=\frac{1}{2nw_p^2+2^nw_c^2}=\frac{n(n+2)^2 2^n}{n^3+2^{n+1}},$$

can be computed as follows. We solve $L_5(n,s)=N_W$, that is

$$\frac{\left((n+2)(n+3)s^2+4(n+2)s-n+1\right)(1-s)}{2s\left(3-(n+2)s^2\right)}=\frac{(n+2)^2 2^n}{n^3+2^{n+1}}$$

to obtain the nodes $(\alpha_i)_{i=0}^2$. Then the quadrature weights $(\rho_i)_{i=0}^2$ are computed by setting the Lagrange basis polynomials in (7). The ULB in dimensions $2\le n\le 7$, calculated for the Newton potential $h(t)=1/(2(1-t))^{(n-2)/2}$, are shown

in the sixth column of Table 3. It is $\mathrm{ULB}_7(W,h)$ for $n=2$, $\mathrm{ULB}_5(W,h)$ for $3\le n\le 6$ and $\mathrm{ULB}_6(W,h)$ for $n=7$. Note that the bound $\mathrm{ULB}_7(W,h)$ is attaned for $n=2$, where it coincides with the ULB for the equi-weighted case [5] (recall that the attaining (C_{qp},W) is an equi-weighted regular 8-gon).

Acknowledgments. The research of the second author was supported, in part, by Bulgarian NSF grant KP-06-N72/6-2023. The research of the third author is supported in part by the Lilly Endowment. The research of the sixth author was supported, in part, by Contract BG-RRP-2.004-0008, Sofia University Marking Momentum for Innovation and Technological Transfer (SUMMIT), Work group 3.2.1. Numerical Analysis, Theory of Approximations and Their Applications (NATATA).

References

1. Barg, A., Boyvalenkov, P., Stoyanova, M., Bounds for the sum of distances in spherical sets of small size, Discr. Math., **346**, art. 113346 (2023).
2. Bilyk, D., Matzke, R. W., On the Fejes Tóth problem about the sum of angles between lines, Proc. Am. Math. Soc. **147**, 51–59 (2019).
3. de Boor, C., Divided differences, Surveys in Approximation Theory **1**, 46–69 (2005).
4. Boyvalenkov, P., Danev, D., Landgev, I., On maximal spherical codes II, J. Combin. Designs **7**, 316–326 (1999).
5. Boyvalenkov, P., Dragnev, P., Hardin, D., Saff, E., Stoyanova, M., Universal lower bounds for potential energy of spherical codes, Constr. Approx. **44**, 385–415 (2016).
6. Boyvalenkov, P., Dragnev, P., Hardin, D., Saff, E., Stoyanova, M., Energy bounds for codes in polynomial metric spaces, Anal. Math. Phys. **9**, 781–808 (2019).
7. Cohn, H., Kumar, A., Universally optimal distribution of points on spheres, J. Amer. Math. Soc. **20**, 99–148 (2007).
8. Delsarte, P., Goethals, J.-M., Seidel, J. J., Spherical codes and designs, Geom. Dedic. **6**, 363–388 (1977).
9. Fejes Tóth, L., On the sum of distances determined by a pointset, Acta Math. Acad. Sci. Hung. **7**, 397–401 (1956).
10. Goethals, J.M., Seidel, J.J. . Cubature Formulae, Polytopes, and Spherical Designs. In: Davis, C., Grünbaum, B., Sherk, F.A. (eds) The Geometric Vein, 204–218, Springer, New York (1981).
11. Hughes, D., Waldron, S., Spherical (t,t)-designs with a small number of vectors, Lin. Alg. Appl. **608**, 84–106 (2021).
12. Levenshtein, V. I., Universal bounds for codes and designs, in Handbook of Coding Theory, V. S. Pless and W. C. Huffman, Eds., Elsevier, Amsterdam, Ch. 6, 499–648 (1998).
13. Salihov, G. N., On the theory of cubature formulas for multidimensional spheres. Dissertation, Acad. Sci. USSR, Novosibirsk 1978 (in Russian).
14. Sobolev, S. L., Cubature formulas on the sphere invariant under finite groups of rotations, Dokl. Akad. Nauk SSSR 146, 310–313 (1962); English translation Soviet Math. Dokl. 3, 1307–1310 (1962).
15. Sobolev, S. L., Introduction to the Theory of Cubature Formulas (Russian). Nauka 1974.
16. Stroud, A. H., Approximate Calculation of Multiple Integrals. Prentice-Hall 1971.
17. Waldron S., An introduction to finite tight frames, New York, Applied and Numerical Harmonic Analysis, Birkhaüser/Springer, 2018.

Optimal s-boxes against alternative operations

Marco Calderini[1], Roberto Civino[2], and Riccardo Invernizzi[3]

[1] University of Trento, marco.calderini@unitn.it
[2] University of L'Aquila, roberto.civino@univaq.it
[3] KU Leuven, riccardo.invernizzi@kuleuven.be

Abstract. Civino et al. have characterised diffusion layers that expose an SPN to vulnerability from differential cryptanalysis when employing alternative operations coming from groups isomorphic to the translation group on the message space. In this study, we present a classification of diffusion layers that exhibit linearity in *parallel* alternative operations for ciphers with 4-bit s-boxes, enabling the possibility of an alternative differential attack simultaneously targeting all the s-boxes within the block. Furthermore, we investigate the differential behaviour with respect to alternative operations for all classes of optimal 4-bit s-boxes, as defined by Leander and Poschmann (2007). Our examination reveals that certain classes contain weak permutations w.r.t. alternative differential attacks, and we leverage these vulnerabilities to execute a series of experiments.

Keywords: Differential cryptanalysis · Alternative operations · 4-bit s-boxes

1 Introduction and preliminaries

Differential cryptanalysis, originally introduced by Biham and Shamir in the late 1980s [4] and subsequently generalised [3, 6, 11, 16], has become one of the cornerstones for evaluating the robustness of various symmetric primitives. The fundamental premise of differential cryptanalysis is that analysing the differences (differentials) between pairs of plaintexts and the corresponding ciphertexts can unveil undesired biases. While differentials can be calculated with respect to any difference operator, regardless of which operation is responsible for performing the sum with the round key during encryption, it is usual for the two operations to coincide. For this reason, classical differential cryptanalysis of a cipher in which the key is xor-ed to the state is typically performed by studying the distribution of xor-differentials, whose propagation is traditionally prevented by the combined action of the linear diffusion layer and the s-box layer. In particular, s-boxes are pivotal for ensuring the security of almost all contemporary block ciphers, serving as the primary non-linear component within the cipher, particularly in the case of SPNs. Equally relevant, the efficiency of a cipher is significantly influenced by the size of the s-boxes. In practical scenarios, s-boxes typically have a size of 4 or 8 bits, with 4 being the most popular choice for ciphers designed to operate on power-constrained devices [1, 2, 5, 14]. It is clear that the selection of appropriate

s-boxes is critical to fortify the cipher against various types of attacks. In this sense, Leander and Poschmann have classified 4-bit s-boxes which are optimal w.r.t. standard criteria that guarantee poor propagation of xor-differentials [12].

A recent line of research is focused on the study of alternative difference operators for the differential cryptanalysis of xor-based ciphers [7, 8, 10, 15]. These new operators are designed to induce a novel operation with respect to which differentials are computed. Within this approach, a large class of possible alternative operations has been studied, all of which have in common that they are induced by a group of translations isomorphic to the group of translations acting on the message space by means of the xor addition with the key. In the context of an SPN, where the encrypted message is generated by iterating through a sequence of s-box layers, (xor)-linear diffusion, and xor-based key addition layers, altering the differential operator yields a dual impact. On one hand, it is highly probable that differentials traverse the s-box layer more effectively, given that its non-linearity is maximised with respect to xor. On the other hand, differentials do not deterministically propagate through the diffusion layer, as observed in classical scenarios. This pivotal limitation effectively restricts the success of the attack only to cases where the target layer is linear not only concerning xor but also with respect to the operation under consideration for computing differentials.

A first successful attempt based on the study of the alternative differential properties of a xor-based toy cipher of the SPN family has shown that it is possible to highlight a bias in the distribution of the differences calculated compared to an alternative operation which is instead not detectable by means of the standard xor-differential-based approach [10]. The target cipher featured five 3-bit s-boxes and the operation used to perform the attack acted as the xor on the last four s-boxes, while on the first one matched with one of the alternative sums defined by Calderini et al. [8], coming from another translation groups. The advantage of employing an alternative operation in this case was only derived from the benefit induced by a single s-box. In a more recent experimental approach [7], we showed that better results in a similar context can be obtained using an *alternative parallel operation*, in which every s-box can be targeted. In this case, the diffusion layer of the cipher was determined through an algorithm, ensuring that it adheres to the constraint of linearity with respect to both xor and the target operation.

In this paper, we establish a general result that, in the context of an SPN with 4-bit s-boxes, characterises all xor-linear maps that are concurrently linear with respect to a parallel alternative operation (Sec. 2). This finding enables the execution of a differential attack wherein each s-box affected by a non-trivial differential contributes to the final differential probability with increased efficacy compared to the conventional xor differentials. Additionally, differentials propagate deterministically through the linear layer in this scenario. Moreover, we examine all possible alternative operations on 4 bits and investigate the differential properties of optimal 4-bit s-boxes, following the classification outlined by Leander and Poschmann (a comparable methodology, albeit in the context of modular addition, was recently employed by Zajac and Jókay [17]). Our analysis

demonstrates that each class comprises potentially weak permutations (Sec. 3). When coupled with a diffusion layer as described earlier, these permutations have the potential to render the cipher susceptible to differential attacks with alternative operations. To substantiate our findings, we conclude the paper by presenting experimental results on a family of toy SPNs (Sec. 4).

1.1 Notation

Let V be an n-dimensional vector space over $\mathbb{F}_2$ which represents the message space. We write $V = V_1 \oplus V_2 \oplus \cdots \oplus V_b$, where each V_j is isomorphic to a vector space B such that $\dim(B) = s$ on which every s-box acts. Thefore we have $n = sb$. We denote by $\{e_i\}_{i=1}^n$ the canonical basis of V. If G is any finite group acting on V, for each $g \in G$ and $v \in V$ we denote the action of g on v as vg, i.e. we use postfix notation for every function evaluation. We denote by $\mathrm{Sym}(V)$ the symmetric group acting on V, i.e. the group of all permutation on the message space, by $\mathrm{GL}(V,+)$ the group of linear transformations, and by $\mathrm{AGL}(V,+)$ the group of affine permutations. The identity matrix of size l is denoted by $\mathbb{1}_l$ and the zero matrix of size $l \times h$ is denoted by $\mathbb{0}_{l,h}$, or simply $\mathbb{0}_l$ if $l = h$. We finally denote by T_+ the group of translations on V, i.e. $T_+ := \{\sigma_a \mid a \in V, x \mapsto x + a\} < \mathrm{Sym}(V)$. We remind that the translation σ_k acts on a vector x in the same way the key-addition layer of an SPN acts xor-ing the round key k to the message x, i.e. $x\sigma_k = x + k$.

1.2 Preliminaries on alternative operations

An alternative operation on V can be defined given any 2-elementary abelian regular subgroup $T < \mathrm{AGL}(V,+)$, that we can write as $T = \{\tau_a \mid a \in V\}$, where τ_a is the unique element in T which maps 0 into a. Consequently, for all $a, b \in$V, we can define $a \circ b := a\tau_b$, resulting in $(V, \circ)$ forming an additive group. The operation $\circ$ induces a vector space structure on V, with the corresponding group of translation being $T_\circ = T$. Additionally, for each $a \in V$, there exists $M_a \in \mathrm{GL}(V,+)$ such that $\tau_a = M_a\sigma_a$, meaning that for every $x \in V$,

$$x \circ a = x\tau_a = xM_a + a.$$

It is also assumed throughout that $T_+ < \mathrm{AGL}(V, \circ)$, where $\mathrm{AGL}(V, \circ)$ is the normaliser in $\mathrm{Sym}(V)$ of $T_\circ$ (i.e., the group of affine permutations w.r.t. $\circ$). This crucial technical assumption renders the key-addition layer an affine operator concerning the new operation, enabling the prediction of how the key addition affects the differentials with a reasonable probability. Further details on this aspect, which may not be directly relevant to the scope of the current paper, can be found in Civino et al. [10]. In this context, we define the *weak keys subspace* as

$$W_\circ := \{a \mid a \in V, \sigma_a = \tau_a\} = \{k \mid k \in V,\ \forall x \in V\ x \circ k = x + k\}.$$

$W_\circ$ is a vector subspace of both $(V,+)$ and $(V,\circ)$. It is known [8, 9] that $W_\circ$ is non empty and that

$$2 - (n \bmod 2) \leq \dim(W_\circ) \leq n - 2. \tag{1}$$

Moreover, up to conjugation we can always assume $W_\circ$ to be the span of the last d canonical vectors of V [8]. This allows to represent the new sum in a canonical way [8]: for each $a \in V$ there exists a matrix $E_a \in \mathbb{F}_2^{(n-d)\times d}$ such that

$$M_a = \begin{pmatrix} \mathbb{1}_{n-d} & E_a \\ \mathbb{0}_{d,n-d} & \mathbb{1}_d \end{pmatrix}. \tag{2}$$

Fixing such an operation as above is therefore equivalent to defining the matrices

$$M_{e_i} = \begin{pmatrix} \mathbb{1}_{n-d} & E_{e_i} \\ \mathbb{0}_{d,n-d} & \mathbb{1}_d \end{pmatrix} = \left(\begin{array}{c|c} & \mathbf{b}_{i,1} \\ \mathbb{1}_{n-d} & \vdots \\ & \mathbf{b}_{i,n-d} \\ \hline \mathbb{0}_{d,n-d} & \mathbb{1}_d \end{array}\right)$$

for $1 \le i \le n$, where $\mathbf{b}_{i,j} \in \mathbb{F}_2^d$. The assumptions on $T_\circ$ and on $W_\circ$ imply that $E_{e_i} = 0$ for $n-d+1 \le i \le n$, $\mathbf{b}_{i,i} = \mathbf{0}$ and $\mathbf{b}_{i,j} = \mathbf{b}_{j,i}$. In conclusion, the following result characterises the criteria that the vectors $\mathbf{b}_{i,j}$ must adhere to in order to define an alternative operation as previously described.

Theorem 1 ([10]). *Let $T_\circ < \mathrm{AGL}(V,+)$ be 2-elementary, abelian, and regular, and let $d \le n-2$. The operation $\circ$ induced by $T_\circ$ is such that $d = \dim(W_\circ)$, $T_+ < \mathrm{AGL}(V,\circ)$, and $W_\circ = \mathrm{Span}\{e_{n-d+1},\dots,e_n\}$ if and only if the matrix $\Theta_\circ \in (\mathbb{F}_{2^d})^{(n-d)\times(n-d)}$ defined as*

$$\Theta_\circ := \begin{pmatrix} \mathbf{b}_{1,1} & \mathbf{b}_{1,2} & \cdots & \mathbf{b}_{n-d,1} \\ \mathbf{b}_{2,1} & \mathbf{b}_{2,2} & \cdots & \mathbf{b}_{n-d,2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{b}_{n-d,1} & \mathbf{b}_{n-d,2} & \cdots & \mathbf{b}_{n-d,d} \end{pmatrix}$$

is zero-diagonal, symmetric and no $\mathbb{F}_2$-linear combination of its columns is the null vector. The matrix $\Theta_\circ$ is also called the defining matrix *for $\circ$.*

In the subsequent discussion, the term *alternative operation* refers to an additive law $\circ$ on V as defined above.

2 Parallel operations and their automorphism groups

Let $\circ$ be an alternative operation on the block-sized space V. As outlined in the introduction, if $\lambda \in \mathrm{GL}(V,+)$ represents a (xor)-linear diffusion layer, and $\Delta \in V$ is an input difference traversing λ, predicting the output difference with respect to $\circ$, i.e.,

$$x\lambda \circ (x \circ \Delta)\lambda,$$

becomes inherently challenging without additional assumptions on λ that ensure a sufficiently high predictive probability. For this reason, the examination of the following object becomes crucial: in cryptographic terms, it contains potential diffusion layers that allow differentials, whether computed with respect to xor or $\circ$, to propagate with a probability of 1.

Definition 1. *Let $\circ$ be an alternative operation on V. Let us define*

$$H_\circ := \{f \in \mathrm{GL}(V,+) \mid \forall a,b \in V : \ (a \circ b)f = af \circ bf\}$$

to be the subgroup of $\mathrm{GL}(V,+)$ *of permutations that are linear w.r.t. the operation* $\circ$*. More precisely, denoting by* $\mathrm{AGL}(V,\circ)$ *the normaliser in* $\mathrm{Sym}(V)$ *of* $T_\circ$ *and by* $\mathrm{GL}(V,\circ)$ *the stabiliser of* 0 *in* $\mathrm{AGL}(V,\circ)$*, we have* $H_\circ = \mathrm{GL}(V,+) \cap \mathrm{GL}(V,\circ)$*.*

The structure of the group $H_\circ$ in its most general case has not been understood yet. This work addresses this challenge in a specific scenario, guided by assumptions that are deemed reasonable within the context of differential cryptanalysis.

Assumption 1: $\circ$ is a parallel operation. While the operation $\circ$ could, in theory, be defined on the entire message space V, studying the differential properties of the s-box layer, considered as a function with 2^n inputs, is impractical for standard-size ciphers. For this reason, we focus on operations applied in a *parallel* way to each s-box-sized block, i.e., $\circ = (\circ_1, \circ_2, \ldots, \circ_b)$, where for each $1 \le j \le b$, $\circ_j$ is an operation on V_j. In this scenario, every operation is acting independently on the s-box space B, regardless of the others. This motivates the following definition.

Definition 2. *Let $\circ$ be an alternative operation on V. We say that $\circ$ is* parallel *if for each $1 \le j \le b$ there exists an alternative operation $\circ_j$ on V_j such that for each $x, y \in V$ we have*

$$x \circ y = \begin{pmatrix} x_1 \\ \vdots \\ x_b \end{pmatrix} \circ \begin{pmatrix} y_1 \\ \vdots \\ y_b \end{pmatrix} = \begin{pmatrix} x_1 \circ_1 y_1 \\ \vdots \\ x_b \circ_b y_b \end{pmatrix},$$

where $x = (x_1, x_2, \ldots, x_b), y = (y_1, y_2, \ldots, y_b)$ and each component belongs to the s-box-sized space, i.e., $x_j, y_j \in V_j \cong B$ for $1 \le j \le b$.

In the notation of Sec. 1.2, up to a block matrix conjugation, we can assume that every element $x \in V$ is associated to a translation $\tau_x = M_x \sigma_x$, with

$$M_x = \begin{pmatrix} M_{x_1}^{\circ_1} & \cdots & \mathbb{0} \\ \vdots & \ddots & \vdots \\ \mathbb{0} & \cdots & M_{x_b}^{\circ_b} \end{pmatrix}$$

where $M_{x_i}^{\circ_i}$ is the matrix associated to the translation τ_{x_i} with respect to the sum $\circ_i$, as defined in Eq. (2). Notice that it can be assumed, without loss of generality, that all the operations $\circ_j$ coincide.

Assumption 2: $\dim(W_{\circ_j}) = s - 2$. According to Eq. (1), every operation $\circ_j$ defined at the s-box level must satisfy the bound $\dim(W_{\circ_j}) \le s - 2$, being $s = \dim(B)$. The situation where the (upper) bound is reached holds particular interest for several reasons, as elaborated further in Civino et al. [10]. Notably,

- if the s-box size s is four, the case where $\dim(W_{\circ_j}) = 2$ is the sole possibility;
- the considered case stands today as the only one for which the structure of $H_{\circ_j}$ is well understood.

For the reader's convenience, we present the classification result for $H_{\circ_j}$ obtained by Civino et al. in the considered case. Additionally, it is worth recalling that, according to Theorem 1, any $\circ_j$ for which $\dim(W_{\circ_j}) = s-2$ is determined by a single non-null vector $\mathbf{b} \in (\mathbb{F}_2)^{s-2}$.

Theorem 2 ([10]). *Let $\circ_j$ be an alternative operation such that $d = \dim(W_{\circ_j}) = s-2$ defined by a vector $\boldsymbol{b} \in (\mathbb{F}_2)^{s-2}$, and let $\lambda \in (\mathbb{F}_2)^{s\times s}$. The following are equivalent:*

- *$\lambda \in H_{\circ_j}$;*
- *there exist $A \in \mathrm{GL}((\mathbb{F}_2)^2, +)$, $D \in \mathrm{GL}((\mathbb{F}_2)^d, +)$, and $B \in (\mathbb{F}_2)^{2\times d}$ such that*

$$\lambda = \begin{pmatrix} A & B \\ \mathbb{0}_{d,2} & D \end{pmatrix}$$

and $\boldsymbol{b}D = \boldsymbol{b}$.

We are now prepared to present the first novel contribution of this work, wherein we characterise the group $H_\circ$ for a parallel operation $\circ = (\circ_1, \circ_2, \ldots, \circ_b)$ with components at the s-box level satisfying $\dim(W_{\circ_j}) = s-2$. For the sake of simplicity and without losing generaly, we assume that the b operations at the s-box level coincide.

Theorem 3. *Let $\circ = (\circ_1, \circ_2, \ldots, \circ_b)$ be a parallel alternative operation on V such that for each $1 \le j \le b$ $\circ_j$ is an alternative operation on V_j. Let us assume that every $\circ_j$ is such that $\dim(W_{\circ_j}) = s-2$ and it is defined by a vector $\boldsymbol{b} \in (\mathbb{F}_2)^{s-2}$. Let $\lambda \in (\mathbb{F}_2)^{n\times n}$. Then, $\lambda \in H_\circ$ if and only if it can be represented in the block form*

$$\lambda = \left(\begin{array}{cc|c|cc} A_{11} & B_{11} & & A_{1b} & B_{1b} \\ C_{11} & D_{11} & \cdots & C_{1b} & D_{1b} \\ \hline \multicolumn{2}{c|}{\vdots} & \ddots & \multicolumn{2}{c}{\vdots} \\ \hline A_{b1} & B_{b1} & & A_{bb} & B_{bb} \\ C_{b1} & D_{b1} & \cdots & C_{bb} & D_{bb} \end{array}\right),$$

where

1. *$A_{ij} \in (\mathbb{F}_2)^{2\times 2}$ such that for each row and each column of blocks there exists one and only one non-zero A_{ij}; moreover, all the non-zero A_{ij} are invertible;*
2. *$B_{ij} \in (\mathbb{F}_2)^{2\times(s-2)}$;*
3. *$C_{ij} = \mathbb{0}_{(s-2)\times 2}$;*
4. *$D_{ij} \in (\mathbb{F}_2)^{(s-2)\times(s-2)}$ such that if A_{ij} is zero, then $\boldsymbol{b}D_{ij} = \boldsymbol{0}$, and if A_{ij} is invertible, then $\boldsymbol{b}D_{ij} = \boldsymbol{b}$. Moreover, the matrix D defined by*

$$D := \begin{pmatrix} D_{11} & \cdots & D_{1b} \\ \vdots & \ddots & \vdots \\ D_{b1} & \cdots & D_{bb} \end{pmatrix}$$

is invertible.

Proof. The proof involves standard linear algebra techniques, but its extensive and laborious nature necessitates omission due to page limitations.

3 Differential properties of optimal s-boxes

In this section we delve into the examination of the differential properties exhibited by all possible 4-bit permutations, with respect to all possible alternative operations defined as in Sec. 1.2. In particular, we set $s = 4$ and therefore consider $B = \mathbb{F}_2^4$. We begin by acknowledging that, despite the compact size of the space, the count of alternative operations on B is considerable:

Proposition 1 ([8]). *There exist 105 different elementary abelian regular subgroups groups $T_\circ$ in $AGL(\mathbb{F}_2^4, +)$. Furthermore, each of them satisfies $T_+ < AGL(\mathbb{F}_2^4, \circ)$ and $\dim W_\circ = s - 2 = 2$.*

We recall that given a permutation $f \in \mathrm{Sym}(B)$ we can define

$$\delta_f(a, b) = \#\{x \in B \mid xf + (x + a)f = b\}.$$

The *differential uniformity* of f is defined as $\delta_f := \max_{a \neq 0} \delta_f(a, b)$ and it represent the primary metric to consider when assessing the resistance of an s-box to differential cryptanalysis [13].

Several cryptographic properties, including differential uniformity, are preserved under affine equivalence for vectorial Boolean functions. Two functions, denoted as f and g, are considered *affine equivalent* if there exist two affine permutations, α and β, in $\mathrm{AGL}(V, +)$ such that $g = \beta f \alpha$.

Leander and Poschmann [12] provided a comprehensive classification (up to affine equivalence) of permutations over $B = \mathbb{F}_2^4$. They identified 16 classes with *optimal* cryptographic properties. All 16 classes exhibit a classical differential uniformity equal to 4, which represents the best possible value for s-boxes in $\mathrm{Sym}(B)$. The representatives of the 16 classes are listed in Table 1, where each vector is interpreted as a binary number, most significant bit first.

3.1 Dealing with affine equivalence

Our goal is to analyse the differential uniformity of each optimal s-box class, with respect to every alternative operation $\circ$ on B. The definitions given above can be generalised in the obvious way setting $\delta_f^\circ(a, b) = \#\{x \in B \mid xf \circ (x \circ a)f = b\}$ and calling $\circ$-*differential uniformity* of f the value $\delta_f^\circ := \max_{a \neq 0} \delta_f^\circ(a, b)$.

It is noteworthy that, unlike in the case of classic differential uniformity, the value of δ_f° is not invariant under affine equivalence. However, verifying the $\circ$-differential uniformity of $g_2 G_i g_1$ for any optimal class and every pair $g_1, g_2 \in \mathrm{AGL}(V, +)$ would be impractical. Therefore, a reduction in the number of permutations to be checked is necessary, and for this purpose, we make the following observations. First, similar to the classical case, the $\circ$-differential uniformity is preserved under affine transformations w.r.t. $\circ$.

Table 1. Optimal 4-bit permutations according to Leander and Poschmann

	0_x 1_x 2_x 3_x 4_x 5_x 6_x 7_x 8_x 9_x A_x B_x C_x D_x E_x F_x
G_0	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x B_x C_x 9_x 3_x E_x A_x 5_x
G_1	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x B_x E_x 3_x 5_x 9_x A_x 12_x
G_2	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x B_x E_x 3_x A_x C_x 5_x 9_x
G_3	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x C_x 5_x 3_x A_x E_x B_x 9_x
G_4	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x C_x 9_x B_x A_x E_x 5_x 3_x
G_5	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x C_x B_x 9_x A_x E_x 3_x 5_x
G_6	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x C_x B_x 9_x A_x E_x 5_x 3_x
G_7	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x C_x E_x B_x A_x 9_x 3_x 5_x
G_8	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x E_x 9_x 5_x A_x B_x 3_x 12_x
G_9	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x E_x B_x 3_x 5_x 9_x A_x 12_x
G_{10}	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x E_x B_x 5_x A_x 9_x 3_x 12_x
G_{11}	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x E_x B_x A_x 5_x 9_x C_x 3_x
G_{12}	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x E_x B_x A_x 9_x 3_x C_x 5_x
G_{13}	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x E_x C_x 9_x 5_x B_x A_x 3_x
G_{14}	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x E_x C_x B_x 3_x 9_x 5_x 10_x
G_{15}	0_x 1_x 2_x D_x 4_x 7_x F_x 6_x 8_x E_x C_x B_x 9_x 3_x A_x 5_x

Proposition 2. *Given $f \in \mathrm{Sym}(B)$ and $g_1, g_2 \in \mathrm{AGL}(B, \circ)$ we have*

$$\delta^{\circ}_{g_1 f g_2}(a, b) = \delta^{\circ}_{f}(g_2(a), g_1^{-1}(b)).$$

Moreover, Proposition 1 establishes that for any $\circ$ derived from a translation group in $\mathrm{AGL}(B, +)$, the $+$-translations are affine with respect to $\circ$. This initial observation allows us to narrow down the analysis to $g_2 G_i g_1$ with $g_1, g_2 \in \mathrm{GL}(B, +)$, which still remains impractical. Furthermore, considering that $H_\circ = \mathrm{GL}(B, +) \cap \mathrm{GL}(B, \circ)$, Proposition 2 establishes that left and right multiplication by elements in $H_\circ$ preserves both $\circ$ and $+$-differential uniformity. It is noteworthy that during this process, the rows of the matrix containing all the $\delta^{\circ}_{f}(a, b)$ (DDT°) are merely shuffled, thereby preserving the highest element of each row. Therefore, the following conclusion can be easily obtained.

Proposition 3. *Let $g_1, g_2 \in \mathrm{GL}(B, +)$ and $f \in \mathrm{Sym}(B)$. For any $g_1' \in g_1 H_\circ$ and $g_2' \in H_\circ g_2$ we have*

$$\delta^{\circ}_{g_2 f g_1} = \delta^{\circ}_{g_2' f g_1'}.$$

Proof. Take $h_1, h_2 \in H_\circ$ such that $g_1' = g_1 h_1$ and $g_2' = h_2 g_2$. Then,

$$x g_2' f g_1' \circ (x \circ a) g_2' f g_1' = x h_2 g_2 f g_1 \circ (x h_2 \circ a h_2 g_2 f g_1) h_1,$$

implying that $\delta^{\circ}_{g_2' f g_1'}(a, b) = \delta^{\circ}_{g_2 f g_1}(a h_2, b h_1^{-1})$. So, $\delta^{\circ}_{g_2' f g_1'} = \delta^{\circ}_{g_2 f g_1}$. □

The final proposition allows us to focus solely on g_1 and g_2 within the left and right cosets of $H_\circ$. These reductions facilitate the analysis of the potential $\circ$-differential uniformities attainable across all classes of optimal permutations for the 105 conceivable alternative sums defined over B. For each of the 105

alternative operations, we systematically explored each of the 16 classes, following the described procedure, and we recorded the $\circ$-differential uniformity for every candidate. To streamline the presentation, we calculated the average across the 105 operations and presented the consolidated results in Tab. 2.

Table 2. Avg. number of functions with given $\circ$-differential uniformity

Class \ δ°	2	4	6	8	10	12	14	16
G_0	0	914	7842	3463	420	19	0	14
G_1	0	1019	10352	4226	560	0	0	18
G_2	0	1003	8604	3805	462	21	0	16
G_3	0	16733	117740	27639	1779	0	0	0
G_4	0	1101	9295	2715	179	0	0	0
G_5	0	2479	24135	5402	639	0	0	0
G_6	0	1632	10842	3071	218	0	0	0
G_7	0	1257	10679	2994	186	28	0	0
G_8	0	1691	12821	6113	583	93	0	24
G_9	0	1228	7734	2693	154	39	0	0
G_{10}	0	1228	8063	2763	166	41	0	0
G_{11}	0	1637	9940	2941	214	0	0	0
G_{12}	0	2541	16832	5308	352	0	0	0
G_{13}	0	1124	9520	2416	217	15	0	0
G_{14}	0	1207	7641	2584	160	51	0	0
G_{15}	0	1227	7776	2630	163	52	0	0

In our examination, we observe that if, for a given operation $\circ$, certain elements within an affine equivalence class yield a $\circ$-differential uniformity δ, then this value δ is achieved by some element in the entire class for all alternative operations. Our analysis reveals that certain optimal functions may exhibit the highest differential uniformity (16) for alternative operations, specifically the classes G_0 (containing, e.g., the s-box S1 of Serpent [2]), G_1 (containing, e.g., the s-box of Present [5]), G_2, and G_8. Conversely, the classes G_3, G_4, G_5, G_6, G_{11}, and G_{12} demonstrate more favorable behavior concerning alternative operations.

4 Experiments on a 16-bit block cipher with 4-bit s-boxes

In this concluding section, we aim to apply the results obtained above to a family of (toy) ciphers. These ciphers may exhibit security under classical differential cryptanalysis but reveal vulnerabilities to the alternative differential approach.

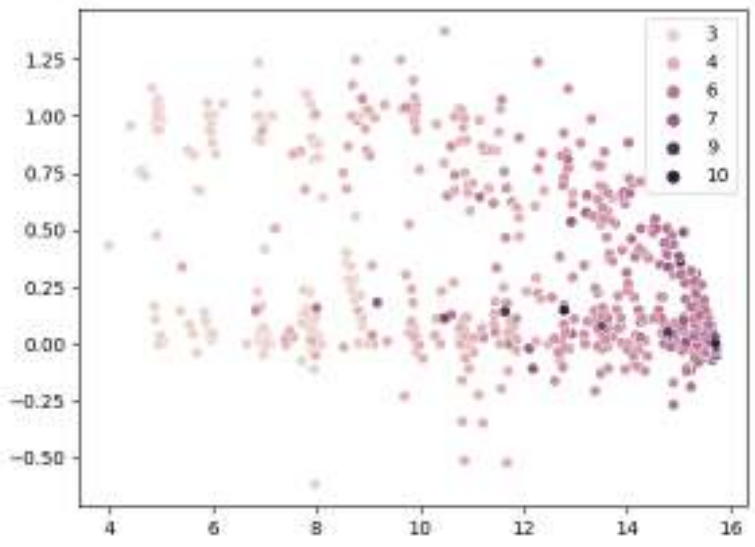
In our experiments, we set $V = \mathbb{F}_2^{16}$, $n = 4$, and $s = 4$, defining $\circ$ as the parallel sum by applying the alternative operation defined by the vector $\mathbf{b} = (0, 1)$ to each 4-bit block. Moreover, all our ciphers will feature the 4-bit permutation $\gamma : \mathbb{F}_2^4 \to \mathbb{F}_2^4$ defined by the sequence (0_x, E_x, B_x, 1_x, 7_x, C_x, 9_x, 6_x, D_x, 3_x, 4_x, F_x, 2_x, 8_x, A_x, 5_x) as its s-box. Precisely, four copies of γ will act on the 16-bit block. Notice that the s-box $\gamma \in$ belongs to G_0 and has $\delta_\gamma = 4$ and $\delta_\gamma^\circ = 16$.

In all the experiments described below, we consider the SPN whose i-th round is obtained by the composition of the parallel application of the s-box γ on every 4-bit block, a 'diffusion layer' λ sampled random from $H_\circ$, and the xor with the i-th random round key. We study the difference propagation in the cipher in a long-key scenario, i.e., the key-schedule selects a random long key $k \in \mathbb{F}_2^{16r}$ where r is the number of rounds. To avoid potential bias from a specific key choice, we conduct our experiments by averaging over 2^{15} random long-key generations. This approach gives us a reliable estimate of the expected differential probability for the best differentials in this cipher.

In 150 distinct executions, spanning a range of rounds from 3 to 10, we calculated the discrepancy between the most effective $\circ$-trail and $+$-trail. To manage computational resources, our focus was narrowed down to input differences with a Hamming weight of 1.

The results are depicted in Fig. 1, where each dot represents an individual simulation. The x axis corresponds to the negative logarithm of the probability of the best $\circ$ differential, while the y axis represents the difference between that value and the negative logarithm of the probability of the best $+$ differential. Darker dots indicate a higher number of rounds, as explained in the legend. Notably, about half of the dots lie above zero, suggesting that the best $\circ$ differential consistently outperforms the best $+$ differential until they become indistinguishable. Interestingly, this convergence often occurs when the $\circ$ probability is already very close to 16, providing potential candidates for our distinguisher attack.

Fig. 1. Comparison of $+$ and $\circ$ trails for random mixing layers



References

1. Banik, S., Bogdanov, A., Isobe, T., Shibutani, K., Hiwatari, H., Akishita, T., Regazzoni, F.: Midori: A block cipher for low energy. In: Advances in Cryptology–ASIACRYPT 2015: 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29–December 3, 2015, Proceedings, Part II 21, pp. 411–436, Springer (2015)
2. Biham, E., Anderson, R., Knudsen, L.: Serpent: A new block cipher proposal. In: International workshop on fast software encryption, pp. 222–238, Springer (1998)
3. Biham, E., Biryukov, A., Shamir, A.: Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials. Journal of Cryptology **18**, 291–311 (2005)
4. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. Journal of CRYPTOLOGY **4**, 3–72 (1991)
5. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J., Seurin, Y., Vikkelsoe, C.: Present: An ultra-lightweight block cipher. In: Cryptographic Hardware and Embedded Systems-CHES 2007: 9th International Workshop, Vienna, Austria, September 10-13, 2007. Proceedings 9, pp. 450–466, Springer (2007)
6. Borisov, N., Chew, M., Johnson, R., Wagner, D.: Multiplicative differentials. In: Fast Software Encryption: 9th International Workshop, FSE 2002 Leuven, Belgium, February 4–6, 2002 Revised Papers 9, pp. 17–33, Springer (2002)
7. Calderini, M., Civino, R., Invernizzi, R.: Differential experiments using parallel alternative operations. Journal of Mathematical Cryptology **18**(1), 20230030 (2024)
8. Calderini, M., Civino, R., Sala, M.: On properties of translation groups in the affine general linear group with applications to cryptography. Journal of Algebra **569**, 658–680 (2021)
9. Caranti, A., Dalla Volta, F., Sala, M.: Abelian regular subgroups of the affine group and radical rings. Publ. Math. Debrecen **69**(3), 297–308 (2006)
10. Civino, R., Blondeau, C., Sala, M.: Differential attacks: using alternative operations. Designs, Codes and Cryptography **87**, 225–247 (2019)
11. Knudsen, L.R.: Truncated and higher order differentials. In: Fast Software Encryption: Second International Workshop Leuven, Belgium, December 14–16, 1994 Proceedings 2, pp. 196–211, Springer (1995)
12. Leander, G., Poschmann, A.: On the classification of 4 bit s-boxes. In: Arithmetic of Finite Fields: First International Workshop, WAIFI 2007, Madrid, Spain, June 21-22, 2007. Proceedings 1, pp. 159–176, Springer (2007)
13. Nyberg, K.: Differentially uniform mappings for cryptography. In: Workshop on the Theory and Application of of Cryptographic Techniques, pp. 55–64, Springer (1993)
14. Shibutani, K., Isobe, T., Hiwatari, H., Mitsuda, A., Akishita, T., Shirai, T.: Piccolo: an ultra-lightweight blockcipher. In: Cryptographic Hardware and

Embedded Systems–CHES 2011: 13th International Workshop, Nara, Japan, September 28–October 1, 2011. Proceedings 13, pp. 342–357, Springer (2011)
15. Teşeleanu, G.: The security of quasigroups based substitution permutation networks. In: International Conference on Information Technology and Communications Security, pp. 306–319, Springer (2022)
16. Wagner, D.: The boomerang attack. In: International Workshop on Fast Software Encryption, pp. 156–170, Springer (1999)
17. Zajac, P., Jókay, M.: Cryptographic properties of small bijective s-boxes with respect to modular addition. Cryptography and Communications **12**, 947–963 (2020)

On the Properties of the Ortho-Derivatives of Quadratic Functions

Alain Couvreur[1,2], Anne Canteaut[1], Léo Perrin[1]

[1] Inria, France
[2] LIX, CNRS UMR 7161, École Polytechnique,
Institut Polytechnique de Paris, France
{anne.canteaut, alain.couvreur, leo.perrin}@inria.fr

Abstract. Quadratic APN vectorial functions are under intense scrutiny due to their role e.g. in the big APN problem. Recently, a new tool has emerged to investigate their differential properties: the ortho-derivative. We present new results about this object. We first generalize it as a family of functions that can be defined for any quadratic function, even if not APN. We highlight a relation between the preimages sets of the ortho-derivative and the set of bent components, and between the ortho-derivative and some EA-invariants recently introduced by Kaleyski. We also show it is possible to reconstruct a quadratic function given its ortho-derivative.
In the APN case, we prove that its algebraic degree is always at most equal to $n-2$ using a previously unknown relation between the ortho-derivatives and cofactor matrices.

Keywords: Boolean Functions · Quadratic · APN · Ortho-derivative

1 Introduction

Let $\mathbb{F}_2 = \{0,1\}$ be the field with two elements and $n > 0$ be an integer. We use $x \cdot y = \sum_{i=0}^{n-1} x_i y_i$ to denote the scalar product of two elements of $\mathbb{F}_2^n$. The functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ are *Boolean functions*, and those mapping $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ are *vectorial Boolean functions*. In this article, we only consider the case where $m = n$. We let $\mathcal{F}_n$ denote the set of all functions from $\mathbb{F}_2^n$ to itself. Each of the coordinates of a vectorial Boolean function has a unique representation as a polynomial of n variables in $\mathbb{F}_2$ called its *algebraic normal form*. The degree of this representation is the *algebraic degree* of the Boolean function, and the algebraic degree of a function of $\mathcal{F}_n$ is the maximum algebraic degree of its coordinates.

A linear combination of some coordinates is a *component*, and the distance between a component $x \mapsto b \cdot F(x)$ and a linear function $x \mapsto a \cdot x$ is given by the *Walsh coefficient* $\mathcal{W}_F(a,b) = \sum_{x \in \mathbb{F}_2^n} (-1)^{a \cdot x + b \cdot F(x)}$. The maximum of $|\mathcal{W}_F(a,b)|$ taken over all $a \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2^n \backslash \{0\}$ is the *linearity* of F, denoted $\mathcal{L}(F)$.

The *derivative* of a vectorial Boolean function F is defined for any $a \in \mathbb{F}_2^n$ and is the function $\Delta_a F$ mapping x to $F(x+a)+F(x)$. The number of solutions x of

the equation $\Delta_a F(x) = b$ is denoted by $\delta_F(a, b)$ and its maximum, taken over all $a \neq 0$ and all b in $\mathbb{F}_2^n$, is called the differential uniformity [19] of the function F and is denoted by u_F. When $u_F = 2$, we say that F is *Almost Perfect Nonlinear (APN)*. The existence of APN permutations when n is even is an open problem (known as *the* big APN problem), except when $n = 6$ where a sporadic solution was found by Dillon *et al.* [6].

A function $F \in \mathcal{F}_n$ is a collection of n *coordinates*, each being a Boolean function mapping $\mathbb{F}_2^n$ to $\mathbb{F}_2$. Each of these coordinates has a unique representation as a polynomial of n variables in $\mathbb{F}_2$ called its *algebraic normal form*. Its degree is the *algebraic degree* of the Boolean function, and the algebraic degree of a function of $\mathcal{F}_n$ is the maximum algebraic degree of its coordinates.

Let $F \in \mathcal{F}_n$ be a quadratic APN function. Then there exists a unique function $\pi_F \in \mathcal{F}_n$ such that $\pi_F(0) = 0$, $\pi_F(a) \neq 0$ for $a \neq 0$, and

$$\text{for any } \ (a, x) \in (\mathbb{F}_2^n)^2, \ \pi_F(a) \cdot \big(F(x) + F(x + a) + F(0) + F(a)\big) = 0. \qquad (1)$$

Functions corresponding to such π_F have been studied before [3,20,17,13,14,11], and π_F was called the *ortho-derivative of* F in [8]. In that paper, the authors showed that the ortho-derivative was a powerful tool to investigate the CCZ- and EA-equivalence of quadratic APN functions, as shown by its later use in [2,21].

In this article, we present some new results on the ortho-derivative. First, we generalize it to any quadratic function, and in particular to non-APN ones (Section 2). The non-trivial ortho-derivatives of a quadratic function F then form a family of functions that reduces to a single function if and only if F is APN. We then discuss the algebraic degree of ortho-derivatives, and in particular prove that this degree is at most equal to $n - 2$ for APN functions (Section 3).

We then shift our focus to more practical aspects. First, we prove that some EA- and CCZ-invariants introduced in [7,16] can be derived from the Hamming weight of the ortho-derivative in the case of quadratic APN functions. Then, we show that it is possible to define and to implement an operation that is the inverse of ortho-derivation, namely the *ortho-integration* (Section 4).

2 The Ortho-Derivatives of any Quadratic Function

We first recall the definition of the ortho-derivative, generalize it to any quadratic function, and describe some basic properties (Section 2.1). Then, we investigate the values an ortho-derivative can take, and highlight some simple relations between those and the Walsh spectrum of the function in Section 2.2.

2.1 Definition and Basic Properties

Definition 1. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a quadratic function. We say that $\pi : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is an* ortho-derivative for F *if, for any x and a in $\mathbb{F}_2^n$,*

$$\pi(a) \cdot \big(F(x) + F(x + a) + F(0) + F(a)\big) = 0 \, .$$

The set of all ortho-derivatives for a given F is denoted by $\Pi(F)$.

It is worth noting that $\Pi(F)$ always contains several ortho-derivatives since $\pi(0)$ can take any value and, for any $a \neq 0$, 0 is a valid value for $\pi(a)$. Therefore, we say in the following that an ortho-derivative $\pi \in \Pi_F$ is *non-trivial* if $\pi(0) = 0$ and $\pi(a) \neq 0$ for all nonzero a.

Among all quadratic functions, APN functions are characterized by the fact that they have a single non-trivial ortho-derivative. More generally, the number of ortho-derivatives depends on the differential spectrum of the quadratic function. In order to establish this, we need the following properties (the proof is omitted due to the page-count limitation).

Proposition 1. *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a quadratic function. For any nonzero $a \in \mathbb{F}_2^n$, we define*

$$P_F(a) = \{x : \exists \pi \in \Pi(F) \text{ with } \pi(a) = x\}.$$

Then, for any $a \in \mathbb{F}_2^n$, $P_F(a)$ is a linear space and $\#P_F(a) = \max_{b \in \mathbb{F}_2^n} \delta_F(a, b)$. Moreover,

$$\#\{(a, b) : \pi(a) = b \text{ for some } \pi \in \Pi(F)\} = 2^{-n} \sum_{a,b \in \mathbb{F}_2^n} \delta_F(a, b)^2.$$

We can then easily establish a link between the APN property and the number of ortho-derivatives.

Corollary 1. *Let $F \in \mathcal{F}_n$ be a quadratic function. Then, the following conditions are equivalent:*

(i) *F is APN.*
(ii) *There exists a unique nontrivial $\pi \in \Pi(F)$.*
(iii) *$\#\{(a, b) : \pi(a) = b \text{ for some } \pi \in \Pi(F)\} = 2^n + 2(2^n - 1)$.*

If we do not assume that F is quadratic, a similar characterization holds for generalized crooked functions (see Definition 4.2 in [9]).

2.2 On the values an ortho-derivative can take

In the following, we study the set

$$\mathcal{S} = \{(a, b) : \pi(a) = b \text{ for some } \pi \in \Pi(F)\} .$$

This set can be partitioned into linear spaces $P_F(a)$ whose dimensions are determined by the differential spectrum of F. The existence of vector space partitions with a given *type*, i.e., such that the dimensions of the involved vector spaces are given, is a well-known problem studied by many authors, see [15] for a survey.

We can also partition the same set according to the values of b, i.e. into sets

$$T_F(b) = \{a : \exists \pi \in \Pi(F) \text{ with } \pi(a) = b\}.$$

This provides another partition of $\mathcal{S}$ into vector spaces, as already established by Gorodilova in [14].

Proposition 2. *Let $F \in \mathcal{F}_n$ be a quadratic function. For any $b \in \mathbb{F}_2^n$, we define*

$$T_F(b) = \{a : \exists \pi \in \Pi(F) \text{ with } \pi(a) = b\} .$$

Then, $T_F(b) = \mathsf{LS}(F_b)$ where $\mathsf{LS}(F_b)$ denotes the set of all linear structures of the component $x \mapsto b \cdot F(x)$. It follows that, for any b, $T_F(b)$ is a linear subspace of $\mathbb{F}_2^n$ whose dimension has the same parity as n.

We deduce that the partition of $\mathcal{S}$ into $T_F(b), b \in \mathbb{F}_2^n$ is derived from the Walsh spectrum of F, while the partition into $P_F(a), a \in \mathbb{F}_2^n$ corresponds to its differential spectrum.

Proposition 3. *Let $F \in \mathcal{F}_n$ be a quadratic function. Then, for any of its nontrivial component F_b, we have $\mathcal{L}(F_b) = 2^{\frac{n+\dim T_F(b)}{2}}$.*

We then easily recover the characterization of APN functions from their Walsh spectrum (see e.g [4, Corollary 1]):

$$\sum_{\mu \neq 0} \sum_{\lambda \in \mathbb{F}_2^n} \mathcal{W}_F^4(\lambda, \mu) = (2^n - 1)2^{3n+1} .$$

Indeed, Proposition 3 combined with Corollary 1 leads to the following equivalent formulation.

Corollary 2. *Let $F : \mathcal{F}_n$ be a quadratic function and, for any $d, 0 \leqslant d \leqslant n$,*

$$\begin{aligned} B_d &= \#\{b \neq 0 : \mathcal{L}^2(F_b) = 2^{n+d}\} \\ &= \#\{b \neq 0 : b \text{ has } 2^d \text{ preimages by some } \pi \in \Pi(F)\} . \end{aligned}$$

Then, $\sum_{d=1}^n B_d(2^d - 1) \geqslant 2^n - 1$, with equality if and only if F is APN.

Proof. Recall that $\pi(0)$ can take any value, and that $\pi(a) = 0$ is a valid value for any a. Let $\widetilde{B_d} = \#\left\{b : b \text{ has } 2^d \text{ preimages by some } \pi \in \Pi(F)\right\}$, and consider the set $\mathcal{S} = \{(a, b) : \pi(a) = b \text{ for some } \pi \in \Pi(F)\}$. Then, the size of $\mathcal{S}$ is given by $\sum_b T_F(b) = \sum_{d=0}^n \widetilde{B_d} 2^d$. Using that $\mathcal{L}(F_0) = 2^n$, we get $\sum_{d=0}^n \widetilde{B_d} 2^d = \sum_{d=0}^n B_d 2^d + 2^n$. Moreover,

$$\sum_{d=0}^n B_d 2^d = \sum_{d=0}^n B_d(2^d - 1) + \sum_{d=0}^n B_d = \sum_{d=1}^n B_d(2^d - 1) + (2^n - 1) .$$

It follows that $\#\mathcal{S} = \sum_{d=1}^n B_d(2^d - 1) + 2^{n+1} - 1$. From Corollary 1, we know that $\#\mathcal{S} \geqslant 2^n + 2(2^n - 1)$ with equality if and only if F is APN. Equivalently, $\sum_{d=1}^n B_d(2^d - 1) \geqslant 2^n - 1$ with equality if and only if F is APN. □

2.3 Relations between the properties of a function and those of its ortho-derivative

n odd. When n is odd, all $T_F(b)$ have an odd dimension, implying that they can never have dimension zero. Then, if F is APN, we have that $B_1 = 2^n - 1$, or equivalently that the only nontrivial π is a permutation. It follows that F is almost bent, as proved in [10].

n even. When n is even, the values taken by $(B_0, \ldots, B_n)$ are not unique. Some conditions on B_0 can be deduced from Corollary 2.

Corollary 3. *Let $n \geqslant 4$ be an even integer and $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a quadratic APN function. Let B_0 denote the number of bent components of F. Then $B_0 \geqslant 2 \times \frac{(2^n-1)}{3}$, with equality if and only if $B_2 = \frac{(2^n-1)}{3}$ and all other B_d vanish.*
Moreover, $B_0 \equiv 2 \bmod 4$, and $B_n = 0$.

Proof. The first statement corresponds to Corollary 3 in [4]. It is a straightforward consequence of $\sum_{d=1}^{n} B_d(2^d - 1) = 2^n - 1$, where $B_d = 0$ for all odd d. Therefore

$$2^n - 1 \geqslant 3\left(\sum_{d=2}^{n} B_d\right) = 3(2^n - 1 - B_0),$$

leading to $B_0 \geqslant 2 \times (2^n - 1)/3$ with equality if and only if $B_d = 0$ for all $d > 2$.

Also, from

$$\sum_{d=2}^{n} B_d(2^d - 1) \;=\; \sum_{d=2}^{n} B_d 2^d - \sum_{d=2}^{n} B_d = 2^n - 1,$$

we deduce that

$$4\sum_{d=2}^{n} B_d 2^{d-2} - (2^n - 1 - B_0) \;=\; 2^n - 1$$

which implies that $B_0 \equiv 2 \bmod 4$. We can also use the same characterization to prove that a quadratic APN function cannot have linearity 2^n. Indeed, if $B_n > 0$, then the only possibility is $B_n = 1$ and $B_0 = 2^n - 2$. Let us assume w.l.o.g that the component of F with linearity 2^n corresponds to the last coordinate. Then, the function F' from $\mathbb{F}_2^n$ into $\mathbb{F}_2^{n-1}$ derived from F by removing the last coordinate is bent, i.e. all its components are bent. However, bent functions only exist if the number of outputs is at most half of the number of inputs, i.e. $n - 1 \leqslant \frac{n}{2}$ [18]. This cannot occur if $n \geqslant 4$. □

The previous corollary shows that, for any quadratic APN function F, $\mathcal{L}(F) \leqslant 2^{n-1}$. It is worth noticing that such quadratic APN functions have been exhibited for $n = 6, 8$. All classes of quadratic APN functions of 6 variables have the lowest possible value of B_0, i.e. $B_0 = 42$, except one which satisfies

$$B_0 = 46,\ B_2 = 16 \text{ and } B_4 = 1\ .$$

For $n = 8$ variables, the quadratic APN functions exhibited in [22,2,1] have six different spectra corresponding to the following $(B_d)_{d \leqslant n}$:

$B_0 = 170,\ B_2 = 85,$
$B_0 = 174,\ B_2 = 80,\ B_4 = 1,$
$B_0 = 178,\ B_2 = 75,\ B_4 = 2,$
$B_0 = 182,\ B_2 = 70,\ B_4 = 3,$
$B_0 = 186,\ B_2 = 65,\ B_4 = 4,$
$B_0 = 190, B_2 = 64, B_4 = 0, B_6 = 1.$

Thanks to Corollary 2, we can deduce some properties of the differential spectrum of the ortho-derivative π_F from the Walsh spectrum of a quadratic APN function F.

Corollary 4. *Let $F \in \mathcal{F}_n$ be a quadratic APN function, π_F be its nontrivial ortho-derivative, and B_d be the number of components of F with squared linearity 2^{n+d}. Then, for all d, we have that at least $B_d(2^d-1)$ entries in the DDT of π_F are greater than or equal to 2^d-2. In particular,*

$$B_d > 0 \implies u_{\pi_F} \geqslant 2^d - 2 \ .$$

For instance, it can be checked that the nontrivial ortho-derivative of the previously mentioned 8-bit APN function with $B_6 = 1$ has differential uniformity 62. The converse statement of this corollary is false: there are 8-bit APN functions with an ortho-derivative with a differential uniformity of 30 but for which $B_d = 0$ for all $d \geqslant 3$, whereas a true converse statement would have implied that $B_4 > 0$.

3 On the degree of the ortho-derivative(s)

In [14], Gorodilova studied the ortho-derivatives of quadratic APN functions and identified several of their properties. In particular, she proved that the algebraic degree of a nontrivial ortho-derivative is either n or at most $n-2$, and as an immediate corollary, that the nontrivial ortho-derivative is of degree at most $n-2$ when n is odd (because then the ortho-derivative is a permutation).

In the case where n is even, she could only conjecture that all nontrivial components of the ortho-derivative have algebraic degree exactly equal to $n-2$ (Conjecture 2 of [14]). In this section, we show that a nontrivial ortho-derivative of an APN function is always of degree at most $n-2$.

Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a quadratic APN function. We define J as the function mapping an element of $\mathbb{F}_2^n$ to the binary $n \times n$ matrix such that, for all $x \in \mathbb{F}_2^n$, $J_{i,j}(x) = \Delta_{e_j} F_i(x) + \Delta_{e_j} F_i(0)$. If F is quadratic, then J is the linear part of its Jacobian matrix. From [8, Prop. 14], we know that, for every x and a,

$$J(x) \times a = J(a) \times x = \Delta_a F(x) + \Delta_a F(0) \ . \tag{2}$$

As established in [8], F is APN if and only if $\mathrm{Rank}(J(a)) = n-1$ for all $a \neq 0$. It is a direct consequence of the fact that a function is APN if and only if the image of $\Delta_a F$ is of dimension $n-1$. We deduce that both the left and the right kernels of $J(a)$ contain a single non-trivial element. First, we remark that Equation (2) immediately implies

$$J(a) \times a \ = \ \Delta_a F(a) + \Delta_a F(0) = F(a+a) + F(a) + F(a+0) + F(0) = 0 \ ,$$

meaning that the right kernel of $J(a)$ is $\{0, a\}$. On the other hand, for any $c \in \mathbb{F}_2^n$, we have that $c^T J(a) x$ is the scalar product of c with an element in the image of $\Delta_a F$. Let π be the nontrivial ortho-derivative of F. By definition, we then have

$\pi(a)^T J(a)x = 0,\ \forall x \in \mathbb{F}_2^n$. We deduce that $\pi(a)^T J(a) = 0$, meaning that the left kernel of $J(a)$ consists of $\{0, \pi(a)^T\}$.

For any $n \times n$ binary matrix M, we denote $\mathsf{Cof}(M)$ its cofactors matrix:

$$\mathsf{Cof}(M) = \begin{bmatrix} \det(C_{0,0}) & \cdots & \det(C_{0,n-1}) \\ \vdots & & \vdots \\ \det(C_{n-1,0}) & \cdots & \det(C_{n-1,n-1}) \end{bmatrix} ,$$

where $\det(C_{i,j})$ is the (i,j) minor of M, *i.e.* the determinant of the submatrix obtained by removing Row i and Column j from M. Furthermore, it is well-known that

$$\mathsf{Cof}(M)^T M \ = \ M\mathsf{Cof}(M)^T \ = \ \mathrm{Id} \times \det(M) \ .$$

Let us apply this equality to $J(a)$. As it is of rank $n-1 < n$, we have that

$$\mathsf{Cof}(J(a))^T J(a) = J(a)\mathsf{Cof}(J(a))^T = 0 \ . \tag{3}$$

Lemma 1. *The cofactors matrix of $J(a)$ can be written*

$$\mathsf{Cof}\left(J(a)\right) \ = \ \begin{bmatrix} \pi_0(a)a_0 & \cdots & \pi_0(a)a_{n-1} \\ \vdots & & \vdots \\ \pi_{n-1}(a)a_0 & \cdots & \pi_{n-1}(a)a_{n-1} \end{bmatrix} \ = \ \begin{bmatrix} \pi_0(a) \\ \vdots \\ \pi_{n-1}(a) \end{bmatrix} [a_0, \ldots, a_{n-1}] \ .$$

Proof. As we have established, the right kernel of $J(a)$ contains only 0 and a. This implies that the row space of $\mathsf{Cof}(J(a))$ is contained in $\{0, a\}$. On the other hand, since the left kernel of $J(a)$ is $\{0, \pi(a)\}$, we have that the column space of $\mathsf{Cof}(J(a))$ must be contained in that space.

Since $J(a)$ has rank $n-1$, it has at least one nonzero minor. Hence, its cofactor matrix is nonzero. Thus, the only possibility is $\mathsf{Cof}(J(a)) = \pi(a) \cdot a^T$. □

A direct corollary of this lemma is that each minor of $J(a)$ is equal to $\det(C_{i,j}) = \pi_i(a)a_j$.

Since F is quadratic, the entries in $J(a)$ are linear functions in a. Next, from Leibniz' formula on an $(n-1) \times (n-1)$ binary matrix M:

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_{n-1}} \prod_{i=0}^{n-2} m_{i,\sigma(i)} \ , \tag{4}$$

we deduce that the degree of each entry in $\mathsf{Cof}(J(a))$ is a Boolean function of degree at most $n-1$ in a as it is a sum of products of $n-1$ linear functions.

We deduce that $\deg(\pi_i(a)a_j)$ is at most $n-1$, for all j. Indeed, suppose that there exists a term of degree $n-1$ in the ANF of π_i for some i, and that it corresponds to $\prod_{k \neq j} a_k$. Then the entry at position (i,j) in the cofactors matrix $\mathsf{Cof}(J(a))$ would be a function of degree n, which is impossible. The next theorem follows.

Theorem 1. *If π is the nontrivial ortho-derivative of a quadratic APN function of $\mathbb{F}_2^n$, then* $\deg(\pi) \leqslant n-2$.

Corollary 5. *Let $F \in \mathcal{F}_n$ be a quadratic function. Then it is not APN if and only if at least one of its nontrivial ortho-derivatives is of algebraic degree n.*

4 On Ortho-Integration

We call *ortho-integration* the process that is the inverse of ortho-derivation.

Definition 2 (Ortho-Integral). *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a quadratic function, and let $\Pi(F)$ be the set of its ortho-derivatives. For all $\pi \in \Pi(F)$, we say that F is an* ortho-integral *of π.*

An ortho-integral is not unique: for any affine function A of $\mathbb{F}_2^n$, F and $F+A$ have the same ortho-derivatives. This does not necessarily imply that their DDTs are identical, though it is worth noting that all known examples of APN functions with the same DDT differ from an affine function [5,13]. In general, it is unclear whether there exists other "collisions", i.e. distinct functions with identical ortho-derivatives.

Recall that, for any π in $\Pi(F)$ and any $x \in \mathbb{F}_2^n$, it holds that $\pi(a)^T J(a) x = 0$. Since this equation is linear in x for all a, it is sufficient to consider elements x in a basis of $\mathbb{F}_2^n$. Then, recovering $J(a)$ for all a can be achieved by solving a linear system with $n(2^n-1)$ equations (one per basis vector and per value $a \neq 0$) and $n\binom{n}{2}$ unknowns in $\mathbb{F}_2$, each modeling the presence of a specific degree 2 product in a specific output coordinate.

We have implemented such an algorithm and included it in `sboxU`.[3] Using it, we could verify that the ortho-derivative of F where F is any of the known quadratic APN functions on 6, 7, and 8 bits has a single ortho-integral (up to the addition of an affine function), namely F itself. Computing all the ortho-integrals of 25,624 8-bit quadratic APN functions takes about 20min on a regular desktop computer, meaning roughly half a second per function.

5 Relation with some EA- and CCZ-Invariants

Kaleyski and his coauthors, motivated by two different problems, introduced several invariants for Sboxes in [7] and [16]. The CCZ-invariant presented in [7], called *distance invariant*, applies to APN functions and provides a lower bound on the distance between two APN functions. It has a simplified form in the case of quadratic APN functions. It is observed that this CCZ-invariant takes many distinct values for APN functions in dimension 8, while it takes the same value for all APN functions in dimension 7 constructed in [12]. *Zero-sum invariants* are EA-invariant presented in [16], and are denoted by $\Sigma_k^F(0)$. It turns out that

[3] `https://github.com/lpp-crypto/sboxU`, see in particular file `quadratic.py` starting from line 62.

they are closely related to ortho-derivatives, as we will see below. First, we recall both definitions.

Proposition 4 (Zero-sum invariants [16]). *Let $\Sigma_k^F(t)$ be the multiset defined by $\Sigma_k^F(t) = \left\{\sum_{i=1}^k F(x_i) : x_1 + \cdots + x_k = t\right\}$. Then, for any $G = A_1 \circ F \circ A_2 + A_3$, where A_1 and A_2 are affine permutations and A_3 an affine function, we have*

$$\Sigma_k^G(0) = \begin{cases} \{A_1(s) + A_1(0) : s \in \Sigma_k^F(0)\} & \text{if } k \text{ is even} \\ \{A_1(s) + A_3(0) : s \in \Sigma_k^F(A_2(0))\} & \text{if } k \text{ is odd.} \end{cases}$$

It follows that the multiplicities of the elements in $\Sigma_k^F(0)$, i.e. the values $M_k^F(s) = \#\{(x_1, \ldots, x_k) : x_1 + \cdots + x_k = 0 \text{ and } \sum_{i=1}^k F(x_i) = s\}$ where $s \in \mathbb{F}_2^n$ are an EA-invariant for any even k. The same property holds for odd k if A_2 is linear.

Proposition 5 (Distance invariant [7]). *Let Φ_F be the multiset defined as $\Phi_F = \{\phi_F(b, c) : b, c \in \mathbb{F}_2^n\}$ where*

$$\phi_F(b, c) = \#\{a \in \mathbb{F}_2^n : \exists x \in \mathbb{F}_2^n,\ F(x) + F(x + a) + F(a + c) = b\}\ .$$

Then Φ_F is invariant under CCZ-equivalence. Moreover, if F is quadratic, then Φ_F is equal to the multiset $\Phi_F^0 = \{\phi_F(b, 0) : b \in \mathbb{F}_2^n\}$ where each element is repeated 2^n times.

In the case of quadratic functions, and when $k = 4$, it holds that these invariants are related with each other, and with the ortho-derivative. To establish this, we first re-write the definition of $M_4^F(s)$ and obtain the following lemma.

Lemma 2. *It holds that $M_4^F(s) = \#\{(x, a, b) : \Delta_a\Delta_b F(x) = s\}$.*

The following proposition then links $M_4^F(s)$ to $\phi_F(s, 0)$, i.e. it links the zero-sum invariant to the distance invariant.

Proposition 6. *F is APN if and only if $M_4(0) = 2^{2n+1} + 2^{2n} - 2^{n+1}$. Moreover, if $\deg F = 2$,*

$$M_4^F(s) = 2^n \#\{(a, b) : \Delta_a\Delta_b F(0) = s\},$$

and if F is a quadratic APN function with $F(0) = 0$, then for any nonzero $s \in \mathbb{F}_2^n$ we have $\phi_F(s, 0) = 2^{-(n+1)} M_4^F(s)$.

Proof. Because of Lemma 2, $M_4^F(s) = \#\{(x, a, b) : \Delta_a\Delta_b F(x) = s\}$.

APN functions are characterized by the fact that their second-order derivatives $\Delta_b\Delta_a F(x)$ never take the value 0 unless $\langle a, b\rangle$ does not form a 2-dimensional vector space, which occurs if $a = 0$, or $b = 0$ or $a = b$, *i.e.* $2(2^n - 1) + 2^n = 3 \times 2^n - 2$ times. For each such case, $\Delta_a\Delta_b F$ is the all-zero function, implying that

$$M_4^F(0) \geqslant 2^{2n+1} + 2^{2n} - 2^{n+1}\ ,$$

with equality if and only if F is APN.

When F is quadratic, all its second-order derivatives are constant, implying that $\Delta_b\Delta_a F(x)$ takes the same value for all $x \in \mathbb{F}_2^n$. Moreover, if $F(0) = 0$, then

$$\phi_F(s, 0) = \#\{a \in \mathbb{F}_2^n : \exists b \in \mathbb{F}_2^n, \Delta_a F(b) + \Delta_a F(0) = s\} .$$

Then, if F is a quadratic APN function, each function $b \mapsto \Delta_a F(b) + \Delta_a F(0)$ is a linear 2-to-1 function when $a \neq 0$, which implies that, for any $s \neq 0$,

$$\begin{aligned} 2^{-n} M_4^F(s) &= \#\{(a, b) : \Delta_a F(b) + \Delta_a F(0) = s\} \\ &= 2\#\{a : \exists b \in \mathbb{F}_2^n, \Delta_a F(b) + \Delta_a F(0) = s\} . \end{aligned}$$

Up to a factor 2^{n+1}, the values in Φ_F^0 then correspond to the values $M_4^F(s)$ when s varies in $\mathbb{F}_2^n \setminus \{0\}$. □

We are now ready to describe the connection between the invariants of Kaleyski and the ortho-derivative. It is described by the following proposition.

Proposition 7. *Let F be a quadratic APN function and π be its nontrivial ortho-derivative. Then, for any $s \in \mathbb{F}_2^n \setminus \{0\}$, $M_4^F(s) = 2^{n+1}(2^n - 1 - wt(\pi_s))$ where π_s denotes the component function $x \mapsto s \cdot \pi(x)$. Most notably, if n is odd, then π is a permutation, which implies the following equivalent statements:*

$$M_4^F(s) = \begin{cases} 2^{2n+1} + 2^{2n} - 2^{n+1} & \text{if } s = 0 \\ 2^{2n} - 2^{n+1} & \text{if } s \neq 0 \end{cases}, \quad \phi_F(b, 0) = \begin{cases} 2^n & \text{if } b = 0 \\ 2^{n-1} - 1 & \text{if } b \neq 0 . \end{cases}$$

Proof. By definition of π, when F is a quadratic APN function, for any nonzero $a \in \mathbb{F}_2^n$, the image set of $b \mapsto \Delta_a\Delta_b F(0)$ is the hyperplane composed of all elements orthogonal to $\pi(a)$. It follows that, when $s \neq 0$,

$$\begin{aligned} M_4^F(s) &= 2^n\#\{(a, b) : \Delta_a\Delta_b F(0) = s\} = 2^n\#\{(a, b), a \neq 0 : \Delta_a\Delta_b F(0) = s\} \\ &= 2^{n+1}\#\{a \neq 0 : s \in \langle \pi(a) \rangle^{\perp}\} = 2^{n+1}\#\{a \neq 0 : s \cdot \pi(a) = 0\} \\ &= 2^{n+1}(2^n - 1 - wt(\pi_s)) . \end{aligned}$$

□

Thus, for quadratic APN functions, up a to simple transformation, the multiset $\{M_4^F(s) : s \neq 0\}$ (and the equivalent invariant Φ_F^0) used in [7,16] is included in the multiset formed by the Walsh spectrum of the ortho-derivative.

6 Conclusion

The practical usefulness of the ortho-derivative was established in [8], and then confirmed in [2] and [21]. In this work, we have shed some more light on the properties of this object, and in particular established an upper bound on its degree. We also showed how to obtain a function given its ortho-derivative, which begs the question: what makes a function an ortho-derivative? It indeed remains an open problem to find other properties to efficiently determine whether a given function may be an ortho-derivative, since such a result could now allow us to construct new quadratic APN functions

References

1. Beierle, C., Leander, G.: New Instances of Quadratic APN Functions in Dimension Eight (Sep 2020). https://doi.org/10.5281/zenodo.4030734, https://doi.org/10.5281/zenodo.4030734
2. Beierle, C., Leander, G.: New instances of quadratic APN functions. IEEE Trans. Inform. Theory **68**(1), 670–678 (2022). https://doi.org/10.1109/TIT.2021.3120698, https://doi.org/10.1109/TIT.2021.3120698
3. Bending, T.D., Fon-Der-Flaass, D.: Crooked functions, bent functions, and distance regular graphs. Electr. J. Comb. **5** (1998), http://www.combinatorics.org/Volume_5/Abstracts/v5i1r34.html
4. Berger, T., Canteaut, A., Charpin, P., Laigle-Chapuy, Y.: On almost perfect nonlinear functions over $\mathbb{F}_2^n$. IEEE Trans. Inform. Theory **52**(9), 4160–4170 (2006)
5. Boura, C., Canteaut, A., Jean, J., Suder, V.: Two notions of differential equivalence on sboxes. Des. Codes Cryptogr. **87**(2-3), 185–202 (2019). https://doi.org/10.1007/S10623-018-0496-Z
6. Browning, K.A., Dillon, J., McQuistan, M.T., Wolfe, A.J.: An APN permutation in dimension six. In: Post-proceedings of the 9-th International Conference on Finite Fields and Their Applications. vol. 518, pp. 33–42. American Mathematical Society (2010)
7. Budaghyan, L., Carlet, C., Helleseth, T., Kaleyski, N.S.: On the distance between APN functions. IEEE Trans. Inform. Theory **66**(9), 5742–5753 (2020). https://doi.org/10.1109/TIT.2020.2983684, https://doi.org/10.1109/TIT.2020.2983684
8. Canteaut, A., Couvreur, A., Perrin, L.: Recovering or testing extended-affine equivalence. IEEE Trans. Inform. Theory **68**(9), 6187–6206 (2022). https://doi.org/10.1109/TIT.2022.3166692, https://doi.org/10.1109/TIT.2022.3166692
9. Canteaut, A., Naya-Plasencia, M.: Structural weaknesses of permutations with a low differential uniformity and generalized crooked functions. In: Finite fields: theory and applications, Contemp. Math., vol. 518, pp. 55–71. Amer. Math. Soc., Providence, RI (2010). https://doi.org/10.1090/conm/518/10196
10. Carlet, C., Charpin, P., Zinoviev, V.: Codes, bent functions and permutations suitable for DES-like cryptosystems. Des. Codes Cryptogr. **15**(2), 125–156 (1998)
11. Charpin, P.: The crooked property. Finite Fields Appl. **81** (2022)
12. Edel, D., Pott, A.: A new almost perfect nonlinear function which is not quadratic. Adv. Math. Commun. **3**(1), 59–81 (2009)
13. Gorodilova, A.: On the differential equivalence of APN functions. Cryptogr. Commun. **11**(4), 793–813 (2019). https://doi.org/10.1007/S12095-018-0329-Y
14. Gorodilova, A.: A note on the properties of associated boolean functions of quadratic APN functions. Прикладная дискретная математика **47**, 16–21 (2020)
15. Heden, O.: A survey of the different types of vector space partitions. Discrete Math., Alg. and Appl. **4**(1) (2012)
16. Kaleyski, N.S.: Invariants for EA- and CCZ-equivalence of APN and AB functions. Cryptogr. Commun. **13**(6), 995–1023 (2021). https://doi.org/10.1007/S12095-021-00541-8, https://doi.org/10.1007/s12095-021-00541-8
17. Kyureghyan, G.M.M.: Crooked maps in $\mathbb{F}_{2^n}$. Finite Fields Appl. **13**(3), 713–726 (2007)
18. Nyberg, K.: Perfect nonlinear S-boxes. In: Davies, D.W. (ed.) EUROCRYPT'91. LNCS, vol. 547, pp. 378–386. Springer, Heidelberg, Germany, Brighton, UK (Apr 8–11, 1991). https://doi.org/10.1007/3-540-46416-6_32

19. Nyberg, K.: Differentially uniform mappings for cryptography. In: Helleseth, T. (ed.) EUROCRYPT'93. LNCS, vol. 765, pp. 55–64. Springer, Heidelberg, Germany, Lofthus, Norway (May 23–27, 1994). https://doi.org/10.1007/3-540-48285-7_6
20. van Dam, E., Fon-Der-Flaass, D.: Codes, graphs, and schemes from nonlinear functions. European J. Combin. **24**(1), 85–98 (2003). https://doi.org/https://doi.org/10.1016/S0195-6698(02)00116-6
21. Yu, Y., Perrin, L.: Constructing more quadratic APN functions with the QAM method. Cryptogr. Commun. **14**(6), 1359–1369 (2022). https://doi.org/10.1007/S12095-022-00598-Z, https://doi.org/10.1007/s12095-022-00598-z
22. Yu, Y., Wang, M., Li, Y.: A matrix approach for constructing quadratic APN functions. Des. Codes Cryptogr. **73**(2), 587–600 (Nov 2014). https://doi.org/10.1007/s10623-014-9955-3, https://doi.org/10.1007/s10623-014-9955-3

On the algebraic degree stability of Boolean functions when restricted to affine spaces *

Claude Carlet[1], Serge Feukoua[2], and Ana Sălăgean[3]

[1] LAGA, Department of Mathematics, University of Paris 8 (and Paris 13 and CNRS), Saint–Denis cedex 02, France, and University of Bergen, Norway.
claude.carlet@gmail.com
[2] Department of Computer Science, University of Loughborough, UK; ENSTP Yaoundé, Cameroon. S.C.Feukoua-Jonzo@lboro.ac.uk; sergefeukoua@gmail.com
[3] Department of Computer Science, University of Loughborough, UK.
A.M.Salagean@lboro.ac.uk

Abstract. We study the n-variable Boolean functions which keep their algebraic degree unchanged when they are restricted to any (affine) hyperplane, or more generally to any affine space of a given co-dimension k. For cryptographic applications it is of interest to determine functions f which have a relatively high degree and also maintain this degree when restricted to affine spaces of co-dimension k for k ranging from 1 to as high a value as possible. This highest value will be called the restriction degree stability of f, denoted by $\deg\_stab(f)$. We give several necessary and/or sufficient conditions for f to maintain its degree on spaces of co-dimension k. The value of $\deg\_stab(f)$ is determined for functions which are direct sums of monomial as well as for functions of degrees $r \in \{1, 2, n-2, n-1, n\}$; we also determine the symmetric functions which maintain their degree on any hyperplane. Finally, using our previous results and some computer assistance, we determine the behaviour of all the functions in 8 variables, therefore determining the optimal ones (i.e. with highest value of $\deg\_stab(f)$) for each degree.

Keywords: Boolean functions· affine spaces· restriction · algebraic degree.

1 Introduction

Boolean functions are used in many research areas; of particular relevance to our work is their use in cryptography, sequences and algebraic coding theory. The existence of some cryptanalysis techniques (such as correlation, fast correlation and algebraic attacks on stream ciphers, linear and differential attacks on block ciphers, see e.g. [10, 5, 1]), leads to several criteria required for the Boolean functions used in symmetric cryptography. These cryptographic functions should in

* The research of the first author is partly supported by the Norwegian Research Council and the two other authors are supported by EPSRC, UK (EPSRC grant EP/W03378X/1)

particular have a high algebraic degree. Indeed, almost all cryptosystems using Boolean functions (e.g. the filter or combiner model of stream ciphers) can be attacked if the considered functions have low algebraic degree (e.g. fast algebraic attacks [4, 1]). It is also important that this algebraic degree remains high even if the function is restricted to an affine hyperplane or to an affine space of low co-dimension to avoid "guess and determine attacks" where the attacker would make assumptions resulting in the fact that the input to the function is restricted to a particular affine space. In [2], an infinite class of functions is described whose algebraic degree remains unchanged when they are restricted to any affine hyperplane, but no general characterization was given.

In this paper, we start a systematic study of the functions which keep their degree unchanged when restricted to affine spaces of a certain co-dimension k. In Section 3 we define a "degree-drop" space of a function f as being an affine space A such that the restriction $f_{|A}$ has degree strictly lower than the degree of f. We also define the notion of restriction degree stability for a function f, denoted by $\deg\_stab(f)$, as being the largest k for which f has no degree-drop space of co-dimension k. The largest value of $\deg\_stab(f)$ over all functions of degree r in n variables will be denoted by $\deg\_stab(r, n)$. Functions which reach this value would be optimal from the point of view of their degree stability, and therefore of interest in cryptographic constructions.

Section 4 gives several necessary and/or sufficient conditions for a function f to have no degree-drop space of co-dimension k. Notably, we show that a space A of co-dimension k is not a degree-drop space for f if and only if $\deg(f1_A) = \deg(f)+k$ (where 1_A is the indicator function of A). Furthermore, we prove that a function f of degree r has a degree-drop hyperplane if and only if it is affine equivalent to a function $x_1g + h$ where g is homogeneous with $\deg(g) = r - 1$ and $\deg(h) < r$. We generalize this condition to any k. We then find more constructive sufficient conditions for the non-existence of degree-drop spaces.

In Section 5 we examine particular classes of functions. The main result, Theorem 5, shows that for functions f which are direct sums of p monomials of degree r (i.e. no variable appears in more than one monomial) the smallest co-dimension for which a degree-drop space exists is p. Therefore $\deg\_stab(f) = p - 1$, which also gives a lower bound of $\deg\_stab(r, n) \geq \lfloor \frac{n}{r} \rfloor - 1$. We also show in Theorem 6 that symmetric functions of degree r with $2 \leq r \leq n - 2$ have no degree-drop hyperplane if and only if r is even.

Finally, in Section 6 we use our previous results alongside computer calculations to determine the number of degree-drop spaces for all the functions in up to 8 variables. To do so, it suffices to examine the representatives of affine equivalence classes computed by [7] and [8]. We determined thus all the 8-variable functions which have optimal restriction degree stability .

2 Preliminaries

We denote by $\mathbb{F}_2$ the finite field with two elements and by $\mathbb{F}_2^n$ the vector space over $\mathbb{F}_2$ of all binary vectors of length n. A n-variable Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$

can be uniquely represented in *Algebraic Normal Form* (in brief, ANF) i.e. as a polynomial function in n variables, of degree at most one in each variable. The *algebraic degree* of f, denoted by $\deg(f)$, is the degree of its ANF. For every n-variable Boolean function f, we denote by $\mathrm{Var}(f)$ the set consisting of all the elements $i \in \{1, \ldots, n\}$ such that x_i appears in at least one term with nonzero coefficient in the ANF of f.

Definition 1. *Two Boolean functions $f, g : \mathbb{F}_2^n \to \mathbb{F}_2$ are said to be affinely equivalent, which is denoted by $f \sim g$, if there exists φ, an affine automorphism of $\mathbb{F}_2^n$, such that $f = g \circ \varphi$ where $\circ$ is the operation of composition.*

The algebraic degree is invariant to affine equivalence.

The set of all n-variable Boolean functions of algebraic degree at most r shall be denoted by $RM(r, n)$. Let $RM(r, n)/RM(r-1, n)$ be the quotient space consisting of all cosets of $RM(r-1, n)$ in $RM(r, n)$. Unless otherwise specified, for each coset we will use as representative the homogeneous polynomial in the coset, i.e. the unique polynomial which contains only monomials of degree r. The equivalence $\sim$ can be extended naturally to an equivalence $\sim_{r-1}$ on $RM(r, n)/RM(r-1, n)$, or more generally on any $RM(d, n)/RM(r-1, n)$ with $d \geq r$. Namely, two $RM(r-1, n)$-cosets are equivalent under $\sim_{r-1}$ if there is a function f_1 in one coset and a function g_1 in the other coset such that $f_1 \sim g_1$. Equivalently, two functions f and g satisfy $f \sim_{r-1} g$ if and only if there is a function h such that $f \sim h$ and $\deg(g - h) \leq r - 1$.

For any monomial m we shall denote by m^c the complement $\prod_{i \in \{1,\ldots n\} \setminus \mathrm{Var}(m)} x_i$ (which also equals $\frac{x_1 \cdots x_n}{m}$). If $f = \sum_{i=1}^{p} m_i$, where the m_i are all monomials of degree r, then we define $f^c = \sum_{i=1}^{p} m_i^c$, also called the complement of f.

Given a function f, let us examine how we can obtain the ANF of $f_{|A}$, the restriction of f to an affine space A. The affine space A can be viewed as the set of solutions of k affine equations, that are, after Gaussian elimination: $x_{i_1} = a_{i_1}(y), \ldots, x_{i_k} = a_{i_k}(y)$, where $i_1, \ldots, i_k$ are distinct and $a_j(y)$ are affine functions in the $n-k$ variables $\{y_1, \ldots, y_{n-k}\} = \{x_1, \ldots, x_n\} \setminus \{x_{i_1}, \ldots, x_{i_k}\}$. We obtain one of the expressions for the ANF of $f_{|A}$ by substituting the variables $x_{i_1}, \ldots, x_{i_k}$ with $a_{i_1}(y), \ldots, a_{i_k}(y)$ in the ANF of f, obtaining a function in the remaining $n-k$ variables $\{y_1, \ldots, y_{n-k}\}$. This function depends on the equations used for defining A (the choice of $i_1, \ldots, i_k$ is not unique), but all choices yield the ANF of functions which are affinely equivalent to each other and therefore have the same degree. Clearly, $\deg(f_{|A}) \leq \deg(f)$.

3 Degree-drop spaces: definition and basic properties

In this paper, we are interested in the behaviour of the algebraic degree of Boolean functions when they are restricted to affine spaces of a certain codimension. We introduce the following terminology:

Definition 2. *Let f be an n-variable Boolean function and A an affine subspace of $\mathbb{F}_2^n$. If $\deg(f_{|A}) < \deg(f)$, then we call A a* degree-drop subspace *for f. The*

largest k such that f has no degree-drop subspace of co-dimension k will be called the restriction degree stability *of f, denoted by* $\deg\_\mathrm{stab}(f)$ *(with* $\deg\_\mathrm{stab}(f) = 0$ *if f has degree-drop hyperplanes).*

Note that for a fixed affine space A, the property of A being a degree-drop space for a function f depends only on the monomials of f of algebraic degree $\deg(f)$. Hence, for any polynomials f and g in the same coset of $RM(r,n)/RM(r-1,n)$, we have that A is a degree-drop subspace of f if and only if A is a degree-drop subspace of g. It suffices therefore to study the degree-drop spaces of homogeneous polynomials.

Notation 1 *For all integers k, r, n with $1 \leq k \leq n$ and $1 \leq r \leq n$, we denote by $K_{k,r,n}$ the subset of $RM(r,n)/RM(r-1,n)$ consisting of the nonzero elements f which admit no degree-drop subspace of co-dimension k.*
We denote by $\deg\_\mathrm{stab}(r,n)$ *the largest value of the restriction degree stability among all functions of degree r in n variables, i.e. the largest k with $K_{k,r,n} \neq \emptyset$.*

We first collect a few preliminary observations:

Lemma 1. *Let f be a homogeneous function of degree r in n variables and let $1 \leq k \leq n$.*
(i) If H is a hyperplane defined by an equation involving a variable which is not in $\mathrm{Var}(f)$*, then H is not a degree-drop hyperplane of f.*
(ii) The hyperplane H defined by the equation $x_j = 0$ is a degree-drop hyperplane for f if and only if $f(x_1, \ldots, x_n) = x_j g(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ for some homogeneous polynomial g of degree $r-1$.
(iii) If f has only one monomial in its ANF, then f has $2^r - 1$ degree-drop linear hyperplanes.
(iv) If $r \geq 2$ and $f = \sum_{i=1}^{p} m_i$ with m_i monomials (i.e. f has p monomials in its ANF), then f has a degree-drop space of co-dimension p and therefore $\deg\_\mathrm{stab}(f) \leq p - 1$.
(v) $K_{k+1,r,n} \subseteq K_{k,r,n}$.
(vi) If $k > n - r$ then any space of co-dimension k is a degree-drop space for f and $K_{k,r,n} = \emptyset$. In particular, if f has degree n then all spaces of any co-dimension $k \geq 1$ are degree-drop spaces for f and $K_{k,n,n} = \emptyset$.
(vii) Any function of degree 1 has degree-drop hyperplanes, i.e. $K_{k,1,n} = \emptyset$.
(viii) $\deg\_\mathrm{stab}(r, n+1) \leq \deg\_\mathrm{stab}(r,n) + 1$.
(ix) $f \in K_{k,r,n}$ if and only if $f \in K_{k,r,n+1}$ (viewing f as a homogeneous function in $n+1$ variables). In this sense, $K_{k,r,n} \subseteq K_{k,r,n+1}$ and therefore $\deg\_\mathrm{stab}(r,n) \leq \deg\_\mathrm{stab}(r,n+1)$.

Allowing a degree drop subspace is clearly an affine invariant property:

Lemma 2. *Let f, g be n-variable functions of degree r such that $f \sim_{r-1} g$, i.e. $g = f \circ \varphi + h$ for some affine automorphism φ of $\mathbb{F}_2^n$ and some function h with* $\deg(h) < r$*. Let A be an affine space of $\mathbb{F}_2^n$ of co-dimension k. Then $g_{|_A} \sim_{r-1} f_{|_{\varphi(A)}}$. Therefore, A is a degree-drop space for g if and only if $\varphi(A)$ is a degree-drop space for f. Consequently, $f \in K_{k,r,n}$ if and only if $g \in K_{k,r,n}$.*

The next result shows that in order to decide whether a function has degree-drop affine spaces it suffices to consider linear spaces.

Lemma 3. *Let f be an n-variable Boolean function and $A = v + E$ be an affine space in $\mathbb{F}_2^n$, where E is a vector subspace and v a vector. Then $deg(f_{|A}) = deg(f)$ if and only if $deg(f_{|E}) = deg(f)$*

4 Necessary and/or sufficient conditions for a function not to admit a degree-drop space

The following necessary condition for a function to belong to $K_{k,r,n}$ was proved in [2] (and it is easy to check):

Lemma 4. *([2]) Let k, r, n satisfy $1 \leq k \leq n$ and $1 \leq r \leq n$ and let $f = \sum_{i=1}^{p} m_i$ where m_i are monomials of degree r. If $f \in K_{k,r,n}$ then*

$$\cap_{i=1}^{p} \mathrm{Var}(m_i) = \emptyset. \tag{1}$$

In particular, $f \in K_{k,r,n}$ implies $p > 1$.

Lemma 4 can be generalized as follows:

Lemma 5. *Let f be a homogeneous n-variable Boolean function of algebraic degree $r \geq 2$ and write $f = \sum_{i=1}^{p} m_i$ with m_i monomials. Let $1 \leq k \leq n$. If $f \in K_{k,r,n}$, then for any set of k distinct variables $x_{j_1}, ..., x_{j_k}$ there is at least one monomial in f which does not contain any of the variables $x_{j_1}, ..., x_{j_k}$*

The conditions in Lemmas 4 and 5 above are necessary, but not sufficient for a function to have no degree-drop spaces of a given co-dimension k. However, they become sufficient when we extend them by affine equivalence.

Theorem 2. *Let f be a Boolean homogeneous function of algebraic degree $r \geq 2$ in n variables and let $a = (a_1, \ldots, a_n) \in \mathbb{F}_2^n \setminus \{\mathbf{0}\}$ and $a_0 \in \mathbb{F}_2$. The following two statements are equivalent:*
(i) f has a degree-drop hyperplane H defined by the equation $a_0 + \sum_{i=1}^{n} a_i x_i = 0$.
(ii) f can be written as $f(x) = (a_0 + \sum_{i=1}^{n} a_i x_i)\, g(x) + c(x)$ for some polynomials g, c with $\deg(g) = r - 1$ and $\deg(c) \leq r - 1$.
Consequently, f has a degree-drop hyperplane if and only if $f(x) \sim_{r-1} x_1 g(x_2, \ldots, x_n)$ for some homogeneous polynomial g of degree $r - 1$ in $n - 1$ variables.

Theorem 3. *Let f be an n-variable homogenous Boolean function of algebraic degree $r \leq n$ and let $k \leq n$. The following statements are equivalent:*
(i) f has a degree-drop affine space of co-dimension k.
(ii) $f \sim_{r-1} g$ for some homogeneous function g of degree r such that each monomial of g contains at least one of the variables $x_1, x_2, ..., x_k$.

Let us recall that for any set $A \subseteq \mathbb{F}_2^n$ the function 1_A, called the indicator of A, is the Boolean function such that $1_A(x) = 1$ if and only if $x \in A$. For any Boolean function f, the product $f1_A$ is the Boolean function which is equal to $f(x)$ if $x \in A$ and is null otherwise. As far as we know, the next lemma has never been explicitly stated in a paper, while it is rather basic.

Lemma 6. *Let f be an n-variable Boolean function. Let A be an affine subspace of $\mathbb{F}_2^n$ and 1_A its indicator function. If $f1_A$ is not the identically zero function, we have*

$$\deg(f1_A) = \deg(f_{|A}) + \deg(1_A).$$

Proof. Let k be the co-dimension of A. By Lemma 2 we can assume that A is defined by the equations $x_{n-k+1} = \cdots = x_n = 1$. Then $1_A(x) = \prod_{i=n-k+1}^{n} x_i$. The expression of the ANF of $f_{|A}$ is obtained from the ANF of f by substituting x_i by 1 for each $i = n-k+1, \ldots, n$. This same expression, when multiplied by $\prod_{i=n-k+1}^{n} x_i$, gives the ANF of $(f1_A)(x)$ (since $(f1_A)(x) = f(x)$ for any $x \in A$ and $(f1_A)(x) = 0$ for any $x \notin A$). Note that the ANF of $f_{|A}$ and the ANF of 1_A have no variables in common, which means that the degree of their product (if the product is not zero) is the sum of their degrees. Since $f1_A \neq 0$, this means $\deg(f1_A) = \deg(f_{|A}) + \deg(1_A)$. □

Proposition 1. *The affine space A of co-dimension k is not a degree-drop space for the n-variable function f if and only if* $\deg(f1_A) = \deg(f) + k$.

Theorems 2 and 3 allow us to easily construct functions that have degree-drop affine spaces. However, if we want to construct a function which does not have any degree-drop hyperplane, they are less useful. For Theorem 2, it is indeed unfeasible (exponential complexity) to check that f is not affine equivalent to any function of the form $x_1 g(x_2, \ldots, x_n)$, and in the case of Theorem 3, things seem still more complex. The following results, which under additional constraints make the conditions in Lemmas 4 and 5 sufficient (but no longer necessary), allow efficient constructions of families of functions with no degree-drop hyperplanes.

Proposition 2. *Let $f = \sum_{i=1}^{p} m_i$ be an n-variable homogeneous Boolean function of algebraic degree r with $2 \leq r \leq n$, where the m_i's are monomials.*
If f satisfies Condition (1) and the condition

$$|\mathrm{Var}(m_i) \cap \mathrm{Var}(m_j)| \leq r-2, \text{ for all } i \neq j \in \{1, \ldots, p\}, \tag{2}$$

then $f \in K_{1,r,n}$ i.e. f has no degree-drop hyperplane.
More generally, for any $k < r$, if

$$|\mathrm{Var}(m_i) \cap \mathrm{Var}(m_j)| \leq r-k-1, \text{ for all } i \neq j \in \{1, \ldots, p\}, \tag{3}$$

and for any set of k distinct variables $x_{j_1}, ..., x_{j_k}$, there is at least one monomial in f which does not contain any of the variables $x_{j_1}, ..., x_{j_k}$, then $f \in K_{k,r,n}$.

The construction in the first part of Proposition 2 can be generalized:

Theorem 4. *Let $f = \sum_{j=1}^{p} m_j$ be a homogenous n-variable function of degree r. If for all $i \in \mathrm{Var}(f)$, there exists a monomial m_{j_i} in f such that:*

- $i \notin \mathrm{Var}(m_{j_i})$
- *for all $t \in \mathrm{Var}(m_{j_i})$, the monomial $\frac{x_i m_{j_i}}{x_t}$ is not in f,*

then, f has no degree-drop hyperplane, i.e. $f \in K_{1,r,n}$.

In general, if f has no degree-drop hyperplanes it does not necessarily mean that the same is true for f^c. We can however give a sufficient condition similar to Proposition 2:

Proposition 3. *Let f be a homogenous Boolean function of algebraic degree r in n variables defined by $f = \sum_{i=1}^{p} m_i$ with m_i monomials. If f satisfies Condition (2) and and is such that $\cup_{i=1}^{p} \mathrm{Var}(m_i) = \{1, 2, ..., n\}$ then $f^c \in K_{1,n-r,n}$.*
More generally, for any $k < r$, if f satisfies Condition (3) and is such that, for any k distinct variables $x_{i_1}, ..., x_{i_k}$, there exists a monomial m_j such that $\{i_1, ..., i_k\} \subseteq \mathrm{Var}(m_j)$ then $f^c \in K_{k,n-r,n}$.

5 Special classes of functions

In this section we study the existence of degree-drop spaces for functions of degree $2, n-2, n-1$ (degrees 1 and n are covered in Lemma 1), functions which are direct sums of monomials, and finally symmetric functions. For degree $r = n-1$ we have:

Lemma 7. *Any n-variable Boolean function of degree $n-1$ has $2^{n-1}-1$ degree-drop linear hyperplanes. Therefore $K_{1,n-1,n} = \emptyset$ and $\mathrm{deg\_stab}(n-1, n) = 0$.*

Concerning the degree 2 functions, recall that any quadratic function f in n variables is equivalent under $\sim_1$ with $x_1x_2 + \cdots + x_{2p-1}x_{2p}$ for some $p \geq 1$ such that $2p \leq n$ (see [9, 1]). The case of $p = 1$ (i.e. one monomial) is trivial by Lemma 1(iii). A part of the next proposition has been addressed in [3, Lemma 3]:

Proposition 4. *Let n and p be such that $2 \leq p \leq \lfloor \frac{n}{2} \rfloor$. The function*

$$f(x_1, \ldots, x_n) = x_1x_2 + x_3x_4 + \cdots + x_{2p-1}x_{2p}$$

has no degree-drop space of co-dimension $p-1$ but has degree-drop spaces of co-dimension p; hence, $\mathrm{deg\_stab}(f) = p-1$ and $\mathrm{deg\_stab}(2, n) = \lfloor \frac{n}{2} \rfloor - 1$.

For the degree $r = n-2$, using the fact that $f \sim_{r-1} g$ if and only if $f^c \sim_{n-r-1} g^c$ (see [7, Section 4]) we have:

Proposition 5. *Let $n \geq 4$ be an integer. We have:*
(i) if n is odd, then $K_{1,n-2,n} = \emptyset$ and therefore $\mathrm{deg\_stab}(n-2, n) = 0$,
(ii) if n is even, then $K_{1,n-2,n} = \{f : f^c \sim_1 x_1x_2 + \cdots + x_{n-1}x_n\}$ and $K_{2,n-2,n} = \emptyset$. Therefore $\mathrm{deg\_stab}(n-2, n) = 1$.

Before giving the general result regarding the direct sum of monomial functions, we shall need the following definition and lemmas. Recall that the rank of a function f in n variables is the minimum integer n_1 such that there is a function g which depends on n_1 variables and $g \sim f$. This notion can be extended to the quotient $RM(r, n)/RM(r-1, n)$, by defining the rank of an element $f \in RM(r, n)/RM(r-1, n)$, denoted by $\mathrm{rank}_{r-1}(f)$, as the minimum integer n_1 such that there exists $g \in RM(r, n)/RM(r-1, n)$ such that g depends on n_1 variables and $f \sim_{r-1} g$. Obviously, rank_{r-1} is invariant under $\sim_{r-1}$.

Lemma 8. *Let f be a homogeneous function of degree r in n variables with $2 \leq r \leq n$. If f has the property that for any two distinct monomials m, m' of f we have $|\mathrm{Var}(m) \cap \mathrm{Var}(m')| \leq r-2$, then $\mathrm{rank}_{r-1}(f) = |\mathrm{Var}(f)|$. In particular, if f is a direct sum of p degree r monomials, then $\mathrm{rank}_{r-1}(f) = pr$.*

Lemma 9. *Let f be a degree-r function in n variables and let A be an affine subspace of $\mathbb{F}_2^n$ of co-dimension k. Then $\mathrm{rank}_{r-1}(f_{|A}) \leq \min(n-k, \mathrm{rank}_{r-1}(f))$.*

Theorem 5. *Let $2 \leq r < n$ and $2 \leq p \leq \lfloor \frac{n}{r} \rfloor$. The function of degree r in n variables which is the direct sum of p monomials*

$$f(x_1, \ldots, x_n) = x_1 x_2 \cdots x_r + x_{r+1} x_{r+2} \cdots x_{2r} + \cdots + x_{(p-1)r+1} x_{(p-1)r+2} \cdots x_{pr}$$

has no degree-drop space of co-dimension $p-1$ but has degree-drop spaces of co-dimension p, i.e. $\deg\_\mathrm{stab}(f) = p-1$. *Consequently* $\deg\_\mathrm{stab}(r,n) \geq \lfloor \frac{n-r}{r} \rfloor$.

Proof. By Lemma 1 (iv), f has degree-drop spaces of co-dimension p. We need to prove that no affine space A of co-dimension $p-1$ can be a degree-drop hyperplane. When $p \leq r$, this follows from Proposition 2. Assume now $p > r$. The case $r = 2$ was proven in Proposition 4, so we can assume $r \geq 3$. Assume, for a contradiction, that A is a degree-drop subspace for f. Let the equations that define A be (in diagonalized form) $x_{i_1} = a_{i_1}(y), \ldots, x_{i_{p-1}} = a_{i_{p-1}}(y)$, where $i_1, \ldots, i_{p-1}$ are distinct and $a_{i_1}(y), \ldots, a_{i_{p-1}}(y)$ are affine functions in the $n-(p-1)$ variables $\{y_1, \ldots, y_{n-(p-1)}\} = \{x_1, \ldots, x_n\} \setminus \{x_{i_1}, \ldots, x_{i_{p-1}}\}$. To compute $f_{|A}$ we substitute the variables $x_{i_1}, \ldots, x_{i_{p-1}}$ with $a_{i_1}(y), \ldots, a_{i_{p-1}}(y)$ respectively. We partition the monomials of f, writing $f = f_0 + f_1 + f_2$ where the monomials in f_0 have no variable to be substituted, the monomials in f_1 have exactly one variable that will be substituted and the monomials in f_2 have two or more variables that will be substituted. Let p_0, p_1, p_2 be the number of monomials in f_0, f_1, f_2 respectively. Obviously $p_0 + p_1 + p_2 = p$. There are $p-1$ variables to be substituted, p_1 of them are in f_1 and at least $2p_2$ are in f_2; therefore $p - 1 \geq 2p_2 + p_1$. Using $p_0 + p_1 + p_2 = p$, this implies $p_0 > p_2$.

As f contains p monomials and no variable appears in more than one monomial, there must exist at least one monomial m which does not contain any of $x_{i_1}, ..., x_{i_{p-1}}$ (i.e. m is a monomial of f_0) and therefore remains unchanged after substitution. To cancel m there must be at least one other monomial m' of f such that all the variables of m' are substituted (therefore m' is a monomial of f_2). In other words, $p_0 \neq 0$ and $p_2 \neq 0$. Obviously $f_{|A} = (f_0)_{|A} + (f_1)_{|A} + (f_2)_{|A}$. Since we assumed A is a degree-drop space for f, we have $\deg(f_{|A}) < r$, which means $f_{|A} \sim_{r-1} 0$ and therefore

$$\mathrm{rank}_{r-1}((f_0)_{|A} + (f_1)_{|A}) = \mathrm{rank}_{r-1}(f_{|A} + (f_2)_{|A}) = \mathrm{rank}_{r-1}((f_2)_{|A}). \qquad (4)$$

Since no variables are substituted in f_0 and only one variable is substituted in each monomial of f_1, the monomials of f_0 do not cancel out in $(f_0)_{|A} + (f_1)_{|A}$, which means $\mathrm{Var}(f_0) \subseteq \mathrm{Var}((f_0)_{|A} + (f_1)_{|A})$. Moreover, any two monomials in $(f_0)_{|A} + (f_1)_{|A}$ have at most one variable in common, and therefore at most $r-2$ variables in common (since $r \geq 3$). We apply Lemma 8 to $(f_0)_{|A} + (f_1)_{|A}$,

obtaining $\mathrm{rank}_{r-1}((f_0)_{|A}+(f_1)_{|A}) = |\mathrm{Var}((f_0)_{|A}+(f_1)_{|A})| \geq |\mathrm{Var}(f_0)| = rp_0$. On the other hand, by Lemma 9, $\mathrm{rank}_{r-1}((f_2)_{|A}) \leq \mathrm{rank}_{r-1}(f_2) = rp_2$. Combining these inequalities with (4), we obtain $rp_0 \leq rp_2$, i.e. $p_0 \leq p_2$, contradicting the inequality $p_0 > p_2$ proven at the end of the previous paragraph. □

Corollary 1. *We have the following bounds when* $2 \leq r \leq n-1$*:*

$$\left\lfloor \frac{n-r}{r} \right\rfloor \leq \mathrm{deg\_stab}(r,n) \leq n-r-1.$$

When $r = 2$ *equality is achieved for the lower bound; when* $r = n-1$ *or when* $r = n-2$ *and* n *is even, equality is achieved for the upper bound.*

Let us consider now the class of symmetric functions i.e. functions which are invariant to all permutations of the variables. It contains the class of majority functions (or more generally threshold functions) which is known for having optimal algebraic immunity (see [6, 1]).

Theorem 6. *Let* f *be a symmetric Boolean function in* n *variables of degree* r*, with* $2 \leq r \leq n-2$*.*
(i) If r *is even, then* f *has no degree-drop hyperplane.*
(ii) If r *is odd, then* f *has exactly one degree-drop linear hyperplane of equation* $x_1 + x_2 + \ldots + x_n = 0$*.*

6 Experimental results

We have computed the number of degree-drop spaces for all the functions in $n = 8$ variables of degrees $3 \leq r \leq n-3$ (the other degrees having been settled for any n in Lemma 1 and Section 5).

In [7], Hou showed that there are 31 non-zero classes of polynomials of degree 3 in 8 variables, under the equivalence $\sim_2$. We recall them in the Appendix. For each function f among the representatives $f_2, \ldots, f_{32}$ of these 31 classes, we considered each linear space V of co-dimension up to 3 and we determined whether V is a degree-drop space for f by computing the degree of the restriction of f to V. We counted the number of degree-drop linear spaces of f of each co-dimension from 1 to 3. Moreover, for each degree-drop subspace of co-dimension $k \in \{2,3\}$ of f, we determined whether it is a "new" degree-drop space, in the sense that it is not a subspace of a degree-drop subspace of co-dimension $k-1$. These 5 values are presented for each function in Table 1, in lexicographically decreasing order. For degree 5 in 8 variables, the representatives under $\sim_4$ are the complements f_2^c,..., f_{32}^c (see [7, Section 4]). Since f_2,..., f_{12} are actually functions in 7 or less variables, when we view them as functions in 8 variables and take the complement we will obtain functions where all the monomials contain the variable x_8, and therefore they have a degree-drop hyperplane.The other 20 representatives f_{13},..., f_{32} have no degree-drop hyperplanes, but they all have degree-drop spaces of co-dimension 2, so $K_{2,5,8} = \emptyset$ and $\mathrm{deg\_stab}(5,8) = 1$.

Table 1. Number of degree-drop linear spaces of f of each co-dimension from 1 to 3 for the 31 representatives of degree 3 in 8 variables

Representative	co-dim 1 lin spaces	co-dim 2 lin spaces	co-dim 2 new lin spaces	co-dim 3 lin spaces	co-dim 3 new lin spaces
f_2	7	875	0	17795	0
f_3	1	187	60	6147	0
f_7	1	127	0	3747	1080
f_4	0	49	49	3059	168
f_5	0	35	35	2371	256
f_6	0	21	21	1683	360
f_8	0	13	13	1427	636
f_9	0	7	7	995	568
f_{13}	0	7	7	847	420
f_{16}	0	7	7	739	312
f_{10}	0	3	3	867	678
f_{29}	0	2	2	459	333
f_{11}	0	1	1	563	500
f_{14}	0	1	1	459	396
f_{15}	0	1	1	351	288
f_{24}	0	1	1	307	244
f_{17}	0	1	1	243	180
f_{28}	0	1	1	243	180
f_{26}	0	1	1	135	72
f_{12}	0	0	0	651	651
f_{31}	0	0	0	243	243
f_{18}	0	0	0	167	167
f_{25}	0	0	0	155	155
f_{19}	0	0	0	151	151
f_{30}	0	0	0	151	151
f_{22}	0	0	0	105	105
f_{23}	0	0	0	91	91
f_{32}	0	0	0	91	91
f_{21}	0	0	0	75	75
f_{20}	0	0	0	45	45
f_{27}	0	0	0	15	15

Langevin and Leander [8] computed a representative from each of the 998 classes of functions of degree 4 in 8 variables under the equivalence $\sim_3$. Again, for each of them we determined by computer calculations the number of linear degree-drop spaces of co-dimension $k = 1, 2, 3$. The function $x_1x_2x_3x_4$ has 15 degree-drop linear hyperplanes, as expected; the functions $x_1x_2(x_3x_4 + x_5x_6)$ and $x_1x_2(x_3x_4 + x_5x_6 + x_7x_8)$ each have 3 degree-drop linear hyperplanes; the functions x_8f_i where $f_i \in \{f_4, f_5, f_6, f_8, \ldots, f_{12}\}$ (with f_i being the representatives of function classes of degree 3 in 7 variables mentioned above) have one degree-drop linear hyperplane. There were 494 classes having degree-drop spaces of co-dimension 2 but not 1, and the remaining 493 classes (that is, about half of all the classes) have degree-drop spaces of co-dimension 3 but not 2, i.e. they have optimal restriction degree stability . Therefore $\deg\_stab(4, 8) = 2$.

References

1. C. Carlet. Boolean Functions for Cryptography and Error Correcting Codes. Monography in *Cambridge University Press*, 2021.
2. C. Carlet and S. Feukoua. Three basic questions on Boolean functions. *Advances in Mathematics of Communications*, Vol 4, No 11, pp. 837-855, 2017.

3. C. Carlet and S. Feukoua, Three parameters of Boolean functions related to their constancy on affine spaces. *Advances in Mathematics of Communications*, Vol 14, No. 4, pp. 651-676, 2020.
4. N. Courtois. Fast Algebraic Attacks on Stream Ciphers with Linear Feedback. *Proceedings of CRYPTO 2003, Lecture Notes in Computer Science 2729,* pp. 177- 194, 2003.
5. N. Courtois and W. Meier. Algebraic Attacks on Stream Ciphers with Linear Feedback. *Proceedings of EUROCRYPT 2003, Lecture Notes in Computer Science 2656*, pp. 346-359, 2003
6. D. K. Dalai, S. Maitra and S. Sarkar. Basic Theory in Construction of Boolean Functions with Maximum possible Annihilator Immunity. *Designs, Codes and Cryptography* 40(1), pp. 41-58, 2006.
7. X. Hou, $GL(m,2)$ Acting on $R(r,m)/R(r-1,m)$, *Discrete Mathematics* 149, pp. 99–122, 1996.
8. P. Langevin and G. Leander. Classification of the quartic forms of eight variables. *In Boolean Functions in Cryptology and Information Security, Svenigorod*, Russia, http://langevin.univ-tln.fr/project/quartics/quartics.html, 2007.
9. F. J. MacWilliams and N. J. Sloane. The theory of error-correcting codes, Amsterdam, North Holland. 1977.
10. M. Matsui. Linear cryptanalysis method for DES cipher. *Proceedings of EUROCRYPT 1993, Lecture Notes in Computer Science 765*, pp. 386-397, 1994.

Appendix

Here we denote the 31 non-zero representatives, of the 31 classes of polynomials of degree 3 in 8 variables in [7], by $f_2, \ldots, f_{32}$ (with f_i being the same as the function denoted by F_i in [7]) and defined as follows, where 123 means $x_1x_2x_3$:

$$f_2 = 123$$
$$f_3 = 123 + 145$$
$$f_4 = 123 + 456$$
$$f_5 = 123 + 245 + 346$$
$$f_6 = 123 + 145 + 246 + 356 + 456$$
$$f_7 = 127 + 347 + 567$$
$$f_8 = 123 + 456 + 147$$
$$f_9 = 123 + 245 + 346 + 147$$
$$f_{10} = 123 + 456 + 147 + 257$$
$$f_{11} = 123 + 145 + 246 + 356 + 456 + 167$$
$$f_{12} = 123 + 145 + 246 + 356 + 456 + 167 + 247$$
$$f_{13} = 123 + 456 + 178;$$
$$f_{14} = 123 + 456 + 178 + 478;$$
$$f_{15} = 123 + 245 + 678 + 147;$$
$$f_{16} = 123 + 245 + 346 + 378;$$

$f_{17} = 123 + 145 + 246 + 356 + 456 + 178;$
$f_{18} = 123 + 145 + 246 + 356 + 456 + 167 + 238;$
$f_{19} = 123 + 145 + 246 + 356 + 456 + 158 + 237 + 678;$
$f_{20} = 123 + 145 + 246 + 356 + 456 + 278 + 347 + 168;$
$f_{21} = 145 + 246 + 356 + 456 + 278 + 347 + 168 + 237 + 147;$
$f_{22} = 123 + 234 + 345 + 456 + 567 + 678 + 128 + 238 + 348 + 458 + 568 + 178;$
$f_{23} = 123 + 145 + 246 + 356 + 456 + 167 + 578;$
$f_{24} = 123 + 145 + 246 + 356 + 456 + 167 + 568;$
$f_{25} = 123 + 145 + 246 + 356 + 456 + 167 + 348;$
$f_{26} = 123 + 456 + 147 + 257 + 268 + 278 + 348;$
$f_{27} = 123 + 456 + 147 + 257 + 168 + 178 + 248 + 358;$
$f_{28} = 127 + 347 + 567 + 258 + 368;$
$f_{29} = 123 + 456 + 147 + 368;$
$f_{30} = 123 + 456 + 147 + 368 + 578;$
$f_{31} = 123 + 456 + 147 + 368 + 478 + 568;$
$f_{32} = 123 + 456 + 147 + 168 + 258 + 348.$

A class of locally recoverable codes over finite chain rings

Giulia Cavicchioni[1], Eleonora Guerrini[2], and Alessio Meneghetti[1]

[1] University of Trento giulia.cavicchioni@unit.it, alessio.meneghetti@unitn.it
[2] LIRMM, Université de Montpellier eleonora.guerrini@lirmm.fr

Keywords: Ring-linear code, Locally recoverable codes, Erasure correction

Abstract. Locally recoverable codes deal with the task of reconstructing a lost symbol by relying on a portion of the remaining coordinates smaller than an information set. We consider the case of codes over finite chain rings, generalizing known results and bounds for codes over fields. In particular, we propose a new family of locally recoverable codes by extending a construction proposed in 2014 by Tamo and Barg, and we discuss its optimality.

1 Introduction

Introduced in [6], locally recoverable codes have garnered attention due to their relevance in distributed and cloud storage systems. Data centers and other modern distributed storage systems use redundant data storage to protect against node failures. Indeed they enable local repair of a coordinate by accessing a maximum of r other coordinates. This set of r coordinates is commonly referred to as the *recovering set* and, if a recovering set exists for every coordinate the code has locality r. Many research efforts [2, 6, 7, 9, 10, 23, 25] have been focused on establishing bounds for the minimum distance and developing construction techniques for locally recoverable codes.
If C is a linear code of length n, dimension k and locality r over the field $\mathbb{F}_q$, then its minimum distance satisfies [6]

$$d \leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2 \ . \tag{1.1}$$

In [6], using a probabilistic argument, the authors proved that the bound (1.1) is tight if the field is large enough. Observe that (1.1) is independent of the alphabet size q. In [4] a bound for the minimum distance of a locally recoverable code depending on q is presented.

A central problem in Coding Theory is to construct optimal codes. We remark that usually a code C is said to be optimal if any other code with equal length and minimum distance has at most the same number of codewords as C.

In this work we follow instead the route established for example in [6], and we say that a length-n code C with M codewords is optimal if no code over the same alphabet and with the same parameters has a strictly larger minimum distance. A code meeting the bound (1.1) is thus an optimal locally recoverable code. Constructions of optimal locally recoverable codes are given in [2, 7, 9, 10, 23, 25]. In all these constructions, the i-th coordinate together with its recovering set form a 1-erasure correcting code. A possible extension is presented in [19], where the authors introduced the (r, ρ)-locality, allowing recovering $\rho - 1$ erasures by looking at other r coordinates. An additional and relevant generalization can be found in [21], where each coordinate has several pairwise disjoint recovering sets.

In this paper, we present a generalization of the theory, allowing the alphabet to be a ring [16, 17, 22, 24], rather than a field as in classical Coding Theory.

2 Preliminaries: codes over rings and locality

Let R be a finite commutative ring, with $q = |R|$. From the structure theorem [14, Theorem VI.2], R decomposes uniquely (up to the order of summands) as a finite direct product of local rings, thus $R = R_1 \times \cdots \times R_w$ and $R^n = R_1^n \times \cdots \times R_w^n$. From now on, let R be a principal ideal ring (PIR), so that the R_i are finite chain rings. An *R-linear code* of length n is an R-submodule $C \subseteq R^n$. An R-linear code C is said to be *free* if C is a free submodule of R^n. The elements of C are called *codewords*. Unless otherwise specified, from now on we consider any code to be an R-linear code.

Remark 1. Given $R = R_1 \times \cdots \times R_w$, we define e_i as the element in R represented by $(0, \ldots, 0, 1, 0, \ldots, 0)$ in $R_1 \times \cdots \times R_w$, with 1 in the i-th position. Let $\pi_i \colon R_1^n \times \cdots \times R_w^n \to R_i^n$ be the i-th canonical projection. If C is a code and $c = (c_1, \ldots, c_w) \in C$, where $c_i = \pi_i(c) \in R_i^n$, then

$$e_i c = (0, \ldots, 0, c_i, 0, \ldots, 0) \in C \ .$$

Hence, up to isomorphism, C can be uniquely written as

$$C_1 \times \cdots \times C_w \subseteq R^n \ \text{ with } \ C_i = \pi_i(C) \ \text{ for all } \ 1 \leq i \leq w \ .$$

Therefore, whenever convenient, we may restrict our focus on codes over finite chain rings. In the classical framework R is a finite field, and in this case an important parameter of a linear code is its dimension as a vector subspace of R^n. In our context, if R is a finite chain ring with p^s elements, we define the *p^s-dimension* k of the code as $k = \log_{p^s} |C|$ [3]. Observe that if R is a finite field then the p^s-dimension and the dimension coincide.

Definition 2. The *rank* of C is the minimum K such that there exists a monomorphism $\phi \colon C \to R^K$ as R-modules. In addition, if ϕ is an isomorphism then C is free and $k = K$.

A *minimal generating set* of a code $C \subseteq R^n$ is a subset of C that generates C as an R-module and it is minimal with respect to inclusion. If R is a finite chain ring, as a consequence of Nakayama's Lemma [14, Theorem V.5], all the minimal generating sets have the same cardinality, equal to the rank K of C. A matrix whose rows form a generating set for the code is a *generator matrix*.

The Hamming metric is a discrete metric counting the number of entries in which two tuples differ, namely, for any $v = (v_1, \ldots, v_n)$ and $u = (u_1, \ldots, u_n)$ in R^n, the distance $\mathrm{d}(v, u) = |\{i : v_i \neq u_i,\ 1 \leq i \leq n\}|$. The so-called *minimum distance* $d = \min_{c_1, c_2 \in C,\, c_1 \neq c_2} \mathrm{d}(c_1, c_2)$ is a relevant parameter related to the error correction capability of C, which is the number of coordinates of $c \in C$ that can be corrupted without compromising our ability of reconstructing c.

Theorem 3 (Singleton bound [13])**.** If C is a (non-necessarily R-linear) code of length n over an alphabet of size q, then

$$d \leq n - \log_q |C| + 1 \ .$$

If R is a finite chain ring with $|R| = p^s$ and C is a linear code of p^s-dimension k in R^n, the previous bound reads

$$d \leq n - k + 1 \ .$$

Only free codes can meet this bound and they are said *maximum distance separable* (MDS) codes. However, in the framework of codes over finite chain rings, the Singleton bound can be improved.

Theorem 4 (Generalized Singleton bound [16])**.** Let R be a finite chain ring and let C be an R-linear code of length n and rank K. Then

$$d \leq n - K + 1 \ .$$

This bound is generally tighter than the Singleton bound, and they coincide if and only if the code is free. A linear code meeting this bound is said to be *maximum distance with respect to rank* (MDR).

For any linear code $C \subseteq R^n$ and for any subset $S \subset \{1, \ldots, n\}$ of the coordinates, we define C_S to be the *punctured code of C in S*, obtained by deleting in each codeword all but the coordinates indexed in S. If $|C| = |C_S|$ then S is an *information set of size* $|S|$ for C. In the following, we will denote with κ the minimal size of an information set. Note that for codes over finite chain rings the minimum size of an information set coincides with the rank and $\kappa = K$.

Corollary 5. Let C be a code with minimum distance d and let S be a subset of coordinates which does not form an information set. Then

$$|S| \leq n - d \ .$$

As briefly stated in the introduction, the goal of a local recovery technique is to enable the retrieval of lost encoded data by using only a small portion of the available information, rather than requiring access to the complete codeword c.

Definition 6. Let C be a (possibly non-linear) code in R^n and let $(c_1, \ldots, c_n)$ be a codeword. We say that the coordinate $i \in \{1, \ldots, n\}$ has *locality* r if there exists a subset $S_i \subseteq \{1, \ldots, n\} \setminus \{i\}$ such that:

- *(locality)* $|S_i| \leq r$,
- *(recovery)* $|C_S| = |C_{S \cup \{i\}}|$.

C is a *locally recoverable code* (LRC) with *locality* r if each coordinate has locality r.

In other words, any symbol c_i of any codeword c can be recovered by accessing at most r other symbols of c. If we are presented with a codeword c that is error-free except for an erasure at position i, we can retrieve the original codeword by only examining the coordinates in S_i. For this reason, S_i is referred to as a *recovering set* for i. If R is a finite chain ring and C is an R-linear code of length n, rank K and locality r, we will say that C is an (n, K, r)-code.
Of course, one can choose $R = \mathbb{F}_q$. In this case we recover the classical theory of locally recoverable codes over finite fields.

In 2014 Tamo and Barg [23] presented a clever construction based on polynomial interpolation for locally recoverable codes over finite fields attaining the bound (1.1). In the following sections we extend this construction to finite chain rings. For the sake of readability, in the following sections we omit the proofs of our results, which can be found in the preprint version of this work [5].

3 Lower bound on the minimum distance

Let R be a commutative ring, let C be a code of length n over R and let κ be the minimum size of its information sets.

Theorem 7. Let C be a code of length n and locality r over R. Then

$$d \leq n - \kappa - \left\lceil \frac{\kappa}{r} \right\rceil + 2 \quad \text{and} \quad \frac{\kappa}{n} \leq \frac{r}{r+1} \, . \tag{3.1}$$

Corollary 8 (LRC bound for R-linear codes)**.** Let R be a finite chain ring and let C be an R-linear code of length n, rank K and locality r. Then

$$d \leq n - K - \left\lceil \frac{K}{r} \right\rceil + 2 \, . \tag{3.2}$$

Note that each coordinate in an R-linear code of rank K has locality at most K. Thus r satisfies $1 \leq r \leq K$. In particular:

- If $r = K$, the LRC bound reduces to the generalized Singleton bound and optimal LRC codes are MDR codes;

– If $r = 1$, bound (3.2) reads $d \leq 2\left(\frac{n}{2} - K + 1\right)$. Therefore, by replicating each symbol twice in an MDR code of length $\frac{n}{2}$ and rank K, we get an optimal linear code with locality $r = 1$.

Codes that attain the LRC bound for finite chain rings can be used as building blocks to construct codes that achieve the LRC bound on finite PIRs and hence, we can focus our studies on LRC codes over finite chain rings.

Theorem 9. Let $R = R_1 \times \cdots \times R_w$ be a finite PIR and let $C = C_1 \times \cdots \times C_w \subseteq R^n$ be an R-linear code. If C_i is an optimal LRC over R_i for all $1 \leq i \leq w$, then C is optimal LRC over R.

4 Extending the Tamo-Barg construction

The construction by Tamo and Barg [23] allows to obtain optimal LRC codes over finite fields using a particular class of polynomials called *good polynomials*. Polynomial interpolation is used in order to recover erased data. In order to present our construction, it is important to recall that polynomials over rings lack some desirable properties of polynomials over fields. For instance, in this framework, polynomial interpolation problems require a greater attention.

If R is a finite chain ring with maximal ideal M and residue field $F = R/M$, we will denote by $\bar{y}$ the image of $y \in R$ under the canonical projection from R to F. In addition, for a set $T \subseteq R$ we define $\overline{T} = \{\bar{t} \mid t \in T\}$. Let $N(R)$ denote the group of units of R.

Definition 10. [18, Definition 2.2] A subset $T \subseteq N(R)$ is said to be *subtractive in* $N(R)$ if, for all distinct $a, b \in T$, $a - b \in N(R)$.

Lemma 11. [18, Lemma 2.5] Given $z, s \in R$, then $\bar{z} \neq \bar{s}$ if and only if $z - s \in N(R)$. Thus, T is a subtractive subset of R if and only if $|T| = |\overline{T}|$.

Definition 12. [1, Section III.B] A set $\{a_1, \ldots, a_n\}$ is *well-conditioned* in R if one of the following conditions is satisfied:

1. $\{a_1, \ldots, a_n\}$ is subtractive in $N(R)$;
2. For some i, $\{a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n\}$ is subtractive in $N(R)$ and a_i is a zero-divisor or $a_i = 0$.

The above Definitions and Lemma are useful to work with polynomial reconstruction, as stated in the following known results.

Proposition 13. [20, Corollary 9] Let $f \in R[x]$ be a polynomial of degree at most $n - 1$ with at least n roots in a well-conditioned set of R. Then $f = 0$.

Corollary 14. [20, Corollary 10] Let $\{a_1, \ldots, a_n\}$ be a well-conditioned set in R and let $\{y_1, \ldots, y_n\}$ be a subset of R. Then there exists a unique polynomial $f \in R[x]$ of degree at most $n - 1$ such that $f(a_i) = y_i$ for all $1 \leq i \leq n$.

Proposition 13 points out that, unlike polynomials over fields, the number of roots of a polynomial over a ring is not bounded by its degree. Nonetheless, for polynomials over local rings, there exists a bound on the number of roots, which depends on the polynomial's degree. The following Corollary is a consequence of the Hensel lifting [14, Chapter XIII, Section (C)].

Corollary 15. Let R be a finite chain ring whose residue field F has size $|F| = p^m$ and $|R| = p^{sm}$. Let $f(x) \in R[x]$ be a polynomial of degree n. The number of roots of f in R is at most $np^{(s-1)m}$.

Let $f \in R[x]$. If f is constant on the set A we will denote by $f(A)$ the value of f on A. Additionally, if the leading coefficient of f is a unit, we will call f a *monic polynomial*.

Definition 16. Let $l \in \mathbb{N}^+$ and $A_1, \ldots, A_l$ pairwise disjoint subsets of R of size $r+1$. A polynomial $g \in R[x]$ such that

- Its degree is $r+1$;
- It is monic;
- It is constant on A_i, i.e., for any $1 \leq i \leq l,\ g(A_i) = c_i$ with $c_i \in R$;

is said to be (r,l)-*good* on the blocks $A_1, \ldots, A_l$.

Theorem 17. Let $r \geq 1$ and let $A_1, \ldots, A_l$ be subsets of R such that $A = \bigcup_{i=1}^{l} A_i$ is well-conditioned. Let $g(x) \in R[x]$ be an (r,l)-good polynomial on the blocks of the partition of A. For $t \leq l$, set $n = (r+1)l$ and $K = rt$. Let $a = (a_{i,j},\ 0 \leq i \leq r-1,\ 0 \leq j \leq t-1) \in R^K$. Define the *encoding polynomial* $f_a(x)$ and the code $\mathcal{C}$ as

$$f_a(x) = \sum_{i=0}^{r-1} \sum_{j=0}^{t-1} a_{i,j} g(x)^j x^i \ , \qquad \mathcal{C} = \left\{ (f_a(\alpha),\ \alpha \in A) \mid a \in R^K \right\} . \tag{4.1}$$

Then $\mathcal{C}$ is a free (n, K, r)-code with minimum distance $d = n - K - \frac{K}{r} + 2$. Hence $\mathcal{C}$ is an optimal locally recoverable code.

Let $a \in R^K$ be the message vector and assume $(f_a(\gamma),\ \gamma \in A)$ is sent. Suppose that the symbol c_α corresponding to the location $\alpha \in A_j$ is erased and let c_β for all $\beta \in A_j \setminus \{\alpha\}$ represent the remaining r symbols in the locations of the set A_j. Since g is an (r,l)-good polynomial on the blocks of the partition of A, $f_a(x)$ is a polynomial of degree at most $r-1$ when restricted to A_j. Hence, in order to find c_α, we find the unique polynomial $\delta(x)$ of degree strictly less than r such that $\delta(\beta) = c_\beta$ for all $\beta \in A_j \setminus \{\alpha\}$ and we set $c_\alpha = \delta(\alpha)$. The polynomial $\delta(x)$ is called the *decoding polynomial* for c_α.

Remark 18.
1. If $r = K$ the construction does not require good polynomials and reduces to Reed-Solomon codes.
2. Analogously to the classical case [23, Section 3A], the construction can be generalized even for the case $r \nmid K$.

5 Construction of good polynomials over Galois ring

It is known that classes of good polynomials over finite fields exist [12, 15, 23]. In particular, Micheli [15] introduced a framework that allows the generation of good polynomials over finite fields. The natural question that arises now is whether there exist good polynomials over rings which are not fields. Indeed, they do exist. Here we construct a class of good polynomials over Galois rings exploiting the structure of its group of units.
Let R be a finite chain ring, $M \subset R$ be the maximal ideal, and let $g \in R$. In accordance with the notation of Section 4, we denote with $\bar{g} \in R/M =: F$ its canonical projection onto F, and we extend this projection to $R[x]$, i.e. if $f \in R[x]$ then $\bar{f} \in F[x]$.
Let p be a prime, and let s, m positive integers. The *Galois ring* $\mathrm{GR}(p^s, m)$ of characteristic p^s and with p^{sm} elements is the quotient ring

$$\mathrm{GR}(p^s, m) \cong \mathbb{Z}_{p^s}[x]/(f) \ ,$$

with $f \in \mathbb{Z}_{p^s}[x]$ a monic irreducible polynomial of degree m such that $\bar{f}$ is irreducible in $\mathbb{Z}_p$. A Galois ring $\mathrm{GR}(p^s, m)$ is a local ring with maximal ideal $M = (p)$ and whose residue field $F = \mathrm{GR}(p^s, m)/M$ is isomorphic to the finite field $\mathbb{F}_{p^m}$. Its group of units has a unique maximal cyclic subgroup having order relatively prime to p (namely $p^m - 1$). Throughout this section let $R = \mathrm{GR}(p^s, m)$.

Lemma 19. [14, Lemma XV.1] Let $f \in R[x]$ be a polynomial which is not a zero divisor. Suppose $\bar{f}$ has a zero $l \in F$. Then f has one and only one zero g such that $\bar{g} = l$.

Proposition 20. [18, Theorem 2.9] Let $l \in F$ be an element of order $j \mid p^m - 1$ in F. Then there exists a unique $g \in R$ such that $g^j = 1$ and $\bar{g} = l$.

Corollary 21. Let $q = p^m - 1$ and let $g \in R$ be a primitive qth root of unity. Then $g^i - g^j$ is a unit for all $0 \leq j < i \leq q - 1$.

Let G be the cyclic subgroup of $N(R)$ whose elements are the roots of the polynomial $x^{p^m-1} - 1 \in R[x]$. Corollary 21 implies that G is a subtractive subset in $N(R)$. Lemma 11 implies that the size of any subtractive subset of $N(R)$ cannot exceed $p^m - 1$. A subtractive subset of $N(R)$ of size $p^m - 1$ is said to be a *maximal subtractive subset*. Thus, G is a maximal subtractive subset.

Proposition 22. Let H be a subgroup of the cyclic group G. The annihilator polynomial of H, $h(x) = \prod_{g \in H}(x - g) = x^{|H|} - 1$, is constant on the cosets of H. The same holds true for the polynomial $h(x) = x^{|H|}$.

The annihilators of subgroups form a class of $\big(|H| - 1, (p^m - 1)/|H|\big)$-good polynomials that can be employed in constructing optimal codes. If $|H| = r + 1$, $r + 1$ divides $|G|$ and $p^m \equiv 1 \mod r + 1$. Thus, the length of the code is always a multiple of $r + 1$. It is worth highlighting that the sizes of the possible subgroups and maximum size of a subtractive subset impose constraints on the parameters of the code.

6 Removing the constraints on code length

6.1 Codes over well-conditioned sets with arbitrary length

If n is the code length and r is the locality, Theorem 17 requires the assumption that $r+1$ divides n. However, we provide a different construction that relaxes this condition.

Let R be a finite chain ring, A be a well-conditioned set, $|A| = n$ with $n \bmod (r+1) = m \neq 0, 1$, and let $h_A(x) = \prod_{a \in A}(x-a)$ be the annihilator polynomial of the set A. Let r, K be positive integers and assume $r \mid K+1$. Let $l = \lceil \frac{n}{r+1} \rceil$ and let $A = \bigcup_{i=1}^{l} A_i$ be a partition of A in l subsets such that $|A_i| = r+1$ for all $\leq i \leq m-1$ and $|A_l| = m < r+1$. Let

$$\mathcal{F}_A = \{f \in R[x] \mid f(A_i) = c_i \forall\ i \in \{1, \ldots, l\}\ , \deg f < |A|\}$$

be the algebra of polynomials over R of degree less than $|A|$ which are constant on the blocks of the partition of A.

Theorem 23. Let $A = \bigcup_{i=1}^{l} A_i$ be a well-conditioned set, with $|A_l| = m < r+1$. Let $g(x)$ be a polynomial of degree $r+1$ vanishing on A_l and whose powers span $\mathcal{F}_A$. Let $a = (a_0, \ldots, a_{m-1}) \in R^K$ be the message vector with $a_i \in R^{\frac{K+1}{r}}$ for $i \neq m-1$ and $a_{m-1} \in R^{\frac{K+1}{r}-1}$. Define the encoding polynomial $f_a(x)$ of a as

$$\sum_{i=0}^{m-2} \sum_{j=0}^{\frac{K+1}{r}-1} a_{i,j} g(x)^j x^i + \sum_{j=1}^{\frac{K+1}{r}-1} a_{m-1,j} g(x)^j x^{m-1} + \sum_{i=m}^{r-1} \sum_{j=0}^{\frac{K+1}{r}-1} a_{i,j} g(x)^j h_{A_l}(x)\ .$$

Let

$$\mathcal{C} = \left\{ (f_a(\alpha),\ \alpha \in A) \mid a \in R^K \right\}\ .$$

Then $\mathcal{C}$ is a free (n, K, r)-LRC code with minimum distance $d \geq n - K - \lceil \frac{K}{r} \rceil + 1$.

6.2 LRC codes from arbitrary MDS codes

We present an alternative construction that relaxes the condition $r+1 \mid n$. In the following, we will construct a code such that its symbols can be partitioned into t MDS codes C_i of length n_i and rank K_i.

Definition 24. [11, Definition 2] Let C be a code whose coordinates are partitioned into l sets A_i of size n_i. Let C_i be the code restricted to the coordinates in A_i. The code C has (r, ρ)-*locality* if for all $1 \leq i \leq l$ we have

- $n_i \leq r + \rho - 1$;
- $d_{C_i} \geq \rho$.

Theorem 25. Let R be a finite chain ring and let C be a linear code of length n, rank K and with (r, ρ)-locality. Then

$$d \leq n - K + 1 - \left(\left\lceil \frac{K}{r} \right\rceil - 1 \right)(\rho - 1)\ . \tag{6.1}$$

Theorem 26. Let $r \geq 1$ and let $A = \bigcup_{i=1}^{l} A_i$ be a partition of the well-conditioned set A into l subsets with $|A_i| = r + \rho - 1$ for all $1 \leq i \leq l$. Let $g(x) \in R[x]$ be an $(r + \rho - 1, l)$-good polynomial on the blocks of the partition of A. For $r \mid K$ (this constraint can be lifted, see Remark 18), let $a = (a_0, \ldots, a_{r-1}) \in R^K$ be the message vector with $a_i \in R^{\frac{K}{r}}$ for all $1 \leq i \leq l$. We define

$$f_a(x) = \sum_{i=0}^{r-1} \sum_{j=0}^{\frac{K}{r}-1} a_{i,j} g(x)^j x^i \ , \qquad \mathcal{C} = \left\{ (f_a(\alpha),\ \alpha \in A) \mid a \in R^K \right\} . \tag{6.2}$$

Then, $\mathcal{C}$ is a free code with (r, ρ)-locality and rank K. Moreover C is an optimal (r, ρ)-LRC code.

The previous construction is a particular case of a more general one based on the Chinese Remainder Theorem for rings.

Proposition 27. [14, Section V] Let $h_1(x), \ldots, h_n(x) \in R[x]$ be polynomials that generate pairwise coprime ideals. Then, for any $a_1(x), \ldots, a_n(x) \in R[x]$, there exists a unique polynomial $f \in R[x]$ of degree at most $\sum_i \deg h_i$ such that $f(x) \equiv a_i(x) \bmod h_i(x)$ for all $1 \leq i \leq n$.

Let R be a finite chain ring, let A be a subtractive subset of $N(R)$, and let $A = \bigcup_{i=1}^{l} A_i$ be a partition of A. Using the Hensel Lemma [14], one can prove that the annihilator polynomials of the A_is generate pairwise coprime ideals.

Theorem 28. Let R be a finite chain ring and let A be a subtractive subset of $N(R)$. Let $A = \bigcup_{i=1}^{l} A_i$ be a partition of A such that $|A_i| = n_i$ for all $1 \leq i \leq l$. Let

$$\psi \colon R^K \to \mathcal{F}_{K_1} \times \cdots \times \mathcal{F}_{K_l}, \quad a \mapsto (a_1(x), \ldots, a_l(x)) \ ,$$

be an injective mapping, where $\mathcal{F}_{K_i}$ is the space of polynomials of degree less than K_i. Let $h_i(x)$ be the annihilator polynomial of A_i. For any message vector $a \in R^K$ we define the encoding polynomial $f_a(x)$ as the unique polynomial of degree less than n such that $f_a(x) = a_i(x) \bmod h_i(x)$. The LRC code

$$\mathcal{C} = \left\{ (f_a(\alpha),\ \alpha \in A) \mid a \in R^K \right\}$$

is a free code of rank K. Moreover it can be partitioned into l disjoint local codes $\mathcal{C}_i$, where $\mathcal{C}_i$ is an (n_i, K_i)-MDS code. The minimum distance of $\mathcal{C}$ is at least the minimum between the distances of the local codes $\mathcal{C}_i$.

6.3 LRC codes with non-well-conditioned sets

The most significant limitation in the previous approaches is the restriction on the code length, which is bounded by the maximum size of a well-conditioned set. Let $R = \mathrm{GR}(p^s, m)$ be a Galois ring with residue field $R/M \cong \mathbb{F}_{p^m}$ and let $N(R)$ denote the group of units of R having size $p^{m(s-1)}(p^m - 1)$.

Let G be the maximal cyclic subgroup of $N(R)$ of order coprime with p and let H be a subgroup of G. The cosets $A_1, \ldots, A_l$ of H in $N(R)$ induce a partition of $N(R) = \bigcup_{i=1}^{l} A_i$. Although H is subtractive in $N(R)$, $N(R)$ is not. However, $N(R)$ contains a maximal subtractive subset. Up to reordering, we can assume $\mathcal{A} = \bigcup_{i=1}^{m} A_i$, $m < l$, to be a maximal subtractive subset in $N(R)$.

Theorem 29. Let $r \geq 1$ and let $N(R) = \bigcup_{i=1}^{l} A_i$ be a partition of $N(R)$ into l subtractive subsets A_i of size $r+1$ for all $1 \leq i \leq l$. Let $\mathcal{A} = \bigcup_{i=1}^{m} A_i$, $m < l$, be a maximal subtractive subset of $N(R)$. Let $g(x) \in R[x]$ be an (r, l)-good polynomial on the blocks of the partition of $N(R)$. For $t \leq l$, set $n = (r+1)l$ and $K = rt$. Let $a = (a_{i,j},\ 0 \leq i \leq r-1,\ 0 \leq j \leq t-1) \in R^K$. We define

$$f_a(x) = \sum_{i=0}^{r-1} \sum_{j=0}^{t-1} a_{i,j} g(x)^j x^i \ , \qquad \mathcal{C} = \left\{ (f_a(\alpha),\ \alpha \in N(R)) \mid a \in R^K \right\} . \quad (6.3)$$

Then $\mathcal{C}$ is a free (n, K, r)-code where $n = |N(R)| = p^{m(s-1)}(p^m - 1)$ and minimum distance

$$d = n - p^{m(s-1)} \left(K + \frac{K}{r} - 2 \right) .$$

Note that this construction does not lead to a wider class of good polynomials. Indeed, if h is the annihilator polynomial of the set $\mathcal{A}$, then the class of (r, l)-good polynomials coincides modulo h to the class of (r, m)-good polynomials. Even though we have lifted the constraint on the maximum code length, the code does not meet the LRC bound and thus it is not known whether it is optimal or not.

6.4 On the maximum length of an optimal LRC over finite chain rings

A natural question arises: is there any constraint on the maximum length of a code meeting the LRC bound, as a function of the alphabet size? In the following, we will see that the problem of determining the maximum possible length of an optimal LRC code over R is closely related to the same problem over fields.
Let R be a finite chain ring, let γ be the generator of the maximal ideal and let s be its nilpotency index. Let F be the residue field of R, i.e. $F = R/(\gamma)$. For any $C \subseteq R^n$ we define the code $(C : t) = \{e \in R^n \mid te \in C\}$. In accordance with the notation of Section 4, let $\overline{(C : t)}$ be the projection of $(C : t)$ over F.

Theorem 30. The maximum possible length of an optimal LRC code C over R is bounded by the maximum possible length of the optimal LRC code $\overline{(C : \gamma^{s-1})}$ over F.

While for small code distances ($d = 3, 4$) optimal LRC codes with unbounded length over any fixed alphabet of size $q \geq r+1$ are known, for $d \geq 5$ there is an upper bound on the length of the optimal LRC as a function of its alphabet size. Guruswami et al. in [8] proved that for $d = 5$ the length of an optimal LRC over an alphabet of size q is at most $\mathcal{O}(q^2)$. Moreover, if $d > 5$ the length is at most $\mathcal{O}(q^3)$.

Bibliography

[1] Marc André Armand. List decoding of generalized Reed-Solomon codes over commutative rings. *IEEE transactions on information theory*, 51(1):411–419, 2005.

[2] Alexander Barg, Kathryn Haymaker, Everett W Howe, Gretchen L Matthews, and Anthony Várilly-Alvarado. Locally recoverable codes from algebraic curves and surfaces. In *Algebraic Geometry for Coding Theory and Cryptography: IPAM, Los Angeles, CA, February 2016*, pages 95–127. Springer, 2017.

[3] Eimear Byrne, Anna-Lena Horlemann, Karan Khathuria, and Violetta Weger. Density of free modules over finite chain rings. *Linear Algebra and its Applications*, 651:1–25, 2022.

[4] Viveck R. Cadambe and Arya Mazumdar. Bounds on the size of locally recoverable codes. *IEEE Transactions on Information Theory*, 61(11):5787–5794, 2015.

[5] Giulia Cavicchioni, Eleonora Guerrini, and Alessio Meneghetti. A class of locally recoverable codes over finite chain rings. preprint: `https://arxiv.org/abs/2401.05286`, 2023.

[6] Parikshit Gopalan, Cheng Huang, Huseyin Simitci, and Sergey Yekhanin. On the locality of codeword symbols. *IEEE Transactions on Information theory*, 58(11):6925–6934, 2012.

[7] Sreechakra Goparaju and Robert Calderbank. Binary cyclic codes that are locally repairable. In *2014 IEEE International Symposium on Information Theory*, pages 676–680. IEEE, 2014.

[8] Venkatesan Guruswami, Chaoping Xing, and Chen Yuan. How long can optimal locally repairable codes be? *IEEE Transactions on Information Theory*, 65(6):3662–3670, 2019.

[9] Lingfei Jin. Explicit construction of optimal locally recoverable codes of distance 5 and 6 via binary constant weight codes. *IEEE Transactions on Information Theory*, 65(8):4658–4663, 2019.

[10] Lingfei Jin, Liming Ma, and Chaoping Xing. Construction of optimal locally repairable codes via automorphism groups of rational function fields. *IEEE Transactions on Information Theory*, 66(1):210–221, 2019.

[11] Govinda M Kamath, N Prakash, V Lalitha, and P Vijay Kumar. Codes with local regeneration and erasure correction. *IEEE Transactions on information theory*, 60(8):4637–4660, 2014.

[12] Jian Liu, Sihem Mesnager, and Lusheng Chen. New constructions of optimal locally recoverable codes via good polynomials. *IEEE Transactions on Information Theory*, 64(2):889–899, 2017.

[13] Florence Jessie MacWilliams and Neil James Alexander Sloane. *The theory of error correcting codes*, volume 16. Elsevier, 1977.

[14] Bernard R McDonald. *Finite rings with identity*, volume 28. Marcel Dekker Incorporated, 1974.

[15] Giacomo Micheli. Constructions of locally recoverable codes which are optimal. *IEEE transactions on information theory*, 66(1):167–175, 2019.

[16] G.H. Norton and A. Salagean. On the hamming distance of linear codes over a finite chain ring. *IEEE Transactions on Information Theory*, 46(3):1060–1067, 2000.

[17] Graham H Norton and Ana Sălăgean. On the structure of linear and cyclic codes over a finite chain ring. *Applicable algebra in engineering, communication and computing*, 10:489–506, 2000.

[18] Graham H Norton and Ana Salagean-Mandache. On the key equation over a commutative ring. *Designs, Codes and Cryptography*, 20:125–141, 2000.

[19] N Prakash, Govinda M Kamath, V Lalitha, and P Vijay Kumar. Optimal linear codes with a local-error-correction property. In *2012 IEEE International symposium on information theory proceedings*, pages 2776–2780. IEEE, 2012.

[20] Guillaume Quintin, Morgan Barbier, and Christophe Chabot. On generalized Reed–Solomon codes over commutative and noncommutative rings. *IEEE transactions on information theory*, 59(9):5882–5897, 2013.

[21] Ankit Singh Rawat, Dimitris S. Papailiopoulos, Alexandros G. Dimakis, and Sriram Vishwanath. Locality and availability in distributed storage. *IEEE Transactions on Information Theory*, 62(8):4481–4493, 2016.

[22] Keisuke Shiromoto. Singleton bounds for codes over finite rings. *Journal of Algebraic Combinatorics*, 12:95–99, 2000.

[23] Itzhak Tamo and Alexander Barg. A family of optimal locally recoverable codes. *IEEE Transactions on Information Theory*, 60(8):4661–4676, 2014.

[24] Jay A Wood. Foundations of linear codes defined over finite modules: the extension theorem and the macwilliams identities. In *Codes over rings*, pages 124–190. World Scientific, 2009.

[25] Chaoping Xing and Chen Yuan. Construction of optimal locally recoverable codes and connection with hypergraph. *arXiv preprint arXiv:1811.09142*, 2018.

Spread Code Constructions from Abelian non-cyclic groups

Joan-Josep Climent[0000−0003−0522−0304], Verónica Requena[0000−0002−1497−6456], and Xaro Soler-Escrivà[0000−0001−7595−7032]

Departament de Matemàtiques, Universitat d'Alacant, Spain
jcliment@ua.es, vrequena@ua.es, xaro.soler@ua.es

Abstract. Given the finite field $\mathbb{F}_q$, for a prime power q, in this paper we present a way of constructing spreads of $\mathbb{F}_q^n$. They will arise as orbits under the action of an Abelian non-cyclic group. First, we construct a family of orbit codes of maximum distance using this group, and then we complete each of these codes to achieve a spread of the whole space having an orbital structure.

Keywords: Network coding · subspace codes · Grassmannian · spreads · group action · general linear group.

1 Introduction

Network coding is a part of information theory that describes a method to maximize the rate of a network which is modelled by a directed acyclic multigraph, with one or multiple sources and multiple receivers. First introduced in [1], the key point of this method is allowing the intermediate nodes of the network to transmit linear combinations of the inputs they receive. The algebraic approach given by Kötter and Kschischang in [15] provided a rigorous mathematical setup for error correction when coding in non-coherent networks and, as a result, this theory was able to advance vastly. In this setting, the transmitted messages (*codewords*) are vector subspaces of a given vector space $\mathbb{F}_q^n$, where $\mathbb{F}_q$ is the finite field of q elements and a *subspace code* is just a collection $\mathcal{C}$ of vector subspaces of $\mathbb{F}_q^n$. When all subspaces of $\mathcal{C}$ have the same dimension, we say that $\mathcal{C}$ is a *constant dimension code*. The minimum distance $d(\mathcal{C})$ of $\mathcal{C}$ is computed in the usual way by using a metric called *subspace distance* in the set of all subspaces of $\mathbb{F}_q^n$. We refer the reader to [4, 14] and references therein for further information regarding network coding and subspace codes.

One of the most important and studied families of subspace codes are *spread codes* (or simply *spreads*). A spread code is a constant dimension code such that all its elements intersect pairwise trivially and their union covers the whole vector space. Spreads are clearly a relevant family of constant dimension codes since they reach the maximum distance and, at the same time, the maximum size for that distance. For this reason, many papers in the literature about subspace codes are devoted to the study and construction of this type of codes (see [8, 11, 13, 18] for instance).

A relevant way of constructing constant dimension codes is by considering the natural action of the general linear group, $\mathrm{GL}_n(\mathbb{F}_q)$, on the Grassmannian $\mathcal{G}_q(k,n)$, which is the set of all k-dimensional vector subspaces of $\mathbb{F}_q^n$ (for an integer $k \in \{1,\ldots,n\}$). Using this technique, the codes arise as orbits under the action of some specific subgroup of $\mathrm{GL}_n(\mathbb{F}_q)$. Constant dimension codes constructed in this way are called *orbit codes*. In the linear network coding setting, they were first introduced in [27], where their main properties are given. Due to the group action point of view, orbit codes have a nice mathematical structure, and they have been investigated by many different authors since then (see for instance [3, 9, 19, 22, 25]). Using powerful tools from group theory, the distance of this kind of codes can be calculated in a simpler way and we can compute the size of the code in terms of the order of the acting group and the order of the corresponding stabiliser subgroup [27]. In addition, there exist different algorithms for decoding orbit codes [21, 25] and several of the known algebraic constructions of constant dimension codes can be seen as orbit codes. Most of the research on this topic focus on the use of cyclic subgroups of $\mathrm{GL}_n(\mathbb{F}_q)$, in which case we speak about *cyclic orbit codes*. In particular, in [27] appears the first construction of spreads with an orbital structure. From here, we can find several works on spreads with an orbital structure provided by a cyclic group (for instance, [6, 26]).

While research on cyclic orbit codes abounds, the same is not true when we want to focus on other types of subgroups of the general linear group. As a first step, in [7] we approach the study of orbit codes through the action of Abelian non-cyclic subgroups of $\mathrm{GL}_n(\mathbb{F}_q)$, giving a specific construction of maximum distance. Pursuing this line of research, the papers [5, 24] are also concerned with Abelian non-cyclic orbit codes. Nevertheless, as far as we know, the only construction on spreads through the action of a non-cyclic Abelian group is given in [5]. In this paper, the authors construct an Abelian non-cyclic orbit code of $\mathbb{F}_q^{2k}$ of dimension k having maximum distance and then obtain a k-spread of $\mathbb{F}_q^{2k}$ just by adding two k-subspaces of $\mathbb{F}_q^{2k}$.

Our main objective in this paper is to pursue the research of orbit codes constructed by using non-cyclic Abelian groups. Specifically, we generalize the results obtained in [5] in the following sense: For an even integer n and k a divisor of n, we firstly construct an Abelian non-cyclic orbit code of $\mathbb{F}_q^n$ of dimension k having maximum distance. Then, we achieve to complete this orbit code with a nice family of k-subspaces of $\mathbb{F}_q^n$ in such a way the resulting code is a k-spread of $\mathbb{F}_q^n$ with an orbital structure which is not cyclic. This generalization is not immediate and has required the use of new techniques, not used in [5].

The paper is structured as follows. In Section 2, we collect all the background on finite fields and subspace codes that will be needed in the subsequent sections. Section 3 is devoted to our orbital constructions of maximum distance codes and is divided into two parts. Firstly, we construct a non-cyclic Abelian group $\mathbf{H}$ and from it a maximum distance family of orbit codes of dimension k in $\mathbb{F}_q^n$. Secondly, we find suitable $\mathbf{H}_1$ and $\mathbf{H}_2$ subgroups of $\mathbf{H}$ that allow us to generate new orbit

codes completing each orbit code previously obtained, until a k-spread of $\mathbb{F}_q^n$ is obtained.

2 Preliminaries

2.1 Finite fields

In this section, we recall some basic notions and results concerning finite fields that will be needed later. The reader can find more details in any basic book on this subject, e.g. [17].

Given a finite field $\mathbb{F}_q$ and a positive integer n, we will denote by $\mathbb{F}_q^{n\times n}$ the set of all $n \times n$ matrices with entries in $\mathbb{F}_q$ and by $\mathrm{GL}_n(\mathbb{F}_q)$ the *general linear group* of degree n over $\mathbb{F}_q$.

Assume that $n = ks$, for some positive integers k, s and consider a *primitive element* α of the field $\mathbb{F}_{q^k}$, that is, a generator of the cyclic group $\mathbb{F}_{q^k}^*$. The *companion matrix* of the minimal polynomial $p(\mathrm{x}) = a_0+a_1\mathrm{x}+\cdots+a_{k-1}\mathrm{x}^{k-1}+\mathrm{x}^k$ of α over $\mathbb{F}_q$ is the matrix

$$M_k = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{k-1} \end{pmatrix} \in \mathrm{GL}_k(\mathbb{F}_q).$$

It turns out that $p(\mathrm{x})$ is the characteristic polynomial of M_k and, then, $\mathbb{F}_{q^k} = \mathbb{F}_q[\alpha]$ can be realized as a set of matrices of $\mathbb{F}_q^{k\times k}$ through the following field isomorphism:

$$\begin{array}{rccc} \phi: & \mathbb{F}_q[\alpha] & \longrightarrow & \mathbb{F}_q\left[M_k\right] \\ & \sum_{i=0}^{k-1} a_i\alpha^i & \mapsto & \sum_{i=0}^{k-1} a_i M_k^i. \end{array} \tag{1}$$

Therefore, M_k can be seen as a primitive element of the finite field $\mathbb{F}_q\left[M_k\right]$. Equivalently, the multiplicative order of M_k, denoted by $o(M_k)$, is $q^k - 1$ and $\mathbb{F}_q\left[M_k\right] = \{0_{k\times k}\} \cup \langle M_k\rangle$, where $\langle M_k\rangle = \left\{I_k, M_k, M_k^2, \ldots, M_k^{q^k-2}\right\}$ is the multiplicative group generated by M_k.

On the one hand, the isomorphism ϕ provided in (1) allows to map vector subspaces of $\mathbb{F}_{q^k}^s$ into vector subspaces of $\mathbb{F}_q^n$ (this is the well-known *field reduction* technique, see [16] for instance). Specifically, each line of $\mathbb{F}_{q^k}^s$, will produce a subspace of dimension k of $\mathbb{F}_q^n$ with the following injective map:

$$\begin{array}{rccc} \varphi: & \mathcal{G}_{q^k}(1,s) & \longrightarrow & \mathcal{G}_q(k,n) \\ & \mathrm{rowsp}\left(u_1 \ldots u_s\right) & \mapsto & \mathrm{rowsp}\left(\phi(u_1)\middle|\ldots\middle|\phi(u_s)\right). \end{array} \tag{2}$$

On the other hand, ϕ is also useful to embed $\mathrm{GL}_s(\mathbb{F}_{q^k})$ into $\mathrm{GL}_n(\mathbb{F}_q)$. By using it, we obtain the following group monomorphism:

$$\begin{array}{rccc} \psi: & \mathrm{GL}_s(\mathbb{F}_{q^k}) & \longrightarrow & \mathrm{GL}_n(\mathbb{F}_q) \\ & \begin{pmatrix} a_{11} & \cdots & a_{1s} \\ \vdots & \ddots & \vdots \\ a_{s1} & \cdots & a_{ss} \end{pmatrix} & \mapsto & \left(\begin{array}{c|c|c} \phi(a_{11}) & \cdots & \phi(a_{1s}) \\ \hline \vdots & \ddots & \vdots \\ \hline \phi(a_{s1}) & \cdots & \phi(a_{ss}) \end{array}\right). \end{array} \tag{3}$$

2.2 Subspace codes

In this section, we present the main results of subspace codes that we need to understand this work.

The Grassmannian $\mathcal{G}_q(k,n)$ can be seen as a metric space endowed with the following *subspace distance* (see [15]):

$$d_S(\mathcal{U},\mathcal{V}) = \dim(\mathcal{U}+\mathcal{V}) - \dim(\mathcal{U}\cap\mathcal{V}) = 2(k-\dim(\mathcal{U}\cap\mathcal{V})),$$

for all $\mathcal{U},\mathcal{V}\in\mathcal{G}_q(k,n)$.

If $\mathcal{C}$ is a non-empty subset of $\mathcal{G}_q(k,n)$ then we say that $\mathcal{C}$ is a *constant dimension code* and its *minimum distance* is defined as

$$d_S(\mathcal{C}) = \min\{d_S(\mathcal{U},\mathcal{V}) \mid \mathcal{U},\mathcal{V}\in\mathcal{C},\ \mathcal{U}\neq\mathcal{V}\} \leq \begin{cases} 2k, & \text{if } 2k\leq n, \\ 2(n-k), & \text{if } 2k\geq n. \end{cases} \tag{4}$$

If $|\mathcal{C}|=1$, we put $d_S(\mathcal{C})=0$; in any other case, $d_S(\mathcal{C})>0$. When the upper bound in expression (4) is attained, we say that $\mathcal{C}$ is a *constant dimension code of maximum distance*. Note that $d_S(\mathcal{C})=2k$ only if $2k\leq n$ and all the codewords in $\mathcal{C}$ intersect trivially. This class of codes of maximum distance are known as *partial spread codes*, and they were introduced in [10] as a generalization of the class of *spread codes*, previously studied in [18]. The code $\mathcal{C}$ will be a spread code, if the subspaces in $\mathcal{C}$ pairwise intersect trivially, and they cover the whole space $\mathbb{F}_q^n$ (see [12]). These codes only occur in the case where $k \mid n$ and have cardinality $\frac{q^n-1}{q^k-1}$. In this way, it can be said that spread codes are partial spreads of maximum size, since the size of a partial spread code of dimension k (or k-partial spread code) is always upper bounded by

$$\frac{q^n-q^m}{q^k-1}, \tag{5}$$

where m is the reminder obtained dividing n by k (see [10]). Notice that any code of lines $\mathcal{C}\subseteq\mathcal{G}_q(1,n)$ with $|\mathcal{C}|\geq 2$ is in particular a partial spread code with size $|\mathcal{C}|\leq\frac{q^n-1}{q-1}$, whereas $\mathcal{G}_q(1,n)$ can be seen as the spread of lines of $\mathbb{F}_q^n$.

In case that we have $\mathcal{C}\subseteq\mathcal{G}_q(k,n)$ with $2k\geq n$, then we can consider the *dual code* of $\mathcal{C}$, that is, the set $\mathcal{C}^\perp=\{\mathcal{V}^\perp \mid \mathcal{V}\in\mathcal{C}\}$, which is a constant dimension code of dimension $n-k$ with the same cardinality and distance as $\mathcal{C}$ (see [15]). In particular, if $d_S(\mathcal{C})=2(n-k)$, then $\mathcal{C}^\perp$ is an $(n-k)$-partial spread code and the size of $\mathcal{C}$ can be also upper bounded in terms of (5). This is the reason why from now on we consider that $2k\leq n$.

An important class of constant dimension codes are those called *orbit codes*, introduced in [27]. These codes are defined as orbits under the action of some subgroup of the general linear group. Consider $V \in \mathbb{F}_q^{k\times n}$ a full-rank matrix generating a subspace $\mathcal{V} = \text{rowsp}(V) \in \mathcal{G}_q(k,n)$. The map

$$\begin{array}{ccc} \mathcal{G}_q(k,n) \times \text{GL}_n(\mathbb{F}_q) & \longrightarrow & \mathcal{G}_q(k,n) \\ (\mathcal{V}, A) & \mapsto & \mathcal{V} \cdot A = \text{rowsp}(VA), \end{array}$$

defines a group action from the right on $\mathcal{G}_q(k,n)$ (see [27]) since it is independent of the choice of the generating matrix V. Given a subgroup $\mathbf{H}$ of $\text{GL}_n(\mathbb{F}_q)$, the *orbit code* $\text{Orb}_{\mathbf{H}}(\mathcal{V})$ is the orbit generated by the action of $\mathbf{H}$ on $\mathcal{V}$, that is,

$$\text{Orb}_{\mathbf{H}}(\mathcal{V}) = \{\mathcal{V} \cdot A \mid A \in \mathbf{H}\} \subseteq \mathcal{G}_q(k,n).$$

The size of this orbit code is $|\text{Orb}_{\mathbf{H}}(\mathcal{V})| = \frac{|\mathbf{H}|}{|\text{Stab}_{\mathbf{H}}(\mathcal{V})|}$, where $\text{Stab}_{\mathbf{H}}(\mathcal{V}) = \{A \in \mathbf{H} \mid \mathcal{V} \cdot A = \mathcal{V}\}$ is the stabilizer subgroup of the subspace $\mathcal{V}$ under the action of $\mathbf{H}$. If $\mathbf{H} = \text{Stab}_{\mathbf{H}}(\mathcal{V})$, then $\text{Orb}_{\mathbf{H}}(\mathcal{V}) = \{\mathcal{V}\}$ and $d_S(\text{Orb}_{\mathbf{H}}(\mathcal{V})) = 0$. In any other case, the minimum distance of the orbit code $\text{Orb}_{\mathbf{H}}(\mathcal{V})$ can be calculated as (see [27]):

$$d_S(\text{Orb}_{\mathbf{H}}(\mathcal{V})) = \min\{d_S(\mathcal{V}, \mathcal{V} \cdot A) \mid A \in \mathbf{H} \setminus \text{Stab}_{\mathbf{H}}(\mathcal{V})\}.$$

Recall that we are assuming $n = ks$ and consider the field reduction map φ defined in (2) which maps lines of $\mathbb{F}_{q^k}^s$ into vector subspaces of dimension k of $\mathbb{F}_q^n$.

Since φ is injective, it preserves intersections and, therefore, it follows that $d_S(\varphi(\mathcal{C})) = k d_S(\mathcal{C}) = 2k$, for any $\mathcal{C} \subseteq \mathcal{G}_{q^k}(1,s)$. In other words, $\varphi(\mathcal{C})$ is a k-partial spread of $\mathbb{F}_q^n$, for any code of lines $\mathcal{C}$ of $\mathbb{F}_{q^k}^s$. In particular, if we consider the spread of all lines of $\mathbb{F}_{q^k}^s$, it turns out that

$$\varphi(\mathcal{G}_{q^k}(1,s)) \subseteq \mathcal{G}_q(k,n), \tag{6}$$

is a k-spread of $\mathbb{F}_q^n$, which is called the *Desarguesian* k-spread of $\mathbb{F}_q^n$ (see [16]). Originally due to Segre (see [23]), in the network coding setting, this construction appears for the first time in [18].

Notice that the field reduction map φ together with the group monomorphism ψ defined in expression (3) allow us to establish a relation between the group action of $\text{GL}_s(\mathbb{F}_{q^k})$ on $\mathcal{G}_{q^k}(1,s)$ and the group action of $\text{GL}_n(\mathbb{F}_q)$ on $\mathcal{G}_q(k,n)$ as follows (see [2, 20]):

$$\varphi(\mathcal{V} \cdot A) = \varphi(\mathcal{V}) \cdot \psi(A), \tag{7}$$

for all $\mathcal{V} \in \mathcal{G}_{q^k}(1,s)$ and $A \in \text{GL}_s(\mathbb{F}_{q^k})$. In particular, when we consider a subgroup $\mathbf{H} \leq \text{GL}_s(\mathbb{F}_{q^k})$ and an orbit code $\text{Orb}_{\mathbf{H}}(\mathcal{V})$, for some $\mathcal{V} \in \mathcal{G}_{q^k}(1,s)$, then

$$\begin{aligned} \varphi(\text{Orb}_{\mathbf{H}}(\mathcal{V})) &= \{\varphi(\mathcal{V} \cdot A) \mid A \in \mathbf{H}\} \\ &= \{\varphi(\mathcal{V}) \cdot \psi(A) \mid \psi(A) \in \psi(\mathbf{H})\} = \text{Orb}_{\psi(\mathbf{H})}(\varphi(\mathcal{V})). \end{aligned} \tag{8}$$

3 Our construction

The main purpose of this section is to show how we can construct the Desarguesian k-spread code of $\mathbb{F}_q^n$ given in (6) starting from the action of a specific Abelian non-cyclic subgroup $\bar{\mathbf{H}}$ of $\mathrm{GL}_n(\mathbb{F}_q)$. To do so, in Subsection 3.1 we will first present our group $\bar{\mathbf{H}}$ and then we will construct a k-partial spread of $\mathbb{F}_q^n$ as an orbit generated by the action of $\bar{\mathbf{H}}$. Following this, in Subsection 3.2 we will explain how to achieve the whole k-spread.

We consider n, k, s, t positive integers such that $n = ks$, with $s = 2t \geq 2$ and $\gcd(t, q^k - 1) = 1$.

3.1 A k-partial spread with an orbital structure

In this subsection we provide an explicit description of a k-partial spread of $\mathbb{F}_q^n$ with an orbital structure. For this purpose, firstly, we will construct an Abelian non-cyclic subgroup $\mathbf{H}$ of $\mathrm{GL}_s(\mathbb{F}_{q^k})$ with a suitable action on certain lines of $\mathbb{F}_{q^k}$.

Let $M_t \in \mathrm{GL}_t(\mathbb{F}_{q^k})$ be the companion matrix of a primitive polynomial of degree t over $\mathbb{F}_{q^k}$. As we introduce in Section 2.1, we have that $\mathbb{F}_{q^{kt}} \cong \mathbb{F}_{q^k}[M_t]$ and, therefore, the order of M_t is $o(M_t) = q^{kt} - 1$. Consider $C = M_t^{q^k - 1}$ whose multiplicative order, clearly, is $r = \frac{q^{kt}-1}{q^k-1}$; and, let $\alpha \in \mathbb{F}_{q^k}$ be a primitive element, whose multiplicative order is $q^k - 1$.

We construct the matrices

$$h_1 = \begin{pmatrix} C & I_t \\ 0_{t\times t} & \alpha I_t \end{pmatrix}, \quad h_2 = \begin{pmatrix} \alpha I_t & -I_t \\ 0_{t\times t} & C \end{pmatrix} \in \mathrm{GL}_s(\mathbb{F}_{q^k}). \tag{9}$$

The following result will be useful in order to compute the multiplicative order of h_1 and h_2.

Lemma 1. *For any positive integer ℓ, one has that*

$$\gcd(\ell, q-1) = 1 \text{ if and only if } \gcd\left(\frac{q^\ell - 1}{q-1}, q-1\right) = 1.$$

Proof. First we notice that if p is a prime dividing $q-1$, then $q \equiv 1 \pmod p$ and so $q^i \equiv 1 \pmod p$ for $i \geq 1$. Therefore,

$$\frac{q^\ell - 1}{q-1} = q^{\ell-1} + q^{\ell-2} + \cdots + q + 1 \equiv \ell \pmod p.$$

Then, it is clear that $p \mid \ell$ if, and only if, $p \mid \frac{q^\ell - 1}{q-1}$ and the result follows.

Note that, since we are assuming that $\gcd(t, q^k - 1) = 1$, Lemma 1 states that $\gcd(r, q^k - 1) = 1$.

The following result is a generalization of Lemmas 3.1, 3.2 and 3.3 of [5]. The proof runs analogously, and thus we omitted it.

Lemma 2. *Consider the matrices h_1, h_2 defined in expression (9). The following statements are satisfied:*

1. *The multiplicative order of h_1 and h_2 is $q^{kt}-1$.*
2. *The matrices h_1 and h_2 commute.*
3. *$\langle h_1\rangle \cap \langle h_2\rangle = \{I_{2t}\}$.*

Notice that we can express the elements of $\mathbf{H} = \langle h_1, h_2\rangle$ as,

$$\mathbf{H} = \langle h_1\rangle\langle h_2\rangle = \left\{h_1^a h_2^b \mid 1 \leq a, b \leq q^{kt}-1\right\},$$

where an arbitrary element can be written as

$$h_1^a h_2^b = \begin{pmatrix} \alpha^b C^a & \sum_{j=1}^{a} \alpha^{j-1} C^{a+b-j} - \sum_{j=1}^{b} \alpha^{j-1} C^{a+b-j} \\ 0_{t\times t} & \alpha^a C^b \end{pmatrix} \tag{10}$$

for any integers $1 \leq a, b \leq q^{kt}-1$. We will denote by $(h_1^a h_2^b)_i$ the i-th row of the matrix $h_1^a h_2^b \in \mathbf{H}$, for $1 \leq i \leq s$.

Let us denote $\boldsymbol{e}_i \in \mathbb{F}_{q^k}^s$ the i-th canonical vector, for $1 \leq i \leq s$. We are interested in the action of the group $\mathbf{H}$ on the lines generated by $\boldsymbol{e}_i$, for $1 \leq i \leq t$. For this reason, we analyse the corresponding stabilizer subgroup. First, we need the following technical lemma.

Lemma 3. *If $1 \leq a, b \leq q^{kt}-1$, then*

$$\sum_{j=1}^{a} \alpha^{j-1} C^{a+b-j} - \sum_{j=1}^{b} \alpha^{j-1} C^{a+b-j} = 0_{t\times t} \text{ if and only if } a = b.$$

As a consequence of this result, we are able to obtain the stabilizer subgroup of $\mathbf{H}$ corresponding to the lines generated by $\boldsymbol{e}_i$, for $i \in \{1, \ldots, t\}$.

Theorem 1. *For all $i \in \{1, \ldots, t\}$, one has that*

$$\mathrm{Stab}_{\mathbf{H}}(\mathrm{rowsp}(\boldsymbol{e}_i)) = \langle (h_1 h_2)^r\rangle = \langle \alpha I_s\rangle.$$

Thus, we construct the following 1-dimensional orbit codes of $\mathbb{F}_{q^k}^s$, for $i \in \{1, \ldots, t\}$:

$$\mathcal{C}_i = \mathrm{Orb}_{\mathbf{H}}(\mathrm{rowsp}(\boldsymbol{e}_i)) = \{\mathrm{rowsp}((h_1^a h_2^b)_i) \mid 1 \leq a, b \leq q^{kt}-1\} \subseteq \mathcal{G}_{q^k}(1, s). \tag{11}$$

For all $i \in \{1, \ldots, t\}$, the size of the orbit code $\mathcal{C}_i$ will be

$$|\mathcal{C}_i| = \frac{|\mathbf{H}|}{|\mathrm{Stab}_{\mathbf{H}}(\mathrm{rowsp}(\boldsymbol{e}_i))|} = \frac{(q^{kt}-1)^2}{q^k-1}.$$

Now, from the lines $\mathrm{rowsp}(\boldsymbol{e}_i) \in \mathbb{F}_{q^k}^s$ and the group $\mathbf{H} \subseteq \mathrm{GL}_s(\mathbb{F}_{q^k})$, the injective map φ defined in (2) and the group monomorphism ψ defined in (3), allow us to construct constant dimension codes of $\mathcal{G}_q(k, n)$. Starting from the

field isomorphism ϕ defined in (1), one has that $\phi(0) = 0_{k\times k}$ and $\phi(1) = I_k$. Thus, we consider the vector subspace

$$\mathcal{U}_{k,i} = \varphi(\text{rowsp}(\boldsymbol{e}_i)) = \text{rowsp}\left(0_{k\times k}|\ldots|I_k|\ldots|0_{k\times k}\right) \subseteq \mathcal{G}_q(k,n).$$

Moreover, we also consider the group $\bar{\mathbf{H}} = \psi(\mathbf{H}) \subseteq \text{GL}_n(\mathbb{F}_q)$. Now, according to expressions (7) and (8), for any $1 \leq i \leq t$, we construct

$$\begin{aligned}\bar{\mathcal{C}}_i = \varphi(\mathcal{C}_i) &= \varphi(\text{Orb}_{\mathbf{H}}(\text{rowsp}(\boldsymbol{e}_i)))\\ &= \text{Orb}_{\psi(\mathbf{H})}(\varphi(\text{rowsp}(\boldsymbol{e}_i))) = \text{Orb}_{\bar{\mathbf{H}}}(\mathcal{U}_{k,i}).\end{aligned} \tag{12}$$

Since φ is an injective map and $d_S(\mathcal{C}_i) = 2$, one deduces that $\bar{\mathcal{C}}_i$ is a k-partial spread of $\mathbb{F}_q^n$, that is, it has dimension k and minimum distance $d_S(\bar{\mathcal{C}}_i) = kd_S(\mathcal{C}_i) = 2k$, for any $1 \leq i \leq t$. Moreover,

$$|\bar{\mathcal{C}}_i| = |\mathcal{C}_i| = \frac{(q^{kt}-1)^2}{q^k-1} = (q^{kt}-1)r.$$

Since a k-spread of $\mathbb{F}_q^n$ has size $\frac{q^n-1}{q^k-1}$ (see Section 2), we can calculate how far each of the $\bar{\mathcal{C}}_i$ codes is from being a k-spread. Specifically, we obtain that

$$|\bar{\mathcal{C}}_i| + 2r = (q^{kt}-1)r + 2r = r(q^{kt}+1) = \frac{q^{kt}-1}{q^k-1}(q^{kt}+1) = \frac{q^n-1}{q^k-1}.$$

3.2 Achieving a k-spread from each $\bar{\mathcal{C}}_i$

In this section, we explain how we can obtain the Desarguesian k-spread of $\mathbb{F}_q^n$ given in (6) starting from each k-partial spread $\bar{\mathcal{C}}_i$ defined in (12), for any $1 \leq i \leq t$. That is, fixing a k-partial spread $\bar{\mathcal{C}}_i$, we explicitly construct $2r$ subspaces of $\mathbb{F}_q^n$ having dimension k and trivial intersection between them and also with the subspaces of $\bar{\mathcal{C}}_i$.

To do so, always starting from our group $\mathbf{H} = \langle h_1 h_2\rangle \leq \text{GL}_s(\mathbb{F}_{q^k})$, for each $i \in \{1,\ldots,t\}$ and $j \in \{t+1,\ldots,s\}$, we construct two sets of lines $\mathcal{A}_i$ and $\mathcal{B}_j$ of $\mathcal{G}_{q^k}(1,s)$, such that $|\mathcal{A}_i| = |\mathcal{B}_j| = r$ and $\mathcal{G}_{q^k}(1,s) = \mathcal{C}_i \cup \mathcal{A}_i \cup \mathcal{B}_j$. Afterward, we will use the field reduction technique to obtain a k-spread of $\mathbb{F}_q^n$.

Recall that the matrices h_1 and h_2 of $\mathbf{H}$ have multiplicative order $q^{kt}-1 = r(q^k-1)$, with $\gcd(r, q^k-1) = 1$. We are going to use the following subgroups of the group $\mathbf{H} \in \text{GL}_s(\mathbb{F}_{q^k})$:

$$\mathbf{H}_2 = \langle h_2^{q^k-1}\rangle, \quad \mathbf{N} = \langle (h_1h_2)^r\rangle \quad \text{and} \quad \mathbf{T} = \langle h_1, h_2^{q^k-1}\rangle = \langle h_1\rangle\mathbf{H}_2.$$

Lemma 4. *Consider the previous groups* $\mathbf{H}$, $\mathbf{H}_2$, $\mathbf{N}$ *and* $\mathbf{T}$. *One has:*

1. $|\mathbf{H}_2| = r$.
2. $\mathbf{N} \cap \mathbf{T} = \{I_s\}$ *and* $\mathbf{H} = \mathbf{NT}$.

Our interest for the subgroup $\mathbf{T}$ is explained in the following result.

Theorem 2. *For each $i \in \{1, \ldots, t\}$ it follows that*

$$\mathcal{C}_i = \mathrm{Orb}_{\mathbf{H}}(\mathrm{rowsp}(\boldsymbol{e}_i)) = \mathrm{Orb}_{\mathbf{T}}(\mathrm{rowsp}(\boldsymbol{e}_i)).$$

From the general expression of elements of $\mathbf{H}$ given in (10), we easily obtain the following expression of elements in $\mathbf{T}$:

$$h_1^a h_2^{(q^k-1)l} = \begin{pmatrix} C^a & D_{a,l} \\ 0_{t\times t} & \alpha^a C^{(q^k-1)l} \end{pmatrix}, \tag{13}$$

where

$$D_{a,l} = \sum_{j=1}^{a} \alpha^{j-1} C^{a+(q^k-1)l-j} - \sum_{j=1}^{(q^k-1)l} \alpha^{j-1} C^{a+(q^k-1)l-j} \in \mathbb{F}_{q^k}[M_t],$$

for any integers $a \in \{1, \ldots, q^{kt}-1\}$ and $l \in \{1, \ldots, r\}$. Since $q^{kt}-1 = (q^k-1)r$, we can put

$$\{1, \ldots, q^{kt}-1\} = A_1 \cup A_2 \cup \cdots \cup A_r, \tag{14}$$

where $A_m = \{a_m r + m \mid 0 \leq a_m \leq q^k - 2\}$, for $m \in \{1, \ldots, r\}$. And notice that these sets form a partition of the set $\{1, \ldots, q^{kt}-1\}$.

Now, according to Theorem 2, for each $i \in \{1, \ldots, t\}$, the orbit code $\mathcal{C}_i$ can be described as the set of lines generated by the i-th row of each matrix of $\mathbf{T}$. We are going to use the partition of $\{1, \ldots, q^{kt}-1\}$ provided in (14) in order to analise these lines.

Lemma 5. *For any integers $a \in \{1, \ldots, q^{kt}-1\}$ and $l \in \{1, \ldots, r\}$, consider the matrices C^a and $D_{a,l}$ given in (13).*

1. *For any $m \in \{1, \ldots, r\}$ it follows that $C^a = C^m$ if and only if $a \in A_m$.*
2. *For each $m \in \{1, \ldots, r\}$, there exists $B_m \in \mathbb{F}_{q^k}[M_t]$ such that $B_m \neq D_{a,l}$, for all $a \in A_m$ and for all $l \in \{1, \ldots, r\}$.*

Next, for each $i \in \{1, \ldots, t\}$, we use the matrices B_m obtained in Lemma 5, in order to define the following sets of lines

$$\mathcal{A}_i = \{\mathrm{rowsp}((C^m | B_m)_i) \mid 1 \leq m \leq r\} \tag{15}$$

Finally, we use the action of the group $\mathbf{H}_2$ on the lines of $\mathbb{F}_{q^k}^s$ generated by the canonical vectors $\boldsymbol{e}_j$, for $j \in \{t+1, \ldots, s\}$ and we consider the orbit codes

$$\mathcal{B}_j = \mathrm{Orb}_{\mathbf{H}_2}(\mathrm{rowsp}(\boldsymbol{e}_j)) \subseteq \mathcal{G}_{q^k}(1, s). \tag{16}$$

We now have all the ingredients to complete each orbit code $\mathcal{C}_i$ until we get the whole space of lines $\mathcal{G}_{q^k}(1, s)$ of $\mathbb{F}_{q^k}^s$.

Theorem 3. *Consider integers $i \in \{1, \ldots, t\}$, $j \in \{t+1, \ldots, s\}$ and the one dimensional codes of $\mathbb{F}_{q^k}^s$ given by $\mathcal{C}_i$, $\mathcal{A}_i$ and $\mathcal{B}_j$ in (11), (15) and (16). One has:*

1. $|\mathcal{A}_i| = |\mathcal{B}_j| = r$.
2. $\mathcal{G}_{q^k}(1, s) = \mathcal{C}_i \cup \mathcal{A}_i \cup \mathcal{B}_j$.

Next, just as we construct the k-partial spreads $\bar{\mathcal{C}}_i$ of $\mathbb{F}_q^n$, we construct now k-partial spreads from the codes $\mathcal{A}_i$ and $\mathcal{B}_j$, for any $i \in \{1, \ldots, t\}$ and $j \in \{t+1, \ldots, s\}$. Denote $\bar{\mathbf{H}}_2 = \psi(\mathbf{H}_2)$. We have that

$$\bar{\mathcal{A}}_i = \varphi(\mathcal{A}_i) = \{\varphi(\mathrm{rowsp}((C^m|B_m)_i)) \mid 1 \leq m \leq r\} \tag{17}$$

and

$$\begin{aligned} \bar{\mathcal{B}}_j &= \varphi(\mathcal{B}_j) = \varphi(\mathrm{Orb}_{\mathbf{H}_2}(\mathrm{rowsp}(\boldsymbol{e}_j))) \\ &= \mathrm{Orb}_{\psi(\mathbf{H}_2)}(\varphi(\mathrm{rowsp}(\boldsymbol{e}_j))) = \mathrm{Orb}_{\bar{\mathbf{H}}_2}(\mathcal{U}_{k,j}). \end{aligned} \tag{18}$$

are k-partial spreads of $\mathbb{F}_q^n$.

Finally, we present our last result, which describes how we can obtain the Desarguesian spread given in (6) from these three partial spreads.

Theorem 4. *For any $i \in \{1, \ldots, t\}$ and $j \in \{t+1, \ldots, s\}$, consider the orbit code $\bar{\mathcal{C}}_i$ defined in* (12)*, the code $\bar{\mathcal{A}}_i$ defined in* (17) *and the orbit code $\bar{\mathcal{B}}_j$ defined in* (18)*. Then, the code $\bar{\mathcal{C}}_i \cup \bar{\mathcal{A}}_i \cup \bar{\mathcal{B}}_j$ is a k-spread of $\mathbb{F}_q^n$.*

4 Conclusions

In this paper, we have dealt with the orbital construction of a k-dimensional spread in $\mathbb{F}_q^n$, where n is an even number and k divides n, using a non-cyclic Abelian group. Our results generalize the results obtained by Chen and Liang in [5] for $\mathbb{F}_q^{2k}$. However, the techniques we have used are new and not easily detached from this work.

Acknowledgments The work of the authors was partially supported by the Spanish I+D+i project PID2022-142159OB-I00 of the Ministerio de Ciencia e Innovación, I+D+i project CIAICO/2022/167 of the Generalitat Valenciana, and the I+D+i project VIGROB-287 of the Universitat d'Alacant.

Bibliography

1. R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network information flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, 2000.
2. C. Alonso-González, M. A. Navarro-Pérez, and X. Soler-Escrivà. An orbital construction of optimum distance flag codes. *Finite Fields and their Applications*, 73:Article 101861, 2021.
3. F. Bardestani and A. Iranmanesh. Cyclic orbit codes with the normalizer of a Singer subgroup. *Journal of Sciences, Islamic Republic of Iran*, 26(1):49–55, 2015.
4. R. Bassoli, H. Marques, J. Rodriguez, K. W. Shum, and R. Tafazolli. Network coding theory: A survey. *IEEE Communications Surveys & Tutorials*, 15(4):1950–1978, 2013.
5. S. Chen and J. Liang. Constructions of spread codes based on Abelian non-cyclic orbit codes. *Linear Algebra and its Applications*, 608:54–67, 2021.
6. S.-d. Chen and J.-y. Liang. New constructions of orbit codes based on the operations of orbit codes. *Acta Mathematicae Applicatae Sinica, English Series*, 36:803–815, 2020.
7. J.-J. Climent, V. Requena, and X. Soler-Escrivà. A construction of Abelian non-cyclic orbit codes. *Cryptography and Communications*, 11(5):839–852, 2019.
8. H. Gluesing-Luerssen and A.-L. Horlemann-Trautmann. Symbol erasure correction in random networks with spread codes. *IEEE Transactions on Information Theory*, 65(4):2075–2091, 2019.
9. H. Gluesing-Luerssen, K. Morrison, and C. Troha. Cyclic orbit codes and stabilizer subfields. *Advances in Mathematics of Communications*, 9(2):177–197, 2015.
10. E. Gorla and A. Ravagnani. Partial spreads in random network coding. *Finite Fields and their Applications*, 26:104–115, 2014.
11. G. Gorla, F. Manganiello, and J. Rosenthal. An algebraic approach for decoding spread codes. *Advances in Mathematics of Communications*, 6(4):443–466, 2012.
12. J. Hirschfeld. *Projective Geometries over Finite Field*. Oxford Mathematical Monographs. Oxford University Press, Oxford, UK, second edition, 1998.
13. T. Honold, M. Kiermaier, and S. Kurz. Partial spreads and vector space partitions. In M. Greferath, M. Osvin Pavăević, N. Silberstein, and M. A. Vázquez-Castro, editors, *Network Coding and Subspace Designs*, Signals and Communication Technology, pages 131–170. Springer International Publishing AG, Cham, Switzerland, 2018.
14. A.-L. Horlemann-Trautmann and J. Rosenthal. Constructions of constant dimension codes. In M. Greferath, M. Osvin Pavăević, N. Silberstein, and M. A. Vázquez-Castro, editors, *Network Coding and Subspace Designs*, Signals and Communication Technology, pages 25–42. Springer International Publishing AG, Cham, Switzerland, 2018.

15. R. Kötter and F. R. Kschischang. Coding for errors and erasures in random network coding. *IEEE Transactions on Information Theory*, 54(8):3579–3591, 2008.
16. M. Lavrauw and G. Van de Voorde. Field reduction and linear sets in finite geometry. *Topics in finite fields*, 632:271–293, 2015.
17. R. Lidl and H. Niederreiter. *Introduction to Finite Fields and Their Applications*. Cambridge University Press, New York, NY, 1986.
18. F. Manganiello, E. Gorla, and J. Rosenthal. Spread codes and spread decoding in network coding. In *Proceedings of the 2008 IEEE International Symposium on Information Theory (ISIT 2008)*, pages 881–885, Toronto, Canada, July 2008. IEEE.
19. F. Manganiello, A.-L. Trautmann, and J. Rosenthal. On conjugacy classes of subgroups of the general linear group and cyclic orbit codes. In *Proceedings of the 2011 IEEE International Symposium on Information Theory (ISIT 2011)*, pages 1916–1920, Saint Pettersburg, July 2011. IEEE.
20. M. A. Navarro-Pérez and X. Soler-Escrivà. Flag codes of maximum distance and constructions using Singer groups. *Finite Fields and their Applications*, 80:Article 102011, 2022.
21. M. H. Poroch and A. A. Talebi. Decoding of orbit codes. *TWMS Journal of Applied and Engineering Mathematics*, 9(2):225–236, 2019.
22. J. Rosenthal and A.-L. Trautmann. A complete characterization of irreducible cyclic orbit codes and their Plücker embedding. *Designs, Codes and Cryptography*, 66:275–289, 2013.
23. B. Segre. Teoria di galois, fibrazioni proiettive e geometrie non desarguesiane. *Annali di Matematica Pura ed Applicata*, 64:1–76, 1964.
24. G. Terra Bastos, R. Palazzo Júnior, and M. Guerreiro. Abelian non-cyclic orbit codes and multishot subspace codes. *Advances in Mathematics of Communications*, 14(4):631–650, 2020.
25. A.-L. Trautmann. Message encoding for spread and orbit codes. In *Proceedings of the 2014 IEEE International Symposium on Information Theory (ISIT 2014)*, pages 2594–2598, Honolulu, Hawaii, June 2014. IEEE.
26. A.-L. Trautmann, F. Manganiello, M. Braun, and J. Rosenthal. Cyclic orbit codes. *IEEE Transactions on Information Theory*, 59(11):7386–7404, 2013.
27. A.-L. Trautmann, F. Manganiello, and J. Rosenthal. Orbit codes – a new concept in the area of network coding. In *Proceedings of the 2010 IEEE Information Theory Workshop (ITW 2010)*, Dublin, Ireland, Aug. 2010. IEEE.

Group Factorisation for Smaller Signatures from Cryptographic Group Actions

Giuseppe D'Alconzo[1], Alessio Meneghetti[2], and Edoardo Signorini[3,1]

[1] Politecnico di Torino, Torino, Italy
[2] University of Trento, Trento, Italy
[3] Telsy, Torino, Italy
giuseppe.dalconzo@polito.it, alessio.meneghetti@unitn.it, edoardo.signorini@telsy.it

Abstract. Cryptographic group actions have gained significant attention in recent years for their application on post-quantum sigma protocols and digital signatures. In NIST's recent additional call for post-quantum signatures, three relevant proposals are based on group actions: LESS, MEDS, and ALTEQ. This work explores signature optimisations leveraging a group's factorisation. We show that if the group admits a factorisation as a semidirect product of subgroups, the group action can be restricted on a quotient space under the equivalence relation induced by the factorisation. If the relation is efficiently decidable, we show that it is possible to construct an equivalent sigma protocol for a relationship that depends only on one of the subgroups. Moreover, if a special class of representative of the quotient space is efficiently computable via a canonical form, the restricted action is effective and does not incur in security loss. Finally, we apply these techniques to LESS and MEDS, showing how they will affect the length of signatures and public keys.

Keywords: Digital signatures · Post-quantum · Code equivalence

1 Introduction

Cryptographic Group Actions. The topic of cryptographic group action has raised a lot of interest in recent years. They represent a generalisation of the Discrete Logarithm Problem, and the underlying problem can be stated as follows: given a group action $(G, X, \star)$ and two elements x, y in X, find, if any, an element g of G such that $y = g \star x$. A first appearance of group actions in cryptography can be found in [4], while in [1] are given the formal assumptions linked to them. This interest has grown since a proposal for a post-quantum Diffie-Hellman is based on the commutative action of the isogenies of elliptic curves CSIDH [6]. After that, many post-quantum proposals have emerged, but the most impactful application is the one related to sigma protocols and digital signatures. For instance, three candidates to the NIST's call for the post-quantum standardisation are based on group actions: LESS [2], MEDS [9] and ALTEQ [16].

Our contribution. The goal of this work is to investigate the cryptographic optimisations taking advantage of a factorisation of the group G. To do this, we introduce a framework that exploits the fact that, to be infeasible to invert, the group action relies only on a part of the group G. More in detail, we show that the group action can be restricted on a quotient space under an appropriate equivalence relation, induced by the group factorisation. From this relation, we propose two optimisation techniques. First, if the relation is decidable in polynomial time, we show that it is possible to define an equivalent sigma protocol for the action $(G, X, \star)$ with shorter responses and without changing the security assumption. Unfortunately, the resulting sigma protocol lacks commitment recoverability, leading to larger signatures. This problem can be overcome with the following technique. We prove that the restricted action can be efficiently computed if an efficiently computable canonical form exists for the equivalence relation. Moreover, we show that this approach can be extended to groups G that are semidirect products of subgroups. We apply these techniques to reduce the size of the public key, secret key and signature of the textbook instantiation of schemes based on code equivalence problems. In particular, we analyse LESS and MEDS. The group acting in the former is $\mathrm{GL}_k(q) \rtimes \mathrm{Mon}(n, q)$, that can be further factorised as $\left(\mathrm{GL}_k(q) \times (\mathbb{F}_q^*)^n\right) \rtimes \mathcal{S}_n$. This, along with the existence of a canonical form for the action of $\mathrm{GL}_k(q) \times (\mathbb{F}_q^*)^n$, implies that the secret can consist of just a permutation of $\mathcal{S}_n$. Moreover, in the sigma protocol, this means that the response of each round is a permutation instead of an element of $\mathrm{GL}_k(q) \rtimes \mathrm{Mon}(n, q)$ or a monomial matrix when the systematic form is involved. Concerning MEDS, we have the action of $\mathrm{GL}_n(q) \times \mathrm{GL}_m(q) \times \mathrm{GL}_k(q)$ on the set of $n \times m$ matrix spaces of dimension k. We consider the factorisation given by $\mathrm{GL}_n(q) \times (\mathrm{GL}_m(q) \times \mathrm{GL}_k(q))$, and, after presenting a canonical form for the action of the group $(\mathrm{GL}_m(q) \times \mathrm{GL}_k(q))$, we describe a compressed variant of the MEDS signature.

Concurrent works. Numerous optimisations for signature schemes based on cryptographic group actions have been proposed. Many of these are generic optimisations that can be applied to any scheme within the framework of Fiat-Shamir signatures. For instance, [12] proposes an approach to reduce the signature size by expanding the public key; while [3] proposes the use of unbalanced challenges when the size of responses varies significantly between distinct challenges. Other optimisations, instead, are closely linked to the specific security assumption. As a reference, LESS includes a variant of the code equivalence introduced in [15] where the size of the signatures is reduced by modifying the commitment generation and the verification procedure. Recently, in [10], the authors introduced a new notion of code equivalence using canonical forms with respect to certain equivalence relations. Compared to our work, the optimisation of [10] cannot be easily extended outside linear equivalence without explicitly proving the reduction from the original assumption. On the other hand, their work exploits a decomposition of the group G that does not require the use of subgroups, leading to increased signature compression.

2 Preliminaries

2.1 Notation

With $\mathcal{S}_n$ and $\mathrm{GL}_n(q)$ we denote the group of permutations acting on n elements and the group of $n \times n$ invertible matrices with coefficients in the finite field with q elements, respectively. $\mathrm{Mon}(n, q)$ is the subgroup of $\mathrm{GL}_n(q)$ of monomial matrices, consisting of matrices with exactly one non-zero element in each row and column. Given a group G, we write $G = G_1 \rtimes G_2$ to denote the internal semidirect product of subgroups G_1, G_2 of G, with G_1 normal in G. If G_2 is normal in G, then $G = G_1 \times G_2$ is an internal direct product of G_1 and G_2.

2.2 Cryptographic Group Actions

We recall the definition of group action and some related properties for their use in cryptography. In the rest of the paper, we will use groups with multiplicative notation.

Definition 1. *Let G be a group, X be a set and $\star$ be a map from $G \times X$ to X. The triple $(G, X, \star)$ is called* group action *if for any g, h in G and x in X, we have $g \star (h \star x) = (gh) \star x$, and, if e is the neutral element of G, then $e \star x = x$ for any x in X.*

In [1] are defined the requirements that a group action must accomplish to be manipulated and used in cryptography. This leads to the definition of *effective group actions.*

Definition 2. *Let λ be a positive integer. Given a group action $(G, X, \star)$ with $\log(|G|) = \mathsf{poly}(\lambda)$ and $\log(|X|) = \mathsf{poly}(\lambda)$, we say that the action is* effective *if the following algorithms are polynomial time computable in λ: unique string representation, sampling and equality testing for both G and X, product and inverse in G and the map $\star$.*

Along with the above polynomial time algorithm, we need some hard problems to use group action in cryptography. The main computational problem related to them is a generalisation of the Discrete Logarithm in the language of group actions.

Definition 3. *Given a group action $(G, X, \star)$, the Group Action Inverse Problem ($GAIP_\star$) takes as input a pair of elements x and y in X and asks to find g in G such that $y = g \star x$, if any.*

Observe that this problem was introduced in [11] with the name of "vectorisation problem" and the related cryptographic assumption is called "one-wayness" of the group action in [1].

2.3 Code Equivalence and related problems

A *k-dimensional linear code* is a subspace of dimension k of a vector space $\mathbb{V}$ endowed with a metric $\mathrm{d} : \mathbb{V} \times \mathbb{V} \to \mathbb{N}$. An *isometry* $\psi : \mathbb{V} \to \mathbb{V}$ for d is a map that does not affect the metric, $\mathrm{d}(\psi(u), \psi(v)) = \mathrm{d}(u, v)$. Two codes are said *equivalent* if there exists an isometry between them and the set of isometries is a group with the group operation given by the composition. This means that a group action on codes can be defined using the group of isometries.
In this work, we will concern linear codes of two types: subspaces of $\mathbb{F}_q^n$ endowed with the *Hamming metric* $\mathrm{d}_H(u, v) = |\{i : v_i - u_i \neq 0\}|$, and subspaces of the vector space of matrices $\mathbb{F}_q^{n \times m}$ endowed with the *rank metric* $\mathrm{d}_{\mathrm{rk}}(U, V) = \mathrm{rank}(V - U)$. Linear codes in the rank metric are also called matrix codes.
We now model the equivalence of codes in the two metrics above as group actions. For the Hamming metric, we have the following.

Definition 4. *Let $G = \mathrm{GL}_k(q) \times \mathrm{Mon}(n, q)$ and $X \subseteq \mathbb{F}_q^{k \times n}$ the set of all full rank $k \times n$ matrices over $\mathbb{F}_q$. The group action is given by $(L, Q) \star G = LGQ$.*

The Group Action Inversion Problem for the above action is usually called *Linear Code Equivalence* (LEP). In the rank metric, we have the following modelling.

Definition 5. *Let $G = \mathrm{GL}_n(q) \times \mathrm{GL}_m(q) \times \mathrm{GL}_k(q)$ and let X be the set of k-dimensional subspaces of $\mathbb{F}_q^{n \times m}$ represented by their bases. The group action is given by $(A, B, C) \star \langle M_1, \ldots, M_k \rangle = \langle AM'_1B, \ldots, AM'_kB \rangle$, where $M'_i = \sum_{j=1}^{k} C_{ij} M_j$.*

The GAIP for this action is known as *Matrix Code Equivalence* (MCE).

3 Equivalence Relations from Groups Factorisations

Given a group action $(G, X, \star)$, suppose that we can write G as $G_1 \rtimes G_2$. Let ψ be the homomorphism from G_2 to the automorphism group of G_1 used in the semidirect product, sending $h \in G_2$ to the automorphism of G_1 $\psi_h : G_1 \to G_1$. In the rest of the paper, we assume that the group factorisation is efficiently computable, i.e. for any $g \in G$, it is feasible to find its decomposition into $(g_1, g_2) \in G_1 \rtimes G_2$. From $(G, X, \star)$, it is natural to define the following relation on $X \times X$

$$x \sim y \iff \exists g_1 \in G_1 \text{ such that } y = (g_1, e) \star x$$

and it is easy to show that $\sim$ is an equivalence relation. Given the quotient space $X_\sim$ with respect to the equivalence $\sim$, we can define a new group action $(G_2, X_\sim, \star_\sim)$ as follows

$$g_2 \star_\sim [x]_\sim \mapsto [(e, g_2) \star x]_\sim. \tag{1}$$

To show that the action is well-defined, let $g_2 \in G_2$ and let $x \sim y$. Then, there exists $g_1 \in G_1$ such that $y = (g_1, e) \star x$ and

$$\begin{aligned} g_2 \star_\sim [y]_\sim &= [(e, g_2) \star ((g_1, e) \star x)]_\sim = [(\psi_{g_2}(g_1), g_2) \star x]_\sim \\ &= [(e, g_2) \star x]_\sim = g_2 \star_\sim [x]_\sim. \end{aligned}$$

Note that if the relation is defined via G_2, the action above is not well-defined. In fact, it is possible to show that to obtain a well-defined action, G_1 must be normal in G.

3.1 Verifying Orbit Equivalence

To deal with the orbits, our first approach requires the existence of an efficient algorithm that checks the equivalence. As an additional feature for the security reductions, on input x_0 and x_1, if they are in the same orbit with respect to $\sim$, we need that this algorithm returns an element g_1 of G_1 such that $x_1 = (g_1, e) \star x_0$.

Definition 6. *Let $(G, X, \star)$ be a group action such that $G = G_1 \rtimes G_2$. An* orbit equivalence algorithm for G_1 *is a polynomial-time computable map $\mathsf{OE}\colon X \times X \to G_1 \cup \{\perp\}$ such that $\mathsf{OE}(x_0, x_1) \in G_1$ and $(\mathsf{OE}(x_0, x_1), e) \star x_0 = x_1$ if and only if x_0 and x_1 are in the same orbit with respect to $\sim$, and $\mathsf{OE}(x_0, x_1) = \perp$ otherwise.*

Restricting the action to G_2 without a canonical representation of the elements in $X_\sim$ would require a new security assumption. However, the existence of an orbit equivalence algorithm allows us to define a modified sigma protocol for the action $(G, X, \star)$, with short responses, without changing the assumptions.

In short, we build a sigma protocol for the following relation

$$\mathcal{R}_{G_1} = \{((x_0, x_1), g_2) \in (X \times X) \times G_2 \mid \exists g_1 \in G_1 \text{ s.t. } (g_1, g_2) \star x_0 = x_1\}.$$

Observe that the existence of an orbit equivalence algorithm OE implies that $\mathcal{R}_{G_1}$ is an NP-relation. Let $\mathcal{R}$ be standard relation of the action $(G, X, \star)$

$$\mathcal{R} = \{((x_0, x_1), g) \in (X \times X) \times G \mid g \star x_0 = x_1\},$$

then $\mathcal{R}_{G_1}$ and $\mathcal{R}$ define the same language in NP. In particular, given a pair (x_0, x_1), the problems of finding a g in G such that $x_1 = g \star x_0$ can be reduced to the problem of finding g_2 in G_2 such that $[x_1]_\sim = g_2 \star_\sim [x_0]_\sim$. Hence, one can store and send only elements in the group G_2 for the secret, without incurring in security losses.

The sigma protocol for $\mathcal{R}_{G_1}$ we define runs as follows. The Prover and the Verifier have a statement $(x_0, x_1) \in X \times X$, while the Prover knows a witness $g_2 \in G_2$ for it. We suppose that an orbit equivalence algorithm OE for G_1 is known.

1. $\mathcal{P}_1((x_0, x_1), g_2)$: picks at random an element $(h_1, h_2) \in G_1 \rtimes G_2$ and sends to the Verifier $\mathsf{com} = (h_1, h_2) \star x_0$ as a commitment.
2. $\mathcal{V}_1((x_0, x_1), \mathsf{com})$: generate a random challenge $\mathsf{ch} \in \{0, 1\}$ and sends it to the Prover.
3. $\mathcal{P}_2((x_0, x_1), g_2, \mathsf{com}, \mathsf{ch})$: if $\mathsf{ch} = 0$ set $\mathsf{rsp} = h_2$, otherwise they set $\mathsf{rsp} = h_2 g_2^{-1}$ and send it to the Verifier.
4. $\mathcal{V}_2((x_0, x_1), \mathsf{com}, \mathsf{ch}, \mathsf{rsp})$: first, they set $y = (e, \mathsf{rsp}) \star x_{\mathsf{ch}}$. Then, they check that $\mathsf{OE}(y, \mathsf{com}) \neq \perp$. If the check succeeds, then they accept; otherwise reject.

Theorem 1. *The sigma protocol for the relation $\mathcal{R}_{G_1}$ presented above is correct, 2-special sound and perfect honest-verifier zero-knowledge.*

Proof. The sigma protocol for the relation $\mathcal{R}_{G_1}$ presented above is a slight modification of the standard one used for group actions, i.e. a generalisation of the protocol for Graph Isomorphism from [13]. Here, we use the action $(G_2, X_\sim, \star_\sim)$ given by the subgroup G_2 on the set $X_\sim$ of the orbits of X under the action of G_1. The existence of the map OE implies that the action of G_1 over X is easy to invert. Moreover, since the initial action $(G, X, \star)$ is effective, so is $(G_2, X_\sim, \star_\sim)$, except for the *unique string representation* property for $X_\sim$. In the sigma protocol, this is addressed using the map OE in point 4. of the algorithm. □

From the above sigma protocol, an identification scheme can be derived. The key generation algorithm, sample at random $(g_1, g_2) \in G$ and $x_0 \in X$, then sets $(x_0, (g_1, g_2) \star x_0)$ as public key and g_2 as private key. To reach a security level of λ bits, the interactive phase is then repeated λ times in parallel.
This scheme can be turned into a digital signature through standard techniques in the ROM. Unfortunately, since the verifier needs to check the orbit equivalence between $(e, \mathsf{rsp}) \star x_{\mathsf{ch}}$ and com, the resulting signature is not *commitment-recoverable*, i.e. the commitment cannot be computed from the knowledge of ch and rsp. Hence, compared to the signatures analysed in Section 4, there would be no gain with respect to signature size.

3.2 Canonical Forms

The second approach concerns a class of functions that leads to efficient orbit equivalence algorithms. To prove that two orbits of $X_\sim$ are the same, we use a special class of representative computable via a canonical form.

Definition 7. *A* canonical form with failure *for a relation $\sim$ on $X \times X$ is a map $\mathsf{CF}_\sim : X \to X \cup \{\perp\}$ such that, for any $x, y \in X$,*

1. *if $x \sim y$ then $\mathsf{CF}_\sim(x) = \mathsf{CF}_\sim(y)$;*
2. *if $\mathsf{CF}_\sim(x) \neq \perp$ then $\mathsf{CF}_\sim(x) \sim x$.*

If $\mathsf{CF}_\sim(x) = \perp$ we say that $\mathsf{CF}_\sim$ fails on the element x. Notice that when $\mathsf{CF}_\sim(x) = \mathsf{CF}_\sim(y) \neq \perp$, the second property implies $x \sim y$. Moreover, if $\sim$ is defined as above, we will use $\mathsf{CF}^*_\sim$ to define the map that returns both the canonical form and the moving element $g_1 \in G$, and we assume that $\mathsf{CF}^*_\sim$ can be always obtained from $\mathsf{CF}_\sim$.

If there exists an efficiently computable *canonical form* CF with low failure probability, such that $x \sim y$ if and only if $\mathsf{CF}(x) = \mathsf{CF}(y)$ for every x and y in X, then the above action is efficiently computable as follows. We identify the orbits of $X_\sim$ with the representatives given by the canonical form CF and the action is given by

$$g_2 \star_\sim x \mapsto \mathsf{CF}((e, g_2) \star x).$$

Similarly to the action of Equation (1), the map above is well-defined, and it leads to an effective group action.

Proposition 1. *If there exists a polynomial-time computable canonical form* CF *for the equivalence* $\sim$*, the Group Action Inverse problems for* $(G, X, \star)$ *and* $(G_2, X_\sim, \star_\sim)$ *are polynomially equivalent.*

Proof. We refer to the Group Action Inverse problems for $(G, X, \star)$ and $(G_2, X_\sim, \star_\sim)$ as $\mathrm{GAIP}_\star$ and $\mathrm{GAIP}_{\star_\sim}$, respectively.
The reduction from $\mathrm{GAIP}_{\star_\sim}$ to $\mathrm{GAIP}_\star$ is as follows. Let (x, y) be an instance of $\mathrm{GAIP}_{\star_\sim}$. This means that y is in canonical form and $\mathsf{CF}(y) = y$. The pair (x, y) can be seen as an instance of $\mathrm{GAIP}_\star$, and finding $g = (g_1, g_2)$ such that $(g_1, g_2) \star x = y$ implies that

$$g_2 \star_\sim x = \mathsf{CF}((e, g_2) \star x) = \mathsf{CF}((g_1, e)(e, g_2) \star x) = \mathsf{CF}((g_1, g_2) \star x) = \mathsf{CF}(y) = y,$$

and we are done.

Vice versa, one can reduce $\mathrm{GAIP}_\star$ to $\mathrm{GAIP}_{\star_\sim}$ as follows. Let (x, y) be an instance of $\mathrm{GAIP}_\star$. For every $z \in X$, let g_z the element of G_1 returned by $\mathsf{CF}^*(z)$, so that $(g_z, e) \star z = \mathsf{CF}(z)$. Let $(x, \mathsf{CF}(y))$, with $\mathsf{CF}(y) = (g_y, e) \star y$, be an instance of $\mathrm{GAIP}_{\star_\sim}$ whose solution is given by g_2. This means that

$$\mathsf{CF}(y) = g_2 \star_\sim x = \mathsf{CF}((e, g_2) \star x) = (\tilde{g}, g_2) \star x,$$

where $\tilde{g}$ is obtained from CF^*. Then, we have that

$$\begin{aligned}(g_y^{-1}\tilde{g}, g_2) \star x &= (g_y^{-1}\psi_e(\tilde{g}), g_2) \star x = (g_y^{-1}, e)(\tilde{g}, g_2) \star x \\ &= (g_y^{-1}, e) \star ((\tilde{g}, g_2) \star x) = (g_y^{-1}, e) \star \mathsf{CF}(y) = y\end{aligned}$$

and we found a solution for the instance (x, y) of $\mathrm{GAIP}_\star$. □

The above results imply that, if one is able to factorise G and a polynomial-time computable canonical form with respect to the relation for a factor G_1, then the induced action $(G_2, X_\sim, \star_\sim)$, where G_2 is the remaining factor, can be used without introducing new computational assumptions. This means that, instead of using elements from the whole group G, one can use elements from G_2, potentially reducing the sizes of the elements involved. This is implicitly used in the Linear Code Equivalence Problem when the systematic form is employed.

4 Applications

4.1 Matrix Code Equivalence

Here we show an application of the above technique to reduce the sizes of MEDS [9], a digital signature scheme based on the equivalence of matrix codes. Its security relies on the hardness of the Matrix Code Equivalence problem, based on the action of Definition 5. Given the group $G = \mathrm{GL}_n(q) \times \mathrm{GL}_m(q) \times \mathrm{GL}_k(q)$, in [9], using the systematic form for the action of the last factor $\mathrm{GL}_k(q)$, they implicitly use only the action of the remaining part of the group $\mathrm{GL}_n(q) \times \mathrm{GL}_m(q)$. Here we go further, quotienting on the factors $\mathrm{GL}_m(q) \times \mathrm{GL}_k(q)$ and using in the

sigma protocol (and hence in the signature) only elements in $\mathrm{GL}_n(q)$.
Let us start by recalling that the action on $n \times m$ k-dimensional matrix codes can be seen as the action of G on $n \times mk$ matrices as follows. Let $M_1, \ldots, M_k$ be a basis of a matrix code, then the action of (A, B, C) is defined as $CM(A^T \otimes B)$, where $M = [M_1 \mid M_2 \mid \ldots \mid M_k] \in \mathbb{F}_q^{n \times mk}$.

One can notice that if we factor $G = G_1 \times G_2$ where $G_1 = \mathrm{GL}_m(q) \times \mathrm{GL}_k(q)$ and $G_2 = \mathrm{GL}_n(q)$, the action of G_1 is equivalent to a special case of the Matrix Space Conjugacy problem that is solvable in polynomial time [5,14,7]. Even if this approach leads to an efficient orbit equivalence algorithm, to obtain a gain in the signature size, we need to present a canonical form for the following relation

$$M \sim N \iff \exists (B, C) \in \mathrm{GL}_m(q) \times \mathrm{GL}_k(q) \text{ such that } N = CM(\mathbf{I}_n \otimes B).$$

From now on, we assume $n = m$ as in the parameter sets from the MEDS submission [8].

Let $M = [M_1 \mid M_2 \mid \ldots \mid M_k] \in \mathbb{F}_q^{n \times nk}$, with $M_i \in \mathbb{F}_q^{n \times n}$, the canonical form (with failure) $\mathsf{CF}(M)$ is computed as follows:

1. Let $1 \leq j \leq k$ be the smallest index for which the j-th block M_j is invertible and compute $\bar{M} = M_j^{-1} M$. If an invertible block does not exist, the procedure fails and returns $\perp$.
2. Let $j' = j + 1 \pmod k$. Find the solution set V of invertible matrices $B \in \mathrm{GL}_n(q)$ such that $B^{-1} \bar{M}_{j'} B$ is equal to the circulant matrix $\mathrm{circ}(e_n)$ on the first $n - 1$ columns. If the solution set is empty, the procedure fails and returns $\perp$.
3. Let $j'' = j + 2 \pmod k$. Given a total ordering on $\mathbb{F}_q^n$, find the unique solution $B \in V$ (up to a constant factor) that minimizes the first column of $B^{-1} \bar{M}_{j''} B$.
4. The canonical form of M is computed as $\mathsf{CF}(M) = (M_j B)^{-1} M(\mathbf{I}_k \otimes B)$.

Proposition 2. CF *is a canonical form for the relation* $\sim$.

Proof. We prove that CF is a canonical form according to Definition 7. Let $M = [M_1 \mid M_2 \mid \ldots \mid M_k] \in \mathbb{F}_q^{n \times nk}$, $M_i \in \mathbb{F}_q^{n \times n}$. Suppose $\mathsf{CF}(M) \neq \perp$, then $\mathsf{CF}(M) = (M_j B)^{-1} M(\mathbf{I}_k \otimes B)$, for some $1 \leq j \leq k$ and $B \in \mathrm{GL}_n(q)$. Then the i-th block of $\mathsf{CF}(M)$ is given by $(M_j B)^{-1} M_i B$, which implies $\mathsf{CF}(M) \sim M$.

Let $M \sim N$, i.e. there exists $X, Y \in \mathrm{GL}_n(q)$ such that $N_i = X M_i Y$, for all $1 \leq i \leq k$. Then, since M_j is invertible, so is N_j and it holds $\mathsf{CF}(N) = (N_j B')^{-1} N(\mathbf{I}_k \otimes B')$ for some $B' \in \mathrm{GL}_n(q)$. Let V (resp. V') be the solution set of invertible matrices $B \in \mathrm{GL}_n(q)$ (resp. B') such that $B^{-1} M_j^{-1} M_{j'} B$ (resp. $B'^{-1} N_j^{-1} N_{j'} B'$) is equal to the circulant matrix $\mathrm{circ}(e_n)$ on the first $n-1$ columns. Then, there is a one-to-one correspondence between V and V' given by $B \mapsto Y^{-1} B$. It follows that

$$\begin{aligned} \mathsf{CF}(N) &= (N_j B')^{-1} N(\mathbf{I}_k \otimes B') = (X M_j Y B')^{-1} X M(\mathbf{I}_k \otimes Y B') \\ &= (Y B')^{-1} M_j^{-1} M(\mathbf{I}_k \otimes Y B') = (M_j B)^{-1} M(\mathbf{I}_k \otimes B) = \mathsf{CF}(M). \quad \square \end{aligned}$$

Table 1. Signature sizes (in bytes) for MEDS.

Parameter set	Sec. Level	MEDS [8]	This work	Gain
MEDS-13220	I	12976	7516	42.1%
MEDS-69497	III	54736	29788	45.6%
MEDS-167717	V	165332	86462	47.7%

Unfortunately, this canonical form, even if it can be computed in expected polynomial time, is not efficient for practical applications. Observe that the most burdensome task is given by step 3 of the computation of CF. To overcome this limitation, we can slightly modify the sigma protocol by including additional information in the response to quickly identify a particular class representative. Consider the standard sigma protocol for a cryptographic group action. The commitment is the element $\mathsf{CF}(h_2 \star_\sim [x]_\sim)$ for a random $h_2 \in G_2$. In the computation of the canonical form, during step 3, the signer chooses the column instead of finding the minimal one as in the algorithm. This column is then sent with the response, so the verifier can efficiently compute the same representative by constraining the choice of column during step 3. This strategy leads to more efficient signing and verifying processes, making the signature doable.

Concerning the version of MEDS using the action of $\mathrm{GL}_n(q) \times \mathrm{GL}_m(q)$ from [9], our proposal allows to reduce the size of the signature of about 45% for the last version of the parameter sets given in [8], as reported in Table 1. The gain in the signature dimensions comes at the cost of running the canonical form algorithm both in the signing and verification phases.

4.2 Linear Code Equivalence

LESS is a digital signature scheme based on the equivalence of linear codes, which can be described in the framework of group actions. For $1 \leq k \leq n$, let $\mathbb{F}_q^{k\times n}$ be the linear space of $k \times n$ matrices over $\mathbb{F}_q$. Let $\mathrm{Mon}(n, q)$ be the group of $n \times n$ monomial matrices over $\mathbb{F}_q$. We consider the group action $\star$ described in Definition 4 of $G = \mathrm{GL}_k(q) \times \mathrm{Mon}(n, q)$ on $X \subseteq \mathbb{F}_q^{k\times n}$, the set of all full rank $k \times n$ matrices over $\mathbb{F}_q$.

It is well known that $\mathrm{Mon}(n, q)$ is isomorphic to the semidirect product $\mathcal{S}_n \ltimes (\mathbb{F}_q^*)^n$, where $(\mathbb{F}_q^*)^n$ is isomorphic to the group of non-singular $n \times n$ diagonal matrices. The group in the previous definition can then be factorised as $G = \mathrm{GL}_k(q) \times (\mathcal{S}_n \ltimes (\mathbb{F}_q^*)^n)$. Observe that G is isomorphic to $(\mathrm{GL}_k(q) \times (\mathbb{F}_q^*)^n) \rtimes \mathcal{S}_n$ and we can apply the framework of the previous section by defining the following relation on $X \times X$:

$$G \sim G' \iff \exists (L, D) \in \mathrm{GL}_k(q) \times (\mathbb{F}_q^*)^n \text{ such that } G' = LGD = (L, (\mathbf{I}_n, D)) \star G.$$

To show that the induced group action $(\mathcal{S}_n, X_\sim, \star_\sim)$ can be efficiently computed, we introduce the following canonical form (with failure) CF for $\sim$.

Let $G \in X$, the canonical form $\mathsf{CF}(G)$ is computed as follows:

Table 2. Signature sizes (in bytes) for LESS.

Parameter set	Sec. Level	LEP	IS-LEP [15]	CF-LEP [10]	This work
LESS-1b	I	15726	8646	2496	9096
LESS-3b	III	30408	17208	5658	18858
LESS-5b	V	53896	30616	10056	34696

1. Compute the *Reduced Row-Echelon Form* (RREF) of G.
2. Let $1 \leq j \leq n$ be the smallest index for which the j-th column $g_j = (g_{1,j}, \ldots, g_{k,j})$ of $\mathrm{RREF}(G)$ has only non-zero elements. If a column of this form does not exist, the procedure fails and returns $\perp$. Compute $D_r = \mathrm{diag}(g_{1,j}^{-1}, \ldots, g_{k,j}^{-1}) \in \mathbb{F}_q^{k \times k}$.
3. Let b_j be the first non-zero element of each column of $D_r\,\mathrm{RREF}(G)$, for $1 \leq j \leq n$. Compute $D_c = \mathrm{diag}(b_1^{-1}, \ldots, b_n^{-1}) \in \mathbb{F}_q^{n \times n}$.
4. The canonical form of G is computed as $\mathsf{CF}(G) = D_r\,\mathrm{RREF}(G) D_c$.

Proposition 3. CF *is an efficient canonical form for the relation* $\sim$.

LESS implicitly use the framework with canonical form by working with the RREF of elements in X. Compared to this basic form, in our version, the response size changes from $n(\lceil \log_2 n \rceil + \lceil \log_2(q-1) \rceil)$ bits, required to represent an element of $\mathrm{Mon}(n, q)$, to $n\lceil \log_2 n \rceil$, required for an element of S_n. However, the version of LESS submitted to NIST includes the Information Set-LEP variant introduced in [15]. With this variant, the commitment generation and the verification procedure are modified so that it is possible to reduce the response size to $k(\lceil \log_2 n \rceil + \lceil \log_2(q-1) \rceil)$ bits. Moreover, the authors of LESS recently presented in [10] a new notion of equivalence for codes and proved that it reduces to linear equivalence. This leads to an even more significant reduction in the size of responses. This last variant can partially be framed within our framework. In particular, let H be a subgroup of G and S be a subset of G such that $e \in S$, and suppose that for each $g \in G$ there exist unique elements $h \in H, s \in S$ such that $g = hs$. Then, as in Section 3, we can take the relation $\sim$ on $X \times X$ induced by H and consider the quotient space $X_\sim$. However, we cannot define a new group action restricted to S since it is not a group. On the other hand, if we know a canonical form $\mathsf{CF}_\sim$ for $\sim$, this is enough to define a sigma protocol based on the original group action, where responses are computed as the factor in S of the considered element in G. This requires the definition of a new security assumption based on a variant of the original problem where the action is taken on $X_\sim$ via the canonical form[4]. Unlike in our framework, this variant must be explicitly shown to be equivalent to the original assumption. Further research should consider the possibility of extending the results of Section 3 to a generic factorisation involving a subset of G. See Table 2 for a comparison.

[4] In the context of LEP, the authors of [10] refer to this variant as Canonical Form-LEP

References

1. Alamati, N., Feo, L.D., Montgomery, H., Patranabis, S.: Cryptographic group actions and applications. In: International Conference on the Theory and Application of Cryptology and Information Security. pp. 411–439. Springer (2020)
2. Barenghi, A., Biasse, J.F., Persichetti, E., Santini, P.: LESS-FM: fine-tuning signatures from the code equivalence problem. In: International Conference on Post-Quantum Cryptography. pp. 23–43. Springer (2021)
3. Beullens, W., Katsumata, S., Pintore, F.: Calamari and Falafl: logarithmic (linkable) ring signatures from isogenies and lattices. In: Advances in Cryptology–ASIACRYPT 2020: 26th International Conference on the Theory and Application of Cryptology and Information Security, Daejeon, South Korea, December 7–11, 2020, Proceedings, Part II. pp. 464–492. Springer (2020)
4. Brassard, G., Yung, M.: One-way group actions. In: Advances in Cryptology-CRYPTO'90: Proceedings 10. pp. 94–107. Springer (1991)
5. Brooksbank, P.A., Luks, E.M.: Testing isomorphism of modules. Journal of Algebra **320**(11), 4020–4029 (2008)
6. Castryck, W., Lange, T., Martindale, C., Panny, L., Renes, J.: CSIDH: an efficient post-quantum commutative group action. In: Advances in Cryptology–ASIACRYPT 2018: 24th International Conference on the Theory and Application of Cryptology and Information Security, Brisbane, QLD, Australia, December 2–6, 2018, Proceedings, Part III 24. pp. 395–427. Springer (2018)
7. Chistov, A., Ivanyos, G., Karpinski, M.: Polynomial time algorithms for modules over finite dimensional algebras. In: Proceedings of the 1997 international symposium on Symbolic and algebraic computation. pp. 68–74 (1997)
8. Chou, T., Niederhagen, R., Persichetti, E., Ran, L., Randrianarisoa, T.H., Reijnders, K., Samardjiska, S., Trimoska, M.: Matrix Equivalence Digital Signature (2023)
9. Chou, T., Niederhagen, R., Persichetti, E., Randrianarisoa, T.H., Reijnders, K., Samardjiska, S., Trimoska, M.: Take your MEDS: digital signatures from matrix code equivalence. In: International Conference on Cryptology in Africa. pp. 28–52. Springer (2023)
10. Chou, T., Persichetti, E., Santini, P.: On linear equivalence, canonical forms, and digital signatures. Cryptology ePrint Archive (2023)
11. Couveignes, J.M.: Hard homogeneous spaces. Cryptology ePrint Archive (2006)
12. De Feo, L., Galbraith, S.D.: SeaSign: compact isogeny signatures from class group actions. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. pp. 759–789. Springer (2019)
13. Goldreich, O., Micali, S., Wigderson, A.: Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. Journal of the ACM (JACM) **38**(3), 690–728 (1991)
14. Ivanyos, G., Karpinski, M., Saxena, N.: Deterministic polynomial time algorithms for matrix completion problems. SIAM journal on computing **39**(8), 3736–3751 (2010)
15. Persichetti, E., Santini, P.: A New Formulation of the Linear Equivalence Problem and Shorter LESS Signatures. In: Guo, J., Steinfeld, R. (eds.) Advances in Cryptology – ASIACRYPT 2023. pp. 351–378. Springer Nature Singapore, Singapore (2023)
16. Tang, G., Duong, D.H., Joux, A., Plantard, T., Qiao, Y., Susilo, W.: Practical Post-Quantum Signature Schemes from Isomorphism Problems of Trilinear Forms. In:

Annual International Conference on the Theory and Applications of Cryptographic Techniques. pp. 582–612. Springer (2022)

Additive twisted codes: new distance bounds and infinite families of quantum codes

Reza Dastbasteh[1] and Petr Lisoněk[2]

[1] University of Navarra rdastbas@sfu.ca
[2] Simon Fraser University plisonek@sfu.ca

Abstract. We provide a new construction of quantum codes that enables integration of a broader class of classical codes into the mathematical framework of quantum stabilizer codes. Next, we present new connections between twisted codes and linear cyclic codes and provide novel bounds for the minimum distance of twisted codes. We show that classical tools such as the Hartmann-Tzeng minimum distance bound are applicable to twisted codes. This enabled us to discover five new infinite families and many other examples of record-breaking, and sometimes optimal, binary quantum codes.

Keywords: additive code, twisted code, cyclic code, quantum code, minimum distance bound.

1 Introduction

Let $\mathbb{F}_q$ be the finite field of q elements. An $\mathbb{F}_2$-linear subspace $C \subseteq \mathbb{F}_4^n$ is called an *additive* code over $\mathbb{F}_4$. Additive codes are especially important due to their application in the construction of binary quantum codes. The class of additive twisted codes is possibly the most structured family of additive codes. They were first introduced as a subclass of additive cyclic codes by Jürgen Bierbrauer and Yves Edel [7]. Twisted codes, like linear cyclic codes, are defined and constructed using (unique) defining sets, and the BCH minimum distance bound holds for them [7]. Moreover, several families and examples of good quantum codes are constructed using dual-containing twisted codes [2]. While the original work on the additive twisted codes has been widely referenced in literature, twisted codes have not been developed much since their invention. This is likely due to their study being technically much more difficult than the study of many other common families of codes.

Quantum error-correcting codes, or simply quantum codes, are applied to protect quantum information from corruption by noise (decoherence) on the quantum channel in a way that is similar to that of classical error-correcting codes. This extended abstract exclusively deals with *binary quantum codes.* The parameters of a binary quantum code that encodes k logical qubits into n physical qubits and has minimum distance d are denoted by $[\![n, k, d]\!]$. The most common approach to construction of quantum codes is by using the stabilizer

formalism which builds a bridge between certain dual-containing additive codes and quantum (stabilizer) codes [4, 8]. One of the main challenges of quantum stabilizer codes is its dual-containment condition, which only allows a small number of classical codes to be used to construct good quantum codes.

In this extended abstract, we first give a novel construction of binary stabilizer quantum codes that makes it possible to also use additive codes that are not dual-containing. Next, we introduce a new perspective on twisted codes by viewing each code as an additive subcode of a particular linear cyclic code. This new approach provides a stronger connection between twisted codes and linear cyclic codes, enabling us to give novel minimum distance lower and upper bounds for twisted codes and show new similarities between twisted codes and linear cyclic codes. In particular, we prove that the Hartmann-Tzeng bound holds for twisted codes. We demonstrate that five infinite families of record-breaking, and sometimes optimal, quantum codes can be constructed from twisted codes using these bounds. To recognize that a quantum code is record-breaking and/or optimal we refer to the tables maintained by Markus Grassl [9].

Remark 1. A comprehensive elaboration of the material introduced in this extended abstract can be found in [6, Chapter 3]. In this extended abstract, we mainly list our main results without a complete proof. For a more in-depth exploration of the details, we suggest consulting the aforementioned reference.

2 Background

Let $\mathbb{F}_4 = \{0, 1, w, w^2\}$ be the field of four elements, where $w^2 = w + 1$. An additive code $C \subseteq \mathbb{F}_4^n$ with $\mathbb{F}_2$-dimension k will be denoted by $(n, 2^k)$. Similar to linear codes, the minimum weight among all non-zero codewords of an additive code C is called the *minimum distance* of C, and it will be denoted $d(C)$.

Let $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n) \in \mathbb{F}_4^n$. The trace map $\mathrm{Tr} : \mathbb{F}_4 \to \mathbb{F}_2$ is defined by $\mathrm{Tr}(x) = x + \overline{x}$, where $\overline{x} = x^2$. We define conjugate of the vector u by $\overline{u} = (\overline{u_1}, \overline{u_2}, \ldots, \overline{u_n})$. The dot product of vectors u and v will be denoted $u \cdot v$. The *trace inner product* of u and v is defined by

$$u * v = \mathrm{Tr}(u \cdot \overline{v}) = (u \cdot \overline{v}) + \overline{(u \cdot \overline{v})} = \sum_{i=1}^{n} (u_i \overline{v_i} + \overline{u_i} v_i). \tag{1}$$

If C is an $(n, 2^k)$ additive code, its *trace dual* with respect to the trace inner product is defined by

$$C^{\perp_t} = \{u \in \mathbb{F}_4^n : u * v = 0 \text{ for all } v \in C\}.$$

It is easy to see that $*$ is non-degenerate and $C^{\perp_t}$ is an $(n, 2^{2n-k})$ additive code. We call an additive code C a *dual-containing* (respectively *self-dual*) code with respect to the trace inner product if $C^{\perp_t} \subseteq C$ (respectively $C^{\perp_t} = C$).

The mathematical formalism of quantum stabilizer code, as described in the next theorem, provides a sufficient condition for constructing binary quantum codes from additive codes over $\mathbb{F}_4$.

Theorem 1. *[4] Let $C \subseteq \mathbb{F}_4^n$ be an $(n, 2^{n+k}, d)$ additive code such that $C^{\perp_t} \subseteq C$. Then an $[\![n, k, d']\!]$ binary quantum stabilizer code can be constructed, where d' is the minimum weight in $C \setminus C^{\perp_t}$ if $k > 0$ and $d' = d$ if $k = 0$.*

Proof. The proof follows from [4, Theorem 2].

If $d = d'$ the above quantum code is called *pure*, and otherwise ($d < d'$) *impure*.

2.1 Additive twisted codes

In this subsection, we provide a brief overview of the construction of additive twisted codes and highlight certain key properties. For more in-depth details, interested readers are encouraged to refer to [1, Section 17.2], [2], or [6, Chapter 3].

Let n be a positive integer such that $n \mid 2^r - 1$ for some positive integer r, and $\mathbb{F}_{2^r}$ be the field of 2^r elements. The surjective $\mathbb{F}_2$-linear map $\phi_\gamma : \mathbb{F}_{2^r} \to \mathbb{F}_2 \times \mathbb{F}_2$ is defined by

$$\phi_\gamma(x) = (\mathrm{Tr}_1^r(x), \mathrm{Tr}_1^r(\gamma x)), \tag{2}$$

where $\gamma \in \mathbb{F}_{2^r} \setminus \mathbb{F}_2$ and Tr_1^r is the trace map from $\mathbb{F}_{2^r}$ to $\mathbb{F}_2$. Since $n \mid 2^r - 1$, the multiplicative group $\mathbb{F}_{2^r}^*$ contains all the n-th roots of unity, namely $W = \{1, \alpha^1, \alpha^2, \ldots, \alpha^{n-1}\}$, where α is a primitive n-th root of unity in $\mathbb{F}_{2^r}^*$. Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$. We define $B(A)$ to be the matrix over $\mathbb{F}_{2^r}$ whose rows and columns are labelled by elements of A and W, respectively, and the entry in row j and column α^i is α^{ij}. Let $C(A)$ be the length n linear cyclic code over $\mathbb{F}_{2^r}$ with the defining set A. Then $B(A)$ is a generator matrix for the code $C(A)^\perp$, the Euclidean dual of $C(A)$. We define $\phi_\gamma(C(A)^\perp)$ to be the $\mathbb{F}_2$-linear code

$$\phi_\gamma(C(A)^\perp) = \{\phi_\gamma(c) : c \in C(A)^\perp\}. \tag{3}$$

Let $v = (v_1, v_2, \ldots, v_n)$ be a vector in $C(A)^\perp$. In this extended abstract, we represent the vector $\phi_\gamma(v)$ by $\phi_\gamma(v) = \big((v_{11}, v_{12}), \ldots, (v_{n1}, v_{n2})\big)$, where $v_{i1} = \mathrm{Tr}_1^r(v_i)$ and $v_{i2} = \mathrm{Tr}_1^r(\gamma v_i)$ for each $1 \le i \le n$.

Definition 1. *Let $\langle , \rangle_s : \mathbb{F}_2^{2n} \times \mathbb{F}_2^{2n} \to \mathbb{F}_2$ be the nondegenerate symplectic $\mathbb{F}_2$-bilinear form defined by*

$$\langle \Big((a_{11}, a_{12}), \ldots, (a_{n1}, a_{n2})\Big), \Big((b_{11}, b_{12}), \ldots, (b_{n1}, b_{n2})\Big) \rangle_s = \sum_{i=1}^{n} a_{i1}b_{i2} - a_{i2}b_{i1}. \tag{4}$$

Definition 2. *Let $n \mid 2^r - 1$ for some integer r and A be a subset of $\mathbb{Z}/n\mathbb{Z}$. The dual of the code $\phi_\gamma(C(A)^\perp)$ with respect to the symplectic inner product $\langle , \rangle_s$ is called a* twisted code *of length n over $\mathbb{F}_2 \times \mathbb{F}_2$. Such a twisted code will be denoted by $\mathscr{C}_\gamma(A)$. In other words,*

$$\mathscr{C}_\gamma(A) = \big(\phi_\gamma(C(A)^\perp)\big)^{\perp_s}.$$

The $\mathbb{F}_2$-linear isomorphism $\psi : \mathbb{F}_2^{2n} \to \mathbb{F}_4^n$ defined by

$$\psi\big((a_{11}, a_{12}), \ldots, (a_{n1}, a_{n2})\big) = (a_{11}\omega + a_{12}\omega^2, a_{21}\omega + a_{22}\omega^2, \ldots, a_{n1}\omega + a_{n2}\omega^2) \tag{5}$$

maps each twisted code into an additive code over $\mathbb{F}_4$. Moreover, we have

$$\langle u, v\rangle_s = \psi(u) * \psi(v)$$

for each $u, v \in \mathbb{F}_2^{2n}$.

In general, the set A in the above definition is not unique, that is, a twisted code can be constructed using different subsets of $\mathbb{Z}/n\mathbb{Z}$. We denote $\kappa = [\mathbb{F}_2(\gamma) : \mathbb{F}_2]$, which is an integer greater than one. For each $a \in \mathbb{Z}/n\mathbb{Z}$, the set

$$Z(a) = \{(a2^j) \bmod n : 0 \le j \le m-1\},$$

where m is the smallest positive integer such that $a2^m \equiv a \pmod n$, is called the 2-*cyclotomic coset* modulo n containing a.

Definition 3. *[1, 2] Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$ and $a \in A$. If $\kappa \mid |Z(a)|$ and $\kappa \mid i-j$ for each $2^i a, 2^j a \in Z(a) \cap A$, then $Z(a) \cap A$ is called unsaturated. Otherwise, $Z(a) \cap A$ is called saturated.*

Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$ and Z be a 2-cyclotomic coset modulo n such that $Z \cap A$ is unsaturated and $a \in Z \cap A$. We define

$$(Z \cap A)^H = \{a2^{\kappa i} : 0 \le i \le \frac{|Z(a)|}{\kappa} - 1\}.$$

In this case, there exists a subgroup H of the Galois group of $\mathbb{F}_{2^r}/\mathbb{F}_2$ of size $\frac{|Z(a)|}{\kappa}$ that acts transitively on $(Z \cap A)^H$. Lemma 3 of [2] shows that the set

$$\tilde{A} = \bigcup_{Z \cap A \ sat} Z \bigcup_{Z \cap A \ unsat} (Z \cap A)^H \tag{6}$$

is the largest defining set (called the *complete defining set*) that the twisted code $\mathscr{C}_\gamma(A)$ can have. Moreover $\mathscr{C}_\gamma(A)^{\perp_s} = \phi_\gamma(C(A)^\perp) = \mathscr{C}_\gamma(A_d)$, where

$$A_d = \bigcup_{Z \cap A = \emptyset} -Z \bigcup_{Z \cap A \ unsat} -((Z \cap A)^H). \tag{7}$$

This observation implies that a twisted code $\mathscr{C}_\gamma(A)$ is dual-containing if and only if $A \subseteq A_d$. The dimension of each twisted code is computed using the following theorem.

Theorem 2. *[2, Theorem 5] Let $n \mid 2^r - 1$ be a positive integer and $A \subseteq \mathbb{Z}/n\mathbb{Z}$. Then the $\mathbb{F}_2$-dimension of $\mathscr{C}_\gamma(A)$ is $\sum_Z c_Z(A)$, where the sum runs over all 2-cyclotomic cosets modulo n and*

$$c_Z(A) = \begin{cases} 2|Z| & \text{if } Z \cap A = \emptyset \\ |Z| & \text{if } Z \cap A \text{ is unsaturated} \\ 0 & \text{if } Z \cap A \text{ is saturated.} \end{cases}$$

Example 1. Let $n = 15$. Note that $n \mid 2^4 - 1$ and therefore γ can be any element of $\mathbb{F}_{16} \setminus \mathbb{F}_2$. The 2-cyclotomic coset of 1 modulo 15 is $Z(1) = \{1, 2, 4, 8\}$ and $\mathbb{Z}/15\mathbb{Z} = Z(0) \cup Z(1) \cup Z(3) \cup Z(5) \cup Z(7)$.

Let $\gamma \in \mathbb{F}_{16} \setminus \mathbb{F}_4$ and $A = \{1, 4\}$. Then $\kappa = [\mathbb{F}_2(\gamma) : \mathbb{F}_2] = 4$. In this case, $A \cap Z(1)$ is saturated and A is an incomplete defining set. So we apply (6) to form the complete defining set of A, namely $\tilde{A} = Z(1)$. The twisted code $\mathscr{C}_\gamma(A)$ has dimension $2|Z(0)| + 2|Z(3)| + 2|Z(5)| + 2|Z(7)| = 22$ over $\mathbb{F}_2$. Applying (7) to A, we get $\mathscr{C}_\gamma(A)^{\perp_s} = \mathscr{C}_\gamma(A_d)$, where $A_d = Z(0) \cup Z(1) \cup Z(3) \cup Z(5)$. Hence the fact that $A \subset A_d$ implies that $\mathscr{C}_{\gamma_2}(A)$ is dual-containing.

Note that choosing $\gamma \in \mathbb{F}_4 \setminus \mathbb{F}_2$ ($\kappa = 2$) implies that A is a complete defining set and $A \cap Z(1)$ is unsaturated. In this case the code $\mathscr{C}_\gamma(A)$ has dimension 26 over $\mathbb{F}_2$.

3 Quantum codes from nearly dual-containing additive codes

In this section, we present a new method of constructing quantum stabilizer codes from additive codes over $\mathbb{F}_4$ that are not dual containing with respect to the trace inner product (1). In reality, there exist many classical codes with good parameters that are not dual-containing but are nearly dual-containing, meaning they contain a large subset of their dual. We quantify this by proving formulas for dual-containment deficiency of a code. Using this formula, we give a novel construction of binary stabilizer quantum codes that makes it possible to also use the additive codes that are not dual-containing as its ingredients to construct quantum codes.

Let C be an additive code over $\mathbb{F}_4$. The *dual-containment deficiency* of the code C is defined by

$$\dim_{\mathbb{F}_2}(C^{\perp_t}) - \dim_{\mathbb{F}_2}(C \cap C^{\perp_t}). \tag{8}$$

In particular, if C is dual-containing, then the value of (8) is zero.

In the next theorem we show that the dual-containment deficiency of additive codes is always an even number. Furthermore, we find a basis for $C^{\perp_t}$ such that the first r vectors form a basis for $C \cap C^{\perp_t}$ and the remaining vectors are paired up such that non-orthogonal vectors occur only as the elements of a pair.

Lemma 1. *Let $C \subseteq \mathbb{F}_4^n$ be an additive code. Then $s = \dim_{\mathbb{F}_2}(C^{\perp_t}) - \dim_{\mathbb{F}_2}(C \cap C^{\perp_t})$ is even. Moreover, if $\dim_{\mathbb{F}_2}(C \cap C^{\perp_t}) = r$ and $s = 2e$, then we can find a basis for $C^{\perp_t}$ in the form $\{V_1, V_2, \ldots, V_r, M_1, M_2, \ldots, M_{2e}\}$ such that*

1. *The set $\{V_1, V_2, \ldots, V_r\}$ forms a basis for $C \cap C^{\perp_t}$.*

2. *For all* $1 \leq i, j \leq 2e$ *we have* $M_i * M_j = 1$ *if and only if* $\{i, j\} = \{2t-1, 2t\}$ *for some* $1 \leq t \leq e$.

Proof. The proof follows from the properties of symplectic bilinear forms. For more details see [6, Lemma 3.2.1].

The next theorem, which is a generalization of the result of [12, Theorem 2] to additive codes, states the parameters of a binary quantum code that is constructed from a nearly dual-containing additive code.

Theorem 3. *Let* C *be an* $(n, 2^{n+k})$ *additive code over* $\mathbb{F}_4$ *and* $\dim_{\mathbb{F}_2}(C^{\perp_t}) - \dim_{\mathbb{F}_2}(C \cap C^{\perp_t}) = 2e$. *Then we can construct an* $[\![n+e, k+e, d]\!]$ *binary quantum code, where*

$$d \geq \min\{d(C), d(C + C^{\perp_t}) + 1\}.$$

Proof. We only give a sketch for the proof, and for more details see [6, Theorem 3.2.3]. We first apply Lemma 1 to find a basis for $C^{\perp_t}$ in the given form. Then we add e new coordinates to make new longer vectors corresponding to M_i, M_j symplectic orthogonal for each $(i, j) = (2t-1, 2t)$ and each $1 \leq t \leq e$. This way we produce a dual-containing additive code of length $n + e$. The results concerning the minimum distance and dimension follow by considering this new basis.

This result is very important as it allows us to construct quantum codes from the twisted codes which are not symplectic dual-containing.

4 Minimum distance bounds for twisted codes

Similar to linear cyclic codes, the minimum distance of twisted codes can be bounded using the BCH bound [2]. Currently, this is the only known minimum distance bound for the family of twisted codes. Throughout this section, n is a positive integer such that $n \mid 2^r - 1$ for some positive integer r. Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$. In this section, we first give a connection between the weight of vectors in the length n twisted code $\mathscr{C}_\gamma(A)$ and $C(A)$, where $C(A)$ is the linear cyclic code of length n over $\mathbb{F}_{2^r}$ with defining set A. Then we present new minimum distance bounds for twisted codes.

Recall that $L \subseteq \mathbb{Z}/n\mathbb{Z}$ is called a consecutive set of length s if there exists an integer c with $\gcd(c, n) = 1$ such that

$$\{(cl) \bmod n : l \in L\} = \{(j + t) \bmod n : 0 \leq j \leq s - 1\}$$

for some $t \in \mathbb{Z}/n\mathbb{Z}$. The next proposition gives the BCH minimum distance bound for twisted codes.

Proposition 1. *[2] Let* A *be a defining set of a twisted code* $\mathscr{C}_\gamma(A)$ *such that* A *contains a consecutive set of size* $t - 1$. *Then* $d(\mathscr{C}_\gamma(A)) \geq t$.

The following theorem establishes a more powerful connection between twisted codes and linear cyclic codes.

Theorem 4. *Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$ be a defining set of a twisted code $\mathscr{C}_\gamma(A)$ of length n over $\mathbb{F}_2 \times \mathbb{F}_2$. Then the following statements are equivalent.*

1. *The vector $y = ((b_{11}, b_{12}), (b_{21}, b_{22}), \ldots, (b_{n1}, b_{n2})) \in \mathscr{C}_\gamma(A)$.*
2. *The vector $x = (\gamma b_{11} + b_{12}, \gamma b_{21} + b_{22}, \ldots, \gamma b_{n1} + b_{n2}) \in C(A)$.*

Proof. Let $y = ((b_{11}, b_{12}), (b_{21}, b_{22}), \ldots, (b_{n1}, b_{n2}))$ and $x = (\gamma b_{11} + b_{12}, \gamma b_{21} + b_{22}, \ldots, \gamma b_{n1} + b_{n2})$ for arbitrary b_{i1} and $b_{i2} \in \mathbb{F}_2$, where $1 \leq i \leq n$. Let $z = (z_1, z_2, \ldots, z_n)$ be an arbitrary element of $C(A)^\perp$. Since Tr_1^r is linear over $\mathbb{F}_2$, one can easily verify that

$$\langle \phi_\gamma(z), y \rangle_s = \sum_{i=1}^{n} (b_{i1} \mathrm{Tr}_1^r(\gamma z_i) + b_{i2} \mathrm{Tr}_1^r(z_i)) = \mathrm{Tr}_1^r(\sum_{i=1}^{n} z_i(\gamma b_{i1} + b_{i2})) = \mathrm{Tr}_1^r(z \cdot x). \tag{9}$$

$1 \Rightarrow 2$: Suppose that $y \in \mathscr{C}_\gamma(A)$. Equation (9) implies that $\langle \phi_\gamma(z), y \rangle_s = \mathrm{Tr}_1^r(z \cdot x) = 0$ holds for each z in $C(A)^\perp$. Hence $z \cdot x = 0$ as otherwise we can find $z' \in C(A)^\perp$ such that $0 = \langle \phi_\gamma(z'), y \rangle_s = \mathrm{Tr}_1^r(z' \cdot x) = 1$, which is a contradiction. Hence $x \in C(A)$.

$2 \Rightarrow 1$: Suppose that $x \in C(A)$. Then for each z in $C(A)^\perp$, we have $z \cdot x = 0$. Now, equation (9) implies that $\langle \phi_\gamma(z), y \rangle_s = \mathrm{Tr}_1^r(z \cdot x) = 0$ for each z in $C(A)^\perp$. Hence $y \in \mathscr{C}_\gamma(A)$.

The following minimum distance bounds are direct consequences of this result.

Corollary 1. *Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$ be a defining set of a twisted code $\mathscr{C}_\gamma(A)$ of length n over $\mathbb{F}_2 \times \mathbb{F}_2$. If $\mathscr{C}_\gamma(A)$ contains a weight t vector, then $C(A)$ also contains a weight t vector. In particular, $d(\mathscr{C}_\gamma(A)) \geq d(C(A))$.*

The Hartmann-Tzeng bound (HT bound) is one of classical bounds on the minimum distance of linear cyclic codes [11, Theorem 4.5.6]. A generalization of the HT bound is given in [6]. Based on this generalization we state the following minimum distance lower bound for twisted codes.

Corollary 2. *Let A be a defining set of a twisted code $\mathscr{C}_\gamma(A)$ of length n over $\mathbb{F}_2 \times \mathbb{F}_2$ such that A contains a subset in the form*

$$B = \{(l + i_1 c_1 + i_2 c_2 + \cdots + i_k c_k) \bmod n : \ 0 \leq i_j \leq s_j, \ \gcd(c_j, n) = 1\},$$

where $l, c_j \in \mathbb{Z}/n\mathbb{Z}$ and s_j is a non-negative integer for $1 \leq j \leq k$. Then $d(\mathscr{C}_\gamma(A)) \geq (\sum_{j=1}^{k} s_j) + 2$.

The next theorem provides a sufficient condition for the twisted code $\mathscr{C}_\gamma(A)$ to have minimum distance at least five.

Theorem 5. *Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$ be a symmetric ($A = -A$) complete defining set of a twisted code of length n over $\mathbb{F}_2 \times \mathbb{F}_2$ such that $0 \notin A$. If $d\big(\mathscr{C}_\gamma(A \cup \{0\})\big) \geq 5$, then $\mathscr{C}_\gamma(A)$ has no codeword of weight 4. If in addition $\gcd(n, 3) = 1$, then $d\big(\mathscr{C}_\gamma(A)\big) \geq 5$.*

Proof. We only give a sketch of the proof, and for a complete proof see [6, Theorem 3.7.2]. Using the result of Theorem 4, each weight 4 codeword of $\mathscr{C}_\gamma(A)$ gives rise to a linear equation over $\mathbb{F}_{2^r}$ involving 4 unknowns. Using the symmetric property of A, one can show that these equations have no solution. The same idea holds for weight 2 and 3 codewords of $\mathscr{C}_\gamma(A)$ when $\gcd(n, 3) = 1$.

In general, the choice of γ is a critical factor in the construction of twisted codes. Interestingly, the literature appears to have ignored the impacts of γ on the parameters of twisted codes. The next example shows that choice of γ can improve the code's parameters.

Example 2. Let $n = 69$, $\kappa = 22$, and $A = \{1, -1\} \cup Z(-3)$. Let α be a primitive element of $\mathbb{F}_{2^{22}}$ defined by the `PrimitiveElement` function in Magma [3], which is a root of $\alpha^{22} + \alpha^{12} + \alpha^{11} + \alpha^{10} + \alpha^9 + \alpha^8 + \alpha^6 + \alpha^5 + 1 = 0$. Let $\gamma_1 = \alpha$ and $\gamma_2 = \alpha^{89}$. (Note that γ_2 is not a primitive element of $\mathbb{F}_{2^{22}}$.) The quantum codes $\mathscr{C}_{\gamma_1}(A)$ and $\mathscr{C}_{\gamma_2}(A)$ are $[\![69, 3, 15]\!]$ and $[\![69, 3, 16]\!]$ binary quantum codes, respectively. The latter is a new *record-breaking* binary quantum code.

Through further computations, we find that the minimum distance 16 is attainable by some elements γ with the algebraic degree 22 over $\mathbb{F}_2$ (both primitive and non-primitive). Choosing various primitive elements as values of γ yields twisted codes with minimum distances of 6, 8, 10, 11, 12, 13, 14, 15, and 16 (this list may not be exhaustive). On the other hand, if the algebraic degree of γ is 11 over $\mathbb{F}_2$, then the minimum distance typically tends to be at most 11, occasionally reaching 12. When $\gamma \in \mathbb{F}_4 \setminus \mathbb{F}_2$, then the minimum distance is 11.

We have also developed new theoretical results regarding the selection of the γ value and equivalence of twisted codes in [6, Section 3.10]. Due to the page limit of this extended abstract, these results will be included in the full version of the paper.

5 Infinite classes of binary quantum codes

In this section we give five infinite families of binary quantum codes that produce good (record-breaking or optimal) binary quantum codes. First, the next theorem gives a secondary construction of binary quantum codes that are constructed from twisted codes.

Theorem 6. *Let $A \subseteq \mathbb{Z}/n\mathbb{Z}$ and $\mathscr{C}_\gamma(A)$ be a pure binary quantum code with parameters $[\![n, k, t]\!]$. Then the following results hold.*

(i) If $d(\mathscr{C}_\gamma(\bar{A})) \geq t+1$, where $\bar{A} = A \cup \{0\}$, then there exists an $[\![n+1, k-1, t+1]\!]$ quantum code.

(ii) If $\kappa = 2$ and $\{a, n-a\}$ is a 2-cyclotomic coset such that $d(\mathscr{C}_\gamma(\bar{A})) \geq t+1$ for $\bar{A} = A \cup \{a\}$, then there exists an $[\![n+1, k-1, t+1]\!]$ quantum code.

Proof. Proof of both cases follow from the fact that the code $\mathscr{C}_\gamma(\bar{A})$ has the dual-containment deficiency $e = 1$. Applying Theorem 3 to $\mathscr{C}_\gamma(\bar{A})$ gives the results.

Next, we construct an infinite family of quantum codes with minimum distance of at least four.

Theorem 7. *Let $r > 5$ be an even integer. Then there exists a binary quantum code with parameters $[\![2^r, 2^r - \frac{3}{2}r - 2, d \geq 4]\!]$.*

Proof. Let $n = 2^r - 1$, $\kappa = \frac{r}{2}$, and $A = \{1, a, b\}$, where $a = 2^{\frac{r}{2}} + 1$ and $b = 2^{\frac{r}{2}}$. The code $\mathscr{C}_\gamma(A)$ is a $[\![2^r - 1, 2^r - \frac{3}{2}r - 1, 3]\!]$ pure quantum code. Moreover, $\bar{A} = \{0, 1, a, b\} = \{0, 1\} + \{0, b\}$. The HT bound implies that $d(\mathscr{C}_\gamma(\bar{A})) \geq 4$ and therefore Theorem 6 gives a $[\![2^r, 2^r - \frac{3}{2}r - 2, d \geq 4]\!]$ quantum code.

For instance, if $r = 6$, this construction gives a *record-breaking* $[\![64, 53, 4]\!]$ quantum code Q. Applying the shortening construction given in [10, Theorem 11] or [13, Theorem 3] to the code Q, we get the following *new quantum codes*

$$[\![38, 27, 4]\!], [\![44, 33, 4]\!], [\![46, 35, 4]\!], [\![48, 37, 4]\!], [\![50, 39, 4]\!], [\![56, 45, 4]\!], \tag{10}$$

which all have *better minimum distance* than the previously best-known quantum codes with the same length and dimension. Next, we present two new infinite families of quantum codes with minimum distance of at least five.

Theorem 8. *(i) Let $t = 2^{2k+1}$ for some integer $k \geq 1$ and $n = t^2 + t + 1$. Then there exists an $[\![n, n - 12k - 6, d \geq 5]\!]$ binary quantum code.*

(ii) Let $t = 2^{2k}$ for some integer $k \geq 1$ and $n = t^2 - t + 1$. Then there exists an $[\![n, n - 12k, d \geq 5]\!]$ binary quantum code.

Proof. We use the result of Theorem 5 to prove the minimum distance of both cases is five.
(i) : Let $\kappa = 2k+1$ and $A = \{\pm 1, \pm t, \pm(t+1)\}$. The code $\mathscr{C}_\gamma(A)$ is dual-containing and has parameters $\big(n, 2^{2n-6(2k+1)}, d \geq 5\big)$. Now the quantum construction given in Theorem 1 implies an $[\![n, n - 12k - 6, d \geq 5]\!]$ binary quantum code.
(ii) : Let $\kappa = 2k$ and $A = \{\pm 1, \pm t, \pm(t-1)\}$. The code $\mathscr{C}_\gamma(A)$ is dual-containing and has parameters $\big(n, 2^{2n-12k}, d \geq 5\big)$. Now the quantum construction given in Theorem 1 gives a quantum code with parameters $[\![n, n - 12k, d \geq 5]\!]$.

We construct one optimal as well as two new record-breaking quantum codes with minimum distance of five using the above results.

Example 3. (i) Let $t = 2^2$, $n = t^2 - t + 1 = 13$, and $A = \{\pm 1, \pm 3, \pm 4\}$. The construction given in part (ii) of Theorem 8 gives an optimal quantum code with parameters $[\![13, 1, 5]\!]$.

(*ii*) Let $t = 2^3$, $n = t^2 + t + 1 = 73$, and $A = \{\pm 1, \pm 8, \pm 9\}$. Then Theorem 8 part (*i*) gives a quantum code with parameters $[\![73, 55, 5]\!]$. This code is a *record-breaking quantum code*. Extending the defining set to $A = \{\pm 1, \pm 8, \pm 9, 20, 14, 39\}$ allows to construct a *record-breaking* binary quantum code with parameters $[\![73, 46, 7]\!]$.

(*iii*) Let $t = 2^4$, $n = t^2 - t + 1 = 241$, and $A = \{\pm 1, \pm 15, \pm 16\}$. The construction given in part (*ii*) of Theorem 8 implies a *record-breaking quantum code* with parameters $[\![241, 217, 5]\!]$. Extending the defining set to

$$A = \{\pm 1, \pm 3, \pm 4, \pm 12, \pm 15, \pm 16, \pm 45, \pm 48, \pm 49, \pm 60, \pm 61, \pm 64\}$$

enables us to construct a *record-breaking* binary quantum code with parameters $[\![241, 193, 8]\!]$. Applying the secondary constructions gives 27 other record-breaking binary quantum codes. Moreover, this code can be used to construct 58 other binary quantum codes with missing constructions (red coloured entries in the tables [9]).

Theorem 9. (*i*) *Let $t \geq 4$ be an even integer and $n = 2^t + 1$. Then there exists a pure quantum code with parameters $[\![2^t + 1, 2^t - 2t + 1, d \geq 4]\!]$.*

(*ii*) *Let $t \geq 3$ be an odd integer and $n = 2^t + 1$. Then there exists a pure quantum code with parameters $[\![2^t + 2, 2^t - 2t, d \geq 4]\!]$.*

Proof. (*i*) Let $A = \{1, 2^{\frac{t}{2}}, -1, -2^{\frac{t}{2}}\}$ and $\kappa = \frac{t}{2}$ (note that $\kappa \mid r = 2t$). Since $A = \{1, -2^{\frac{t}{2}}\} + \{0, 2^{\frac{t}{2}} - 1\}$, the HT bound implies $d(\mathscr{C}_\gamma(A)) \geq 4$. Thus the quantum construction given in Theorem 1 implies a binary quantum code with parameters $[\![2^t + 1, 2^t + 1 - 2t, d \geq 4]\!]$.

(*ii*) A similar proof as above by considering $A = \{-1, 1\}$ and $\kappa = t$ gives the result.

The following codes are all obtained from the above construction:

$$[\![10, 2, 4]\!], [\![17, 9, 4]\!], [\![34, 22, 4]\!], [\![65, 53, 4]\!]. \tag{11}$$

The *quantum Hamming bound* [8] states that a pure $[\![n, k, d]\!]$ binary quantum code with $e = \lfloor \frac{d-1}{2} \rfloor$ satisfies

$$\sum_{j=0}^{e} \binom{n}{j} 3^j \leq 2^{n-k}. \tag{12}$$

Recently it has been shown that each quantum code (pure or impure) with minimum distance $d < 127$ satisfies the quantum Hamming bound [5]. Let Q be an $[\![n, k, d]\!]$ binary quantum code. The code Q will be called *optimal*, if there is no $[\![n, k, d']\!]$ binary quantum code with $d' > d$. Next, we prove that some of our quantum codes are optimal.

Theorem 10. *All quantum codes in (10) and (11) along with the code $[\![64, 53, 4]\!]$ are optimal quantum codes.*

Proof. The proof follows from the fact that improving the minimum distance in each of the mentioned quantum codes implies a contradiction with the Hamming bound of (12).

Bibliography

[1] J. Bierbrauer. *Introduction to coding theory*. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2005.

[2] J. Bierbrauer and Y. Edel. Quantum twisted codes. *J. Combin. Des.*, 8(3):174–188, 2000.

[3] W. Bosma, J. Cannon, and C. Playoust. The Magma algebra system I: The user language. *Journal of Symbolic Computation*, 24(3-4):235–265, 1997.

[4] A. R. Calderbank, E. M. Rains, P. Shor, and N. J. Sloane. Quantum error correction via codes over GF(4). *IEEE Transactions on Information Theory*, 44(4):1369–1387, 1998.

[5] E. Dallas, F. Andreadakis, and D. Lidar. No $((n, k, d < 127))$ code can violate the quantum Hamming bound. *IEEE BITS the Information Theory Magazine*, 2023.

[6] R. Dastbasteh. New quantum codes, minimum distance bounds, and equivalence of codes. *PhD Thesis, Simon Fraser University*, 2023, https://summit.sfu.ca/item/36338.

[7] Y. Edel and J. Bierbrauer. Twisted BCH-codes. *J. Combin. Des.*, 5(5):377–389, 1997.

[8] D. Gottesman. Class of quantum error-correcting codes saturating the quantum Hamming bound. *Physical Review A*, 54(3):1862, 1996.

[9] M. Grassl. Code Tables: Bounds on the parameters of various types of codes. http://www.codetables.de/.

[10] M. Grassl, T. Beth, and M. Rötteler. On optimal quantum codes. *International Journal of Quantum Information*, 2(01):55–64, 2004.

[11] W. C. Huffman and V. Pless. *Fundamentals of error-correcting codes*. Cambridge University Press, Cambridge, 2003.

[12] P. Lisoněk and V. Singh. Quantum codes from nearly self-orthogonal quaternary linear codes. *Des. Codes Cryptogr.*, 73(2):417–424, 2014.

[13] E. M. Rains. Nonbinary quantum codes. *IEEE Trans. Inform. Theory*, 45(6):1827–1832, 1999.

Further Results on Orbits and Incidence matrices for the Class $\mathcal{O}_6$ of Lines External to the Twisted Cubic in PG$(3,q)$

Alexander A. Davydov[1][0000−0002−5827−4560], Stefano Marcugini[2][0000−0002−7961−0260], and Fernanda Pambianco[2][0000−0001−5476−5365]

[1] Kharkevich Institute for Information Transmission Problems Russian Academy of Sciences, Moscow, 127051, Russian Federation alexander.davydov121@gmail.com
[2] Department of Mathematics and Computer Science, Perugia University, Perugia, 06123, Italy {stefano.marcugini,fernanda.pambianco}@unipg.it

Abstract. In the literature, lines of the projective space PG$(3,q)$ are partitioned into classes, each of which is a union of line orbits under the stabilizer group of the twisted cubic. The least studied class is named $\mathcal{O}_6$. This class contains lines external to the twisted cubic which are not its chords or axes and do not lie in any of its osculating planes. For even and odd q, we propose a new family of orbits of $\mathcal{O}_6$ and investigate in detail their stabilizer groups and the corresponding submatrices of the point-line and plane-line incidence matrices. To obtain these submatrices, we explored the number of solutions of cubic and quartic equations connected with intersections of lines (including the tangents to the twisted cubic), points, and planes in PG$(3,q)$.

Keywords: Twisted cubic · Projective space · Incidence matrix · Orbits of lines.

1 Introduction

In the three-dimensional projective space PG$(3,q)$ over a Galois field $\mathbb{F}_q$ with q elements, the normal rational curve $\mathscr{C}$, named twisted cubic, has as many as $q+1$ points. Up to a change of the projective frame of PG$(3,q)$, these points are $P_t=(t^3,t^2,t,1)$, $t\in\mathbb{F}_q$, together with $P_\infty=(1,0,0,0)$. In particular, they form a complete $(q+1)$-arc in PG$(3,q)$. The twisted cubic has many interesting properties and is connected with distinct combinatorial and applied problems, which led this curve to be widely studied, see for instance [3–5, 7, 15, 17, 19]. A novel application of twisted cubic aimed at the construction of covering codes has been the motivation for the study of certain submatrices of the point-plane incidence matrix of PG$(3,q)$ arising from the action of the stabilizer group $G_q\cong$ PGL$(2,q)$ of $\mathscr{C}$ in PG$(3,q)$. The investigation, based on the known classification of the point and plane orbits of G_q given in [17], was initiated by D. Bartoli and the present authors in 2020 [1] and produced optimal multiple covering codes.

The results in [1] were also an important ingredient to classify the cosets of the $[q+1, q-3, 5]_q$ generalized doubly-extended Reed-Solomon code of codimension 4 by means of their weight distributions [8].

For the study of the plane-line and the point-line incidence matrices, an explicit description of line orbits is necessary. In [17], a partition of the lines in $\mathrm{PG}(3,q)$ into classes is given, each of which is a union of line orbits under G_q. Apart from one class denoted by $\mathcal{O}_6$, containing lines external to the twisted cubic that are not its chords or axes and do not lie in its osculating planes, the number and the structure of the orbits forming those unions are simultaneously and independently obtained by distinct methods in [11] (for all $q \geq 2$), [2] (for all $q \geq 23$), and [16] (for finite fields of characteristic > 3); see also the references therein.

The results of [2,11,16] are collected in [9, Sect. 2.2, Tab. 1] where texts from arXiv.org corresponding to [2, 11, 16] are used.

Using the representation of the line orbits in [11], for all $q \geq 2$ and apart from the lines in class $\mathcal{O}_6$, the *point-line* and *plane-line* incidence matrices of $\mathrm{PG}(3,q)$ are obtained in [9,10].

In [16], apart from the lines in class $\mathcal{O}_6$, for odd $q \not\equiv 0 \pmod 3$, the numbers of distinct planes through distinct lines (called "the plane orbit distribution of a line") and the numbers of distinct points lying on distinct lines (called "the point orbit distribution of a line") of $\mathrm{PG}(3,q)$ are obtained. For finite fields of characteristic > 3, the results of [16] on "the plane orbit distribution of a line" and "the point orbit distribution of a line" are in accordance with those from [9,10] on the point-line and the plane-line incidence matrices.

In [11] stabilizer groups for the considered orbits are obtained and described in detail whereas in [2,16] the stabilizer groups are not considered. Also, in [11, Conj. 8.2], for the all fields $\mathbb{F}_q$, $q \geq 5$, the detailed conjecture on the sizes and the number of line orbits in the class $\mathcal{O}_6$ is formulated; for $5 \leq q \leq 37$ and $q = 64$ the conjecture has been proved by an exhaustive computer search; see [11, Th. 8.1].

In $\mathrm{PG}(3,q)$, for $q = 2^n$, $n \geq 3$, the $(q+1)$-arc $\mathcal{A} = \{(1, t, t^{2^h}, t^{2^h+1}) \,|\, t \in \mathbb{F}_q^+\}$, $\mathbb{F}_q^+ = \mathbb{F}_q \cup \{\infty\}$, with $gcd(n,h) = 1$ (twisted cubic for $h = 1$), has been considered in a recent paper [6], where it is shown that the orbits of points and of planes under the projective stabilizer G_h of $\mathcal{A}$ are similar to those under G_1 described in [17]; moreover, the point-plane incidence matrix with respect to G_h-orbits mirrors the case h=1 described in [1]. In [6], our conjecture of [11, Conj. 8.2] has also been proved for even q.

In [12], for all even and odd q, a so-called family $\mathscr{O}_\mu$ of line orbits of the class $\mathcal{O}_6$ is obtained using a line family, called ℓ_μ-*lines*, where a parameter μ runs over $\mathbb{F}_q^* \setminus \{1\}$, $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$. Also, one more orbit $\mathscr{O}_\mathcal{L}$, based on a line $\mathcal{L}$ with another description, is given. The orbits $\mathscr{O}_\mu$ and $\mathscr{O}_\mathcal{L}$ are based on the analysis of the stabilizer groups of the corresponding lines. These orbits include an essential part of all $\mathcal{O}_6$ orbits; e.g. they include about one-half and one-third of all lines of $\mathcal{O}_6$ for q even and for $q \equiv 0 \pmod 3$, respectively.

In [13], using the properties of the orbits $\mathscr{O}_\mu$ and $\mathscr{O}_{\mathcal{L}}$ from [12], we determine all the plane-line and point-line incidence matrices connected with these orbits.

In a quite recent paper [18], for a field $\mathbb{F}_q$ of characteristic > 3, our conjecture of [11, Conj. 8.2] on the sizes and the number of orbits in the class $\mathcal{O}_6$ has been proved. For it, the open problem of classifying binary quartic forms over $\mathbb{F}_q$ into G_q-orbits is solved and used. Also, the Plücker embedding for the Klein quadric is applied. So, in [18], the methods and approaches are different with our published articles [9–13] and with this paper. Also, in [18], unlike this paper, the incidence matrices are not considered and the stabilizer groups on language of $\mathrm{PG}(3,q)$ are not presented.

In this paper, we continue and develop approaches of [12, 13]. We propose a new family of lines $\mathscr{L}_\rho$, where ρ is a parameter running over $\mathbb{F}_q^*$. This family is a generalization of the line $\mathcal{L}$ from [12]. For even and odd $q \not\equiv 0 \pmod 3$ the lines $\mathscr{L}_\rho$ belong to the class $\mathcal{O}_6$. The detailed investigation of the stabilizer groups of the lines $\mathscr{L}_\rho$ for all q and ρ allows us to calculate the sizes of the orbits under G_q, containing the lines $\mathscr{L}_\rho$. Also, the parameters of the point-line and plane-line incidence submatrices are being considered with the help of the research of the number of solutions of cubic and quartic equations connected with intersections of the lines $\mathscr{L}_\rho$ and distinct planes, points, and the tangents to the twisted cubic. Formulas for the numbers of the solutions are being described. It is shown when the lines $\mathscr{L}_\rho$ generate new orbits in comparison with the orbits $\mathscr{O}_\mu$ [12].

The detailed version of this paper with proofs of lemmas and theorems can be found in [14].

2 Preliminaries

2.1 Twisted cubic in $\mathbf{PG}(3,q)$

Here we cite some results from [17] useful in this paper.

Let $\boldsymbol{\pi}(c_0,c_1,c_2,c_3)$ be the plane of $\mathrm{PG}(3,q)$ with equation $c_0x_0+c_1x_1+c_2x_2+c_3x_3=0$, $c_i \in \mathbb{F}_q$. Let $\mathbf{P}(x_0,x_1,x_2,x_3)$ be a point of $\mathrm{PG}(3,q)$ with homogeneous coordinates $x_i \in \mathbb{F}_q$. Let $\mathscr{C} \subset \mathrm{PG}(3,q)$ be the *twisted cubic* consisting of $q+1$ points no 4 of which are coplanar. We consider $\mathscr{C}$ in the canonical form

$$\mathscr{C} = \{P(t) = \mathbf{P}(t^3,t^2,t,1) \,|\, t \in \mathbb{F}_q^+\}.$$

A *chord* of $\mathscr{C}$ through the points $P(t_1)$ and $P(t_2)$ is a line joining either a pair of real points of $\mathscr{C}$, possibly coincident, or a pair of complex conjugate points. If $t_1 = t_2 = t$ then it is the *tangent* $\mathcal{T}_t$ to $\mathscr{C}$ at the point $P(t)$.

In the point $P(t) \in \mathscr{C}$, we have the *osculating plane*

$$\pi_{\mathrm{osc}}(t) = \boldsymbol{\pi}(1,-3t,3t^2,-t^3),\ t \in \mathbb{F}_q^+.$$

For $q \not\equiv 0 \pmod 3$, the osculating planes form the *cubic developable* Γ to $\mathscr{C}$.

For $q \not\equiv 0 \pmod 3$, the *null polarity* $\mathfrak{A}$ [17, Th. 21.1.2] is given by

$$\mathbf{P}(x_0,x_1,x_2,x_3)\mathfrak{A} = \boldsymbol{\pi}(x_3,-3x_2,3x_1,-x_0).$$

For $q \not\equiv 0 \pmod 3$, an *axis* of Γ is a line of $\mathrm{PG}(3,q)$ which is the intersection of a pair of real or complex conjugate planes of Γ.

Notation 1 *Throughout the paper, we consider* $q \equiv \xi \pmod 3$, $\xi \in \{-1, 0, 1\}$. *The following notation is used for* $\xi \neq 0$.

G_q	*the group of projectivities in* $\mathrm{PG}(3, q)$ *fixing* $\mathscr{C}$;
$\#S$	*the cardinality of a set* S;
$\overline{AB}$	*the line through the points* A *and* B;
$\triangleq$	*the sign "equality by definition".*
Γ*-plane*	*an osculating plane of* Γ;
$d_{\mathscr{C}}$*-plane*	*a plane containing* exactly $d_{\mathscr{C}}$ *distinct points of* $\mathscr{C}$, $d_{\mathscr{C}} \in \{0, 2, 3\}$;
$\overline{1_{\mathscr{C}}}$*-plane*	*a plane not in* Γ *containing* exactly 1 *point of* $\mathscr{C}$;
$\mathfrak{P}$	*the list of possible types* π *of planes,* $\mathfrak{P} \triangleq \{\Gamma, 2_{\mathscr{C}}, 3_{\mathscr{C}}, \overline{1_{\mathscr{C}}}, 0_{\mathscr{C}}\}$;
π*-plane*	*a plane of the type* $\pi \in \mathfrak{P}$.
$\mathscr{C}$*-point*	*a point of* $\mathscr{C}$;
μ_Γ*-point*	*a point off* $\mathscr{C}$ *lying on* exactly μ *distinct osculating planes,* $\mu_\Gamma \in \{0_\Gamma, 1_\Gamma, 3_\Gamma\}$;
T*-point*	*a point off* $\mathscr{C}$ *on a tangent to* $\mathscr{C}$;
$\mathfrak{M}$	*the list of possible types* $\mathfrak{p}$ *of points,* $\mathfrak{M} \triangleq \{\mathscr{C}, 0_\Gamma, 1_\Gamma, 3_\Gamma, \mathrm{T}\}$;
$\mathfrak{p}$*-point*	*a point of the type* $\mathfrak{p} \in \mathfrak{M}$.
$\mathscr{N}_\pi$	*the orbit of* π*-planes under* G_q, $\pi \in \mathfrak{P}$;
$\mathscr{M}_\mathfrak{p}$	*the orbit of* $\mathfrak{p}$*-points under* G_q, $\mathfrak{p} \in \mathfrak{M}$;
EnΓ*-line*	*a line, external to the cubic* $\mathscr{C}$*, not in a* Γ*-plane, that is neither a chord nor an axis, see* [17, Lem. 21.1.4] *and its context;*
$\mathcal{O}_6 = \mathcal{O}_{\mathrm{En}\Gamma}$	*the union (class) of all orbits of* EnΓ*-lines.*

Theorem 1. [17, Ch. 21] *The following properties of the twisted cubic* $\mathscr{C}$ *hold:*

- *A matrix* $\mathbf{M}$ *corresponding to a projectivity of* G_q *has the general form*

$$\mathbf{M} = \begin{bmatrix} a^3 & a^2c & ac^2 & c^3 \\ 3a^2b & a^2d + 2abc & bc^2 + 2acd & 3c^2d \\ 3ab^2 & b^2c + 2abd & ad^2 + 2bcd & 3cd^2 \\ b^3 & b^2d & bd^2 & d^3 \end{bmatrix}, \ a, b, c, d \in \mathbb{F}_q, \ ad - bc \neq 0. \quad (2.1)$$

- *Under* G_q, $q \geq 5$, *there are the following five orbits* $\mathscr{N}_\pi$ *of planes:*

$\mathscr{N}_1 = \mathscr{N}_\Gamma = \{\Gamma$*-planes*$\}$, $\mathscr{N}_2 = \mathscr{N}_{2_\mathscr{C}} = \{2_\mathscr{C}$*-planes*$\}$, $\mathscr{N}_3 = \mathscr{N}_{3_\mathscr{C}} = \{3_\mathscr{C}$*-planes*$\}$,
$\mathscr{N}_4 = \mathscr{N}_{\overline{1_\mathscr{C}}} = \{\overline{1_\mathscr{C}}$*-planes*$\}$, $\mathscr{N}_5 = \mathscr{N}_{0_\mathscr{C}} = \{0_\mathscr{C}$*-planes*$\}$.

- *For* $q \not\equiv 0 \pmod 3$, *under* G_q, *there are five orbits* $\mathscr{M}_j$ *of points:*

$\mathscr{M}_1 = \mathscr{M}_\mathscr{C} = \{\mathscr{C}$*-points*$\}$, $\mathscr{M}_2 = \mathscr{M}_\mathrm{T} = \{$T*-points*$\}$, $\mathscr{M}_3 = \mathscr{M}_{3_\Gamma} = \{3_\Gamma$*-points*$\}$,

$\mathscr{M}_4 = \mathscr{M}_{1_\Gamma} = \{1_\Gamma\text{-}points\},\ \mathscr{M}_5 = \mathscr{M}_{0_\Gamma} = \{0_\Gamma\text{-}points\};\ \mathscr{M}_j\mathfrak{A} = \mathscr{N}_j, j = 1, \dots, 5.$

• *The lines of* $\mathrm{PG}(3,q)$ *can be partitioned into classes called* $\mathcal{O}_i$ *and* $\mathcal{O}'_i = \mathcal{O}_i\mathfrak{A}$*, each of which is a union of orbits under* G_q*. In particular, for all* q*, there is the class* $\mathcal{O}_6 = \mathcal{O}_{\mathrm{En}\Gamma} = \{\mathrm{En}\Gamma\text{-}lines\}$*,* $\#\mathcal{O}_6 = (q^2-q)(q^2-1)$*,* $\mathcal{O}_6 = \mathcal{O}'_6 = \mathcal{O}_6\mathfrak{A}$*.*

2.2 An EnΓ-line $\mathcal{L}$ and its orbit $\mathscr{O}_\mathcal{L}$, $q \not\equiv 0 \pmod 3$

Here we present some results from [12] useful in this paper.
Let $Q_\beta = \mathbf{P}(1,0,\beta,1),\ \beta \in \mathbb{F}_q^+$, be a point of $\mathrm{PG}(3,q)$. We consider the EnΓ-line

$$\mathcal{L} = \overline{Q_0Q_\infty} = \overline{\mathbf{P}(1,0,0,1)\mathbf{P}(0,0,1,0)} = \{\mathbf{P}(1,0,\beta,1) \mid \beta \in \mathbb{F}_q^+\}. \tag{2.2}$$

Theorem 2. [12, Sect. 3] *Let* $\mathscr{O}_\mathcal{L}$ *be the orbit of the line* $\mathcal{L}$ *under* G_q*. Let* $q \equiv \xi \pmod 3$*,* $\xi \neq 0$*. The orbit* $\mathscr{O}_\mathcal{L}$ *has size*

$$\#\mathscr{O}_\mathcal{L} = \begin{cases} (q^3-q)/3 & \text{if } \xi = 1,\ q \text{ is even or } 2 \text{ is a non-cube in } \mathbb{F}_q; \\ (q^3-q)/12 & \text{if } \xi = 1,\ q \text{ is odd and } 2 \text{ is a cube in } \mathbb{F}_q; \\ q^3-q & \text{if } \xi = -1,\ q \text{ is even}; \\ (q^3-q)/2 & \text{if } \xi = -1,\ q \text{ is odd}. \end{cases}$$

2.3 A family of EnΓ-lines ℓ_μ and their orbits $\mathscr{O}_\mu$

Here we cite some results from [12,13] useful in this paper.

Let $q \equiv \xi \pmod 3$. Let $\mu \in \mathbb{F}_q^* \setminus \{1\}$ if q is even or $\xi = 0$; $\mu \in \mathbb{F}_q^* \setminus \{1, 1/9\}$ if q is odd, $\xi \neq 0$. Let $R_{\mu,\gamma} = \mathbf{P}(\gamma,\mu,\gamma,1)$, $\gamma \in \mathbb{F}_q^+$, be a point of $\mathrm{PG}(3,q)$. We consider the EnΓ-line

$$\ell_\mu = \overline{R_{\mu,0}R_{\mu,\infty}} = \overline{\mathbf{P}(0,\mu,0,1)\mathbf{P}(1,0,1,0)} = \{\mathbf{P}(\gamma,\mu,\gamma,1) | \gamma \in \mathbb{F}_q^+,\ \mu \text{ is fixed}\}.$$

Theorem 3. [12, Sects. 4–7], [13, Sects. 5–7] *Let* $q \equiv \xi \pmod 3$*.*
• *Let* $\mathscr{O}_\mu$ *be the orbit of the line* ℓ_μ *under* G_q*. Let the condition* $\Upsilon_{q,\mu}$ *have the form* $\mu = -1/3,\ q \equiv 1 \pmod{12}$*,* $-1/3$ *is a fourth power. The orbit* $\mathscr{O}_\mu$ *has size*

$$\#\mathscr{O}_\mu = \begin{cases} (q^3-q)/2 & \text{if } q \text{ is even or } \mu \text{ is a non-square in } \mathbb{F}_q; \\ (q^3-q)/4 & \text{if } \mu \text{ is a square in } \mathbb{F}_q \text{ and } \xi = 0; \\ (q^3-q)/4 & \text{if } q \text{ is odd}, \mu \text{ is a square in } \mathbb{F}_q, \xi \neq 0, \Upsilon_{q,\mu} \text{ does not hold}; \\ (q^3-q)/12 & \text{if } q \text{ is odd},\ \xi \neq 0, \Upsilon_{q,\mu} \text{ holds}; \end{cases}$$

• *Let* q *be odd,* $\xi \neq 0$*. Every line of* $\mathscr{O}_\mu$ *contains* $\mathfrak{n}_q(\mu) \in \{0,2,4\}$ T*-points;*
• *Let* q *be even. Then every line of the orbit* $\mathscr{O}_\mu$ *contains* $\mathfrak{n}_q(\mu) = 2$ T*-points.*

3 The description of EnΓ-lines $\mathscr{L}_\rho$. Useful relations

Let $K_{\rho,\gamma} = \mathbf{P}(\rho,0,\gamma,1),\ \gamma \in \mathbb{F}_q^+,\ \rho \in \mathbb{F}_q^*$ be a point of $\mathrm{PG}(3,q)$, $K_{\rho,0} = \mathbf{P}(\rho,0,0,1),\ K_{\rho,\infty} = \mathbf{P}(0,0,1,0)$. We introduce the line $\mathscr{L}_\rho$ through $K_{\rho,0}, K_{\rho,\infty}$.

$$\mathscr{L}_\rho = \overline{\mathbf{P}(\rho,0,0,1)\mathbf{P}(0,0,1,0)} = \{\mathbf{P}(\rho,0,\gamma,1) | \gamma \in \mathbb{F}_q^+,\ \rho \text{ is fixed}\}. \tag{3.1}$$

By (3.1), the coordinate vector L_ρ of $\mathscr{L}_\rho$ is $L_\rho = (0, \rho, 0, 0, 0, -1)$.

If $\rho = 0$ the line $\mathscr{L}_0$ is the tangent $\mathcal{T}_0$. This explains why we consider $\rho \in \mathbb{F}_q^*$.

By (2.2), (3.1), the line $\mathcal{L}$ of [12] is the line $\mathscr{L}_1$.

Lemma 1. *We have* $\mathscr{L}_\rho = \mathscr{L}_\rho \mathfrak{A}$.

Lemma 2. • *For* $q \not\equiv 0 \pmod 3$, *the line* $\mathscr{L}_\rho$ *is an* EnΓ*-line.*

• *For* $q \equiv 0 \pmod 3$, *the line* $\mathscr{L}_\rho$ *is not an* EnΓ*-line.*

From now on, we consider the lines $\mathscr{L}_\rho$ *for* $q \not\equiv 0 \pmod 3$.

We fix a primitive element α of the field $\mathbb{F}_q$. The discrete logarithm log of $\beta \in \mathbb{F}_q^*$ is the integer $b \in [0, \dots, q-1]$ such that $\alpha^b = \beta$. Let $\mathfrak{R}_m$, $m = 0, 1, 2$, be a class of the values of ρ such that $\mathfrak{R}_m \triangleq \{\rho \in \mathbb{F}_q^* \mid \log \rho \equiv m \pmod 3\}$.

Lemma 3. *Let* $q \equiv 1 \pmod 3$. *Then* $\beta \in \mathbb{F}_q^*$ *is a cube if and only if* $\log \beta \equiv 0 \pmod 3$.

For $\beta \in \mathbb{F}_q^*$, we define the quadratic character η of $\mathbb{F}_q^*$ as follows:
$\eta(\beta) = 1$ if β is a square of an element in $\mathbb{F}_q^*$; $\eta(\beta) = -1$ otherwise.

Notation 2 *The following notation, where* ρ *will be clear by the context, is used:*

$\mathbb{O}_\rho$	*the orbit under* G_q *generated by the line* $\mathscr{L}_\rho$;
Π_π	*the number of* π*-planes through a line from the orbit* $\mathbb{O}_\rho$, $\pi \in \mathfrak{P}$;
Λ_π	*the number of lines from the orbit* $\mathbb{O}_\rho$ *in a* π*-plane*, $\pi \in \mathfrak{P}$;
$\mathbb{P}_\mathfrak{p}$	*the number of* $\mathfrak{p}$*-points on a line from the orbit* $\mathbb{O}_\rho$, $\mathfrak{p} \in \mathfrak{M}$;
$\mathbb{L}_\mathfrak{p}$	*the number of lines from the orbit* $\mathbb{O}_\rho$ *through a* $\mathfrak{p}$*-point*, $\mathfrak{p} \in \mathfrak{M}$.

Proposition 1. *For the orbit* $\mathbb{O}_\rho$, *generated by a line* $\mathscr{L}_\rho$, *we have*

$$\Pi_\Gamma = \Lambda_\Gamma = \mathbb{P}_\mathscr{C} = \mathbb{L}_\mathscr{C} = 0;\ \mathbb{P}_\mathrm{T} = \Pi_{2_\mathscr{C}},\ \mathbb{P}_{0_\Gamma} = \Pi_{0_\mathscr{C}},\ \mathbb{P}_{1_\Gamma} = \Pi_{\overline{1_\mathscr{C}}}, \tag{3.2}$$

$$\mathbb{P}_{3_\Gamma} = \Pi_{3_\mathscr{C}};\ \mathbb{L}_\mathrm{T} = \Lambda_{2_\mathscr{C}},\ \mathbb{L}_{0_\Gamma} = \Lambda_{0_\mathscr{C}},\ \mathbb{L}_{1_\Gamma} = \Lambda_{\overline{1_\mathscr{C}}},\ \mathbb{L}_{3_\Gamma} = \Lambda_{3_\mathscr{C}}. \tag{3.3}$$

Theorem 4. *For orbits* $\mathbb{O}_\rho$, *generated by lines* $\mathscr{L}_\rho$, *the plane-line incidence matrix contains, according to* (3.2), (3.3), *the same values of the point-line incidence matrix, but in this case they refer to* Π_π, Λ_π *instead of* $\mathbb{P}_\mathfrak{p}, \mathbb{L}_\mathfrak{p}$.

4 Intersections of $\mathscr{L}_\rho$-lines and tangents

The mutual invariant [17, Sect. 15.2] of $\mathscr{L}_\rho$ and the tangent $\mathcal{T}_t$ to $\mathscr{C}$ at the point $P(t)$ is $\varpi(\mathscr{L}_\rho, \mathcal{T}_t) = -2\rho t - t^4$, $t \in \mathbb{F}_q$, $\rho \in \mathbb{F}_q^*$; $\varpi(\mathscr{L}_\rho, \mathcal{T}_\infty) = -1$. The lines $\mathscr{L}_\rho$ and $\mathcal{T}_t$ intersect if and only if $\varpi(\mathscr{L}_\rho, \mathcal{T}_t) = 0$, i.e. $t^4 + 2\rho t = 0$; we denote

$$\mathfrak{n}_q(\rho) \triangleq \#\{t \,|\, t^4 + 2\rho t = 0,\ t \in \mathbb{F}_q,\ \rho \in \mathbb{F}_q^*,\ q \not\equiv 0 \pmod 3\}. \tag{4.1}$$

Theorem 5. (i) *The value* $\mathfrak{n}_q(\rho)$ *is the number of* T*-points on the* $\mathscr{L}_\rho$*-line. Let* $\mathfrak{n}_q(\mu)$ *be as in Theorem* 3. *If* $\mathfrak{n}_q(\rho) \neq \mathfrak{n}_q(\mu)$ *then the orbits* $\mathbb{O}_\rho$ *and* $\mathscr{O}_\mu$ *are distinct.*

(ii) *Let q be even or q be odd, $q \not\equiv 0 \pmod 3$, and -2ρ be a non-cube in $\mathbb{F}_q$. Then the unique solution of $\varpi(\mathscr{L}_\rho, \mathcal{T}_t) = 0$ is $t = 0$, i.e. $\mathfrak{n}_q(\rho) = 1$ and every orbit $\mathbb{O}_\rho$ is different from any orbit $\mathscr{O}_\mu$.*

(iii) *Let q be odd, $q \equiv \xi \pmod 3$, $\xi \neq 0$. Let also -2ρ be a cube in $\mathbb{F}_q$. Then the roots of the equation $\varpi(\mathscr{L}_\rho, \mathcal{T}_t) = 0$ are $t = 0,\ t = \sqrt[3]{-2\rho}$.*

5 Stabilizers of $\mathscr{L}_\rho$-lines and sizes of orbits $\mathbb{O}_\rho$

We denote by G_q^∞ the subgroup of G_q fixing the point $K_{\rho,\infty} = \mathbf{P}(0,0,1,0)$. Let $\mathbf{M}^\infty$ be the matrix corresponding to a projectivity of G_q^∞.

Lemma 4. *The general form of the matrix $\mathbf{M}^\infty$ is as follows:*

$$\mathbf{M}^\infty = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & 0 & d^2 & 0 \\ 0 & 0 & 0 & d^3 \end{bmatrix},\ d \in \mathbb{F}_q^*. \tag{5.1}$$

We are looking for the stabilizer group of $\mathscr{L}_\rho$ and its orbit under G_q. Let $G_q^{\mathscr{L}_\rho}$ be the subgroup of G_q fixing $\mathscr{L}_\rho$. Let $\mathbf{M}^{\mathscr{L}_\rho}$ be the matrix corresponding to a projectivity of $G_q^{\mathscr{L}_\rho}$.

Lemma 5. *Let q be even or let -2ρ be a non-cube in $\mathbb{F}_q$. Then the general form of the matrix $\mathbf{M}^{\mathscr{L}_\rho}$ corresponding to a projectivity of $G_q^{\mathscr{L}_\rho}$ is as in* (5.1) *with d a cubic root of unity.*

Lemma 6. *Let $q \equiv -1 \pmod 3$. Then all $\mathscr{L}_\rho$ lines belong to the same orbit $\mathbb{O}_1$.*

Lemma 7. *Let $q \equiv -1 \pmod 3$, q odd. Then $G_q^{\mathscr{L}_\rho}$ has order 2 and the matrix $\mathbf{M}^{\mathscr{L}_\rho}$ corresponding to the non-trivial projectivity of $G_q^{\mathscr{L}_\rho}$ has the form* (2.1) *with $a = \sqrt[3]{1/2\rho},\ b = 1,\ c = \sqrt[3]{2/\rho^2},\ d = -\sqrt[3]{1/2\rho}$.*

Lemma 8. *Let $q \equiv 1 \pmod 3$, q odd, and let -2ρ be a cube in $\mathbb{F}_q$. Then $G_q^{\mathscr{L}_\rho}$ has order 12 and is isomorphic to the group $\mathbf{A}_4$. A matrix $\mathbf{M}^{\mathscr{L}_\rho}$ of $G_q^{\mathscr{L}_\rho}$ either has the form as in* (5.1) *with d a cubic root of unity, or has the form* (2.1) *with $a = \sqrt[3]{1/2\rho},\ b = 1,\ c = -d/\rho a^2,\ d = \sqrt[3]{-1/2\rho}$.*

Theorem 6. (i) *Let $q \equiv 1 \pmod 3$. Let q be even or let -2ρ be a non-cube in $\mathbb{F}_q$. Then the size of the subgroup $G_q^{\mathscr{L}_\rho}$ of G_q fixing the $\mathrm{En}\Gamma$-line $\mathscr{L}_\rho$ is $\#G_q^{\mathscr{L}_\rho} = 3$. The size of the orbit of $\mathscr{L}_\rho$ under G_q is equal to $(q^3 - q)/3$.*

(ii) *Let $q \equiv 1 \pmod 3$. Let q be odd and let -2ρ be a cube in $\mathbb{F}_q$. Then the size of the subgroup $G_q^{\mathscr{L}_\rho}$ of G_q fixing the $\mathrm{En}\Gamma$-line $\mathscr{L}_\rho$ is $\#G_q^{\mathscr{L}_\rho} = 12$ and $G_q^{\mathscr{L}_\rho} \cong \mathbf{A}_4$. The size of the orbit of $\mathscr{L}_\rho$ under G_q is equal to $(q^3 - q)/12$.*

(iii) *Let $q \equiv -1 \pmod 3$. Let q be even. Then $\#G_q^{\mathscr{L}_\rho} = 1$ and the size of the orbit of $\mathscr{L}_\rho$ under G_q is equal to $q^3 - q$.*

(iv) *Let $q \equiv -1 \pmod 3$. Let q be odd. Then $\#G_q^{\mathscr{L}_\rho} = 2$ and the size of the orbit of $\mathscr{L}_\rho$ under G_q is equal to $(q^3 - q)/2$.*

6 A cubic equation, incidence matrices and orbits, even q

Lemma 9. *Let q be even. Let $\gamma, t \in \mathbb{F}_q$. Let the point $K_{\rho,\gamma} = \mathbf{P}(\rho, 0, \gamma, 1)$ belong to the osculating plane $\pi_{osc}(t)$. Then the values of ρ, γ, t satisfy the cubic equation*

$$\widetilde{F}_{\rho,\gamma}(t) = t^3 + \gamma t^2 + \rho = 0,\ \gamma \in \mathbb{F}_q,\ \rho \in \mathbb{F}_q^*,\ q \text{ is even.} \tag{6.1}$$

We denote $\widetilde{\mathbb{N}}_m(\rho)$ the number of γ such that the equation $\widetilde{F}_{\rho,\gamma}(t)$ has exactly m distinct solutions t in $\mathbb{F}_q$, $m = 0, 1, 2, 3$.

Let $\mathrm{Tr}_2(\beta)$ be the absolute trace of an element $\beta \in \mathbb{F}_q$, q even. We denote

$$\widetilde{\mathbb{W}}_q(\rho) \triangleq \#\left\{\gamma \mid \mathrm{Tr}_2\left(\frac{\gamma^3}{\rho} + 1\right) = 1,\ \gamma \in \mathbb{F}_q,\ q = 2^c,\ \rho \in \mathbb{F}_q^* \text{ is fixed}\right\}. \tag{6.2}$$

Lemma 10. • *Let $q = 2^{2m-1}$, $m \geq 2$. Then $\widetilde{\mathbb{N}}_1(\rho) = \widetilde{\mathbb{W}}_{2^{2m-1}}(\rho) = q/2$.*
• *Let $q = 2^{2m} \equiv 1 \pmod 3$, $m \geq 2$. We have*

$$\widetilde{\mathbb{N}}_1(\rho) = \widetilde{\mathbb{W}}_{2^{2m}}(\rho) = \begin{cases} 2^{2m-1} + (-2)^m & \text{if } \rho \text{ is a cube in } \mathbb{F}_q \\ 2^{2m-1} + (-2)^{m-1} & \text{if } \rho \text{ is a non-cube in } \mathbb{F}_q \end{cases}. \tag{6.3}$$

Theorem 7. *Let q be even. For the orbit $\mathbb{O}_\rho$, generated by a line $\mathscr{L}_\rho$, we have*

$$\mathbb{P}_{\mathrm{T}} = 1,\ \mathbb{P}_{0_\Gamma} = \widetilde{\mathbb{N}}_0(\rho),\ \mathbb{P}_{1_\Gamma} = \widetilde{\mathbb{N}}_1(\rho) = \widetilde{\mathbb{W}}_q(\rho),\ \mathbb{P}_{3_\Gamma} = \widetilde{\mathbb{N}}_3(\rho). \tag{6.4}$$

Theorem 8. *Let q be even. Let the orbit $\mathbb{O}_\rho$ be generated by a line $\mathscr{L}_\rho$. Then, for the point-line incidence matrix corresponding to the orbit the following holds:*
• *Let $q = 2^{2m-1}$. Then $\#\mathbb{O}_\rho = q^3 - q$ for all ρ; $\widetilde{\mathbb{W}}_q(\rho) = q/2$, and*

$$\mathbb{P}_{\mathrm{T}} = 1, \mathbb{L}_{\mathrm{T}} = q - 1;\ 2\mathbb{P}_{1_\Gamma} = \mathbb{L}_{1_\Gamma} = q;\ 6\mathbb{P}_{3_\Gamma} = \mathbb{L}_{3_\Gamma} = q - 2;\ 3\mathbb{P}_{0_\Gamma} = \mathbb{L}_{0_\Gamma} = q + 1.$$

• *Let $q = 2^{2m}$. Then $\#\mathbb{O}_\rho = \frac{1}{3}(q^3 - q)$ for all ρ, $\widetilde{\mathbb{W}}_q(\rho)$ is as in* (6.3)*, and*

$$\mathbb{P}_{\mathrm{T}} = 1,\ \mathbb{L}_{\mathrm{T}} = \frac{1}{3}(q-1);\ \mathbb{P}_{1_\Gamma} = \widetilde{\mathbb{W}}_q(\rho),\ \mathbb{L}_{1_\Gamma} = \frac{2}{3}\widetilde{\mathbb{W}}_q(\rho);\ \mathbb{P}_{3_\Gamma} = \frac{q - 1 - \widetilde{\mathbb{W}}_q(\rho)}{3},$$
$$\mathbb{L}_{3_\Gamma} = \frac{2(q - 1 - \widetilde{\mathbb{W}}_q(\rho))}{3};\ \mathbb{P}_{0_\Gamma} = \mathbb{L}_{0_\Gamma} = \frac{2q - 2\widetilde{\mathbb{W}}_q(\rho) + 1}{3}.$$

The plane-line incidence matrix contains, according to (3.2), (3.3)*, the same values of the point-line incidence matrix, but in this case they refer to Π_π, Λ_π instead of $\mathbb{P}_{\mathtt{p}}, \mathbb{L}_{\mathtt{p}}$.*

Theorem 9. *Let q be even. Then all the orbits generated by $\mathscr{L}_\rho$-lines are different from the ones generated by ℓ_μ-lines. Moreover, the following holds:*
• *Let $q = 2^{2m}$. Two lines $\mathscr{L}_{\rho'}$ and $\mathscr{L}_{\rho''}$ belong to different orbits of G_q if and only if $\log \rho' \not\equiv \log \rho'' \pmod 3$. Let α be a primitive element of $\mathbb{F}_q$. There are 3 distinct $\frac{1}{3}(q^3 - q)$-orbits generated by $\mathscr{L}_\rho$-lines with $\rho = \alpha^j$, $j = 0, 1, -1$.*
• *If $q = 2^{2m-1} \equiv -1 \pmod 3$, all $\mathscr{L}_\rho$-lines generate the same $(q^3 - q)$-orbit.*

7 A cubic equation and incidence matrices, odd q

Lemma 11. • *For odd* $q \not\equiv 0 \pmod 3$, *let the point* $K_{\rho,\gamma} = \mathbf{P}(\rho, 0, \gamma, 1)$ *belong to the osculating plane* $\pi_{osc}(t)$, $\gamma, t \in \mathbb{F}_q^*$. *Then* ρ, γ, t *satisfy the equation*

$$F_{\rho,\gamma}(t) = t^3 - 3\gamma t^2 - \rho = 0,\ \gamma, t, \rho \in \mathbb{F}_q^*,\ q \not\equiv 0 \pmod 3. \tag{7.1}$$

• *Let* $q \equiv \xi \pmod 3$. *The equation* $F_{\rho,\gamma}(t)$ *has exactly* 1 *root* $t \in \mathbb{F}_q$ *if and only if* $4\gamma^3 + \rho \neq 0$ *and* $1 + 4\rho^{-1}\gamma^3$ *is a square* (*resp. non-square*) *in* $\mathbb{F}_q$ *for* $\xi = -1$ (*resp.* $\xi = 1$).
• *Let* $\mathbb{N}_1(\rho)$ *be the number of* $\gamma \in \mathbb{F}_q^*$ *such that* $F_{\rho,\gamma}(t)$ *has exactly 1 root* $t \in \mathbb{F}_q^*$.
•• *Let* $q \equiv -1 \pmod 3$ *be odd. Then* $\mathbb{N}_1(\rho) = (q-3)/2$.
•• *Let* $q \equiv 1 \pmod 3$ *be odd. Let* $\eta(\beta)$ *be as in Section 3. Then* $\mathbb{N}_1(\rho) = \mathfrak{N}_{q,\rho}$;

$$\mathfrak{N}_{q,\rho} \triangleq \#\{\gamma \,|\, \gamma \in \mathbb{F}_q^*,\ \eta(1 + 4\rho^{-1}\gamma^3) = -1\},\ q \equiv 1 \pmod 3. \tag{7.2}$$

Theorem 10. *Let* q *be odd. Let* $\mathfrak{N}_{q,\rho}$ *be as in* (7.2). *For the orbit* $\mathbb{O}_\rho$, *generated by a line* $\mathscr{L}_\rho$, *the following holds.*
• $\mathbb{P}_{\mathrm{T}} = 2$ *if* $q \equiv -1 \pmod 3$; $\mathbb{P}_{\mathrm{T}} = 1$ *if* $q \equiv 1 \pmod 3$ *and* -2ρ *is a non-cube in* $\mathbb{F}_q$; $\mathbb{P}_{\mathrm{T}} = 4$ *if* $q \equiv 1 \pmod 3$ *and* -2ρ *is a cube in* $\mathbb{F}_q$.
• $\mathbb{P}_{1_\Gamma} = \mathbb{N}_1(\rho) + 1 = (q-1)/2$ *if* $q \equiv -1 \pmod 3$;
• $\mathbb{P}_{1_\Gamma} = \mathbb{N}_1(\rho) = \mathfrak{N}_{q,\rho}$ *if* $q \equiv 1 \pmod 3$.

Theorem 11. *Let* q *be odd. Let the orbit* $\mathbb{O}_\rho$ *be generated by a line* $\mathscr{L}_\rho$. *Then, for the point-line incidence matrix corresponding to the orbit the following holds:*
• *Let* $q \equiv -1 \pmod 3$. *Then* $\#\mathbb{O}_\rho = (q^3 - q)/2$ *for all* ρ *and we have*

$$\mathbb{P}_{\mathrm{T}} = 2,\ \mathbb{L}_{\mathrm{T}} = q - 1;\ \mathbb{P}_{1_\Gamma} = \mathbb{L}_{1_\Gamma} = \frac{q-1}{2};\ 6\mathbb{P}_{3_\Gamma} = 2\mathbb{L}_{3_\Gamma} = q - 5;$$
$$3\mathbb{P}_{0_\Gamma} = 2\mathbb{L}_{0_\Gamma} = q + 1.$$

• *Let* $q \equiv 1 \pmod 3$. *Let* -2ρ *be a non-cube in* $\mathbb{F}_q$. *Then* $\#\mathbb{O}_\rho = (q^3 - q)/3$;

$$\mathbb{P}_{\mathrm{T}} = 1,\ \mathbb{L}_{\mathrm{T}} = \frac{q-1}{3};\ \mathbb{P}_{1_\Gamma} = \mathfrak{N}_{q,\rho},\ \mathbb{L}_{1_\Gamma} = \frac{2}{3}\mathfrak{N}_{q,\rho};\ \mathbb{P}_{3_\Gamma} = \frac{q - 1 - \mathfrak{N}_{q,\rho}}{3},$$
$$\mathbb{L}_{3_\Gamma} = \frac{2(q - 1 - \mathfrak{N}_{q,\rho})}{3};\ \mathbb{P}_{0_\Gamma} = \mathbb{L}_{0_\Gamma} = \frac{2q + 1 - 2\mathfrak{N}_{q,\rho}}{3}.$$

• *Let* $q \equiv 1 \pmod 3$. *Let* -2ρ *be a cube in* $\mathbb{F}_q$. *Then* $\#\mathbb{O}_\rho = (q^3 - q)/12$ *and*

$$\mathbb{P}_{\mathrm{T}} = 4,\ \mathbb{L}_{\mathrm{T}} = \frac{q-1}{3};\ \mathbb{P}_{1_\Gamma} = \mathfrak{N}_{q,\rho},\ \mathbb{L}_{1_\Gamma} = \frac{1}{6}\mathfrak{N}_{q,\rho};\ \mathbb{P}_{3_\Gamma} = \frac{q - 7 - \mathfrak{N}_{q,\rho}}{3},$$
$$\mathbb{L}_{3_\Gamma} = \frac{q - 7 - \mathfrak{N}_{q,\rho}}{6};\ \mathbb{P}_{0_\Gamma} = \frac{2(q - 1 - \mathfrak{N}_{q,\rho})}{3},\ \mathbb{L}_{0_\Gamma} = \frac{q - 1 - \mathfrak{N}_{q,\rho}}{6}.$$

The plane-line incidence matrix contains, according to (3.2), (3.3), *the same values of the point-line incidence matrix, but in this case they refer to* Π_π, Λ_π *instead of* $\mathbb{P}_{\mathtt{p}}, \mathbb{L}_{\mathtt{p}}$.

8 Orbits $\mathbb{O}_\rho$, odd q

Lemma 12. *Let $q \equiv 1 \pmod 3$ be odd. Let $\mathfrak{R}_m$ be as in Section 3. Then the values of $\rho \in \mathbb{F}_q^*$ can be partitioned into three classes $\mathfrak{R}_0, \mathfrak{R}_1, \mathfrak{R}_2$ such that $\#\mathfrak{R}_m = (q-1)/3$, $m = 0, 1, 2$. Moreover, let $\log(-2) \equiv \psi \pmod 3$, $\psi \in \{0, 1, 2\}$. Then we have the following.*

- *There exist two classes $\mathfrak{R}_m$ such that -2ρ is a non-cube in $\mathbb{F}_q$ for $\rho \in \mathfrak{R}_m$: $\mathfrak{R}_1, \mathfrak{R}_2$ if $\psi = 0$; $\mathfrak{R}_0, \mathfrak{R}_1$ if $\psi = 1$; $\mathfrak{R}_0, \mathfrak{R}_2$ if $\psi = 2$.*
- *There exists one class $\mathfrak{R}_m$ such that -2ρ is a cube in $\mathbb{F}_q$ for $\rho \in \mathfrak{R}_m$: $\mathfrak{R}_0$ if $\psi = 0$; $\mathfrak{R}_2$ if $\psi = 1$; $\mathfrak{R}_1$ if $\psi = 2$.*

Theorem 12. *Let $q \equiv 1 \pmod 3$ be odd. Then we have the following.*

- *Let $\rho_1 \neq \rho_2$. Then two lines $\mathscr{L}_{\rho_1}$ and $\mathscr{L}_{\rho_2}$ belong to distinct orbits under G_q if and only if $\log \rho_1 \not\equiv \log \rho_2 \pmod 3$, i.e. ρ_1, ρ_2 belong to distinct classes $\mathfrak{R}_m$. All $\mathscr{L}_\rho$-lines generate three distinct orbits $\mathbb{O}_\rho$ every of which contains $(q-1)/3$ $\mathscr{L}_\rho$-lines with ρ belonging to the same class $\mathfrak{R}_m$.*
- *Two orbits $\mathbb{O}_\rho$, say $\mathbb{O}_\rho^{(1)}$ and $\mathbb{O}_\rho^{(2)}$, have size $\frac{1}{3}(q^3 - q)$ and are generated by lines $\mathscr{L}_\rho$ such that -2ρ is a non-cube in $\mathbb{F}_q$, in accordance with Lemma 12. The orbits $\mathbb{O}_\rho^{(1)}$ and $\mathbb{O}_\rho^{(2)}$ are different from any orbit $\mathscr{O}_\mu$ of* [12, Section 7]*, see also Theorem 3.*
- *The third orbit $\mathbb{O}_\rho$, say $\mathbb{O}_\rho^{(3)}$, has size $\frac{1}{12}(q^3 - q)$ and is generated by a line $\mathscr{L}_\rho$ such that -2ρ is a cube in $\mathbb{F}_q$, in accordance with Lemma 12.*
- *If $q \not\equiv 1 \pmod{12}$ or $-1/3$ is not a fourth degree in $\mathbb{F}_q$, i.e. the condition $\Upsilon_{q,\mu}$ of Theorem 3 does not hold, then the orbit $\mathbb{O}_\rho^{(3)}$ is different from any orbit $\mathscr{O}_\mu$.*
- *If $q \equiv 1 \pmod{12}$ and $-1/3$ is a fourth degree in $\mathbb{F}_q$, then $\mathbb{O}_\rho^{(3)} = \mathscr{O}_{-1/3}$.*

Theorem 13. *Let $q \equiv -1 \pmod 3$ be odd. Then all $\mathscr{L}_\rho$-lines generate the same $\frac{1}{2}(q^3 - q)$-orbit $\mathbb{O}_1$ that is the orbit $\mathscr{O}_\mathcal{L}$* [12, Lemma 3.4(i), Theorem 3.5(iv)]*. Moreover, this orbit $\mathbb{O}_1$ is different from any orbit $\mathscr{O}_\mu$ of* [12, Section 7] *except when $q \equiv -1 \pmod{12}$; in this case the orbit $\mathbb{O}_1$ coincides with the orbit $\mathscr{O}_{-1/3}$ generated by the line $\ell_{-1/3}$ of* [12].

Acknowledgments. The research of S. Marcugini and F. Pambianco was supported in part by the Italian National Group for Algebraic and Geometric Structures and their Applications (GNSAGA - INDAM) (Contract No. U-UFMBAZ-2019-000160, 11.02.2019) and by University of Perugia (Project No. 98751: Strutture Geometriche, Combinatoria e loro Applicazioni, Base Research Fund 2017–2019; Fighting Cybercrime with OSINT, Research Fund 2021).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bartoli, D., Davydov, A.A., Marcugini, S., Pambianco, F.: On planes through points off the twisted cubic in PG(3,q) and multiple covering codes, Finite Fields Appl. **67**, Article 101710 (2020)

2. Blokhuis, A., Pellikaan, R., Szönyi, T.: The extended coset leader weight enumerator of a twisted cubic code, Des. Codes Cryptogr. **90**(9), 2223–2247 (2022)
3. Bonoli, G., Polverino, O.: The twisted cubic in $\mathrm{PG}(3,q)$ and translation spreads in $H(q)$, Discrete Math. **296**, 129–142 (2005)
4. Bruen, A.A., Hirschfeld, J.W.P.: Applications of line geometry over finite fields I: The twisted cubic, Geom. Dedicata **6**, 495–509 (1977)
5. Caputo, S., Korchmáros, G., Sonnino, A.: Multilevel secret sharing schemes arising from the normal rational curve, Discrete Appl. Math. **284**, 158–165 (2020)
6. Ceria, M., Pavese, F.: On the geometry of a $(q+1)$-arc of $\mathrm{PG}(3,q)$, q even, Discrete Math. **346**, Article 113594 (2023)
7. Cossidente, A., Hirschfeld, J.W.P., Storme, L.: Applications of line geometry, III: The quadric Veronesean and the chords of a twisted cubic, Austral. J. Combin. **16**, 99–111 (1997)
8. Davydov, A.A., Marcugini, S., Pambianco, F.: On cosets weight distributions of the doubly-extended Reed-Solomon codes of codimension 4, IEEE Trans. Inform. Theory, **67**(8), 5088–5096 (2021)
9. Davydov, A.A., Marcugini, S., Pambianco, F.: Twisted cubic and point-line incidence matrix in $\mathrm{PG}(3,q)$, Des. Codes Cryptogr. **89**(10), 2211–2233 (2021)
10. Davydov, A.A., Marcugini, S., Pambianco, F.: Twisted cubic and plane-line incidence matrix in $\mathrm{PG}(3,q)$, J. Geom. **113**(2), Article 29 (2022)
11. Davydov, A.A., Marcugini, S., Pambianco, F.: Orbits of lines for a twisted cubic in $\mathrm{PG}(3,q)$, Mediterr. J. Math. **20**(3), Article 132 (2023)
12. Davydov, A.A., Marcugini, S. Pambianco, F.: Orbits of the class $\mathcal{O}_6$ of lines external to the twisted cubic in $\mathrm{PG}(3,q)$, Mediterr. J. Math. **20**(3), Article 160 (2023)
13. Davydov, A.A., Marcugini, S., Pambianco, F.: Incidence matrices for the class $\mathcal{O}_6$ of lines external to the twisted cubic in $\mathrm{PG}(3,q)$, J. Geom. **114**(2), Article 21 (2023)
14. Davydov, A.A., Marcugini, S., Pambianco, F.: Further results on orbits and incidence matrices for the class $\mathcal{O}_6$ of lines external to the twisted cubic in $\mathrm{PG}(3,q)$, arXiv:2401.00333 [math.CO] (2023)
15. Giulietti, M., Vincenti, R.: Three-level secret sharing schemes from the twisted cubic, Discrete Math. **310**, 3236–3240 (2010)
16. Günay, G., Lavrauw, M.: On pencils of cubics on the projective line over finite fields of characteristic > 3, Finite Fields Appl. **78**, Article 101960 (2022)
17. Hirschfeld, J.W.P.: Finite Projective Spaces of Three Dimensions, Oxford Univ. Press, Oxford (1985)
18. Kaipa, K., Patanker, N., Pradhan, P.: On the $PGL_2(q)$-orbits of lines of $\mathrm{PG}(3,q)$ and binary quartic forms, arXiv:2312.07118 [math.CO] (2023)
19. Korchmáros, G., Lanzone, V., Sonnino, A.: Projective k-arcs and 2-level secret-sharing schemes, Des. Codes Cryptogr. **64**(1), 3–15 (2012)

New Models for the Cryptanalysis of ASCON
Extended Abstract

Mathieu Degré[1], Patrick Derbez[1], Lucie Lahaye[2], and André Schrottenloher[1]

[1] Univ Rennes, Inria, CNRS, IRISA (firstname.lastname@irisa.fr)
[2] ENS de Lyon (firstname.lastname@ens-lyon.fr)

Abstract. This paper focuses on the cryptanalysis of the ASCON family using automatic tools. We analyze two different problems with the goal to obtain new modelings, both simpler and less computationally heavy than previous works (all our models require only a small amount of code and run on regular desktop computers).
The first problem is the search for Meet-in-the-middle attacks on reduced-round ASCON-XOF. Starting from the MILP modeling of Qin et al. (EUROCRYPT 2023 & ePrint 2023), we rephrase the problem in SAT, which accelerates significantly the solving time and removes the need for the "weak diffusion structure" heuristic. This allows us to reduce the memory complexity of Qin et al.'s attacks and to prove some optimality results.
The second problem is the search for lower bounds on the probability of differential characteristics for the ASCON permutation. We introduce a lossy MILP encoding of the propagation rules based on the Hamming weight, in order to find quickly lower bounds which are comparable to the state of the art. We find a small improvement over the existing bound on 7 rounds.

Keywords: ASCON, Symmetric Cryptanalysis, Meet-in-the-middle Cryptanalysis, Differential Cryptanalysis, Mixed Integer Linear Programming, SAT

1 Introduction

Ascon [4] is a family of permutation-based authenticated encryption and hashing, which was selected in 2023 as the winner of the NIST Lightweight Encryption standardization process [15]. The Ascon permutation operates on a 320-bit state represented as an array of bits with 64 columns and 5 rows. Variants of the permutation are obtained by iterating the round function $p = p_L \circ p_S \circ p_C$ where the linear layer p_L applies row-wise, the S-Box layer p_S (of degree 2) column-wise and p_C is a constant addition. The n-round Ascon permutation is then simply written as p^n.

Our work targets both the permutation and the Ascon-XOF function, which is defined using a Sponge mode of operation with a 12-round permutation, a rate $r = 64$ and a variable output length size. We select a size $h = 128$ bits, whereas the hash function Ascon-Hash has an output of 256 bits. The inner part is located on the first row of the state.

Automatic Tools for Cryptanalysis. Automatic tools have been widely used in order to find and optimize attacks on the Ascon family, either reducing the search of an attack to a SAT, SMT or MILP problem [14] or an ad hoc problem which is solved automatically [9]. The main issue with Ascon is its large state size (320 bits) and weakly aligned structure, which essentially requires a model to define at least 320 variables for each round. The resulting SAT or MILP problems can be complex, computationally heavy to solve, and often do not terminate.

Contributions. In this work, we target two cryptanalytic problems with the aim of simplifying the models and reducing their runtime. The first problem is the optimization of Meet-in-the-middle preimage attacks on Ascon-XOF, which currently allow to attack the 3- or 4-round versions. Following a framework of Qin et al. [13,14], we design a simple SAT modeling (whereas they used MILP). With this new modeling, we reduce the memory footprint of the attacks and prove some impossibility results (under the assumption of a symmetry in the attack paths).

Note that in an independent line of work, Li et al. [11] optimized a different algebraic preimage attack using a SAT modeling as well. This led to small improvements in time complexity with respect to [14] and reduction of the memory to negligible. Therefore, the attacks that we present here are not strictly the best preimage attacks on Ascon-XOF, but the best MITM attack paths.

The second problem is the search for differential characteristics of the permutation. We observe that a simple MILP model, with a lossy approximation of the differential transition table of the S-Box, allows to recover quite good lower bounds on the probabilities. We improve the current best lower bound on 7 rounds of Ascon.

The code of our models is available at:

https://gitlab.inria.fr/capsule/ascon-new-models

2 New SAT Model for MITM Preimage Attacks

The Meet-in-the-middle attacks on Ascon-XOF using automatic tools were developed by Qin et al. in two works [13,14][3]. They provide currently the best preimage attacks on Ascon-XOF together with the algebraic attacks of Li et al. [11].

In Ascon-XOF, the attack focuses on the second-to-last permutation call, which is reduced to R rounds instead of 12. The last message block M_3 is absorbed and the first hash block H_1 is returned. The goal is to find a block M_3 so that H_1 matches the target image. Afterwards, the attack is repeated until H_2 also matches. This situation is represented in Figure 1.

[3] We note that after our work, the ePrint report [14] was withdrawn, but to the best of our knowledge, its results remain valid.

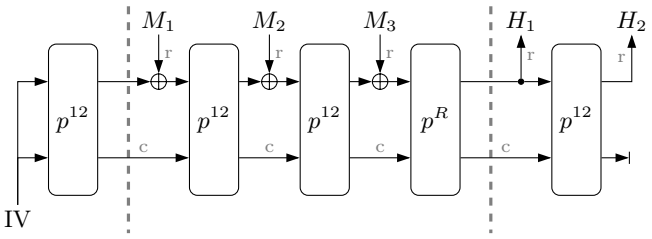


Fig. 1: Representation of the Ascon-XOF structure in the context of this attack. This figure uses the "TikZ library for crypto" of [5].

2.1 Overview of the Attack

At M_3, the 64 bits of the inner part are separated into three sets of bits: *red* bits x_R, *blue* bits x_B and *gray* x_G bits (the remaining ones). Red and blue bits will form two independent computation paths, starting from M_3 (and the current state of the outer part). Gray bits are fixed constants (this includes the padding bits in M_3).

After a configuration of M_3 is selected, the propagation of red and blue bits throughout p^R is a deterministic process. In subsequent rounds, a bit remains blue if it can be computed entirely from the knowledge of the initial blue bits (and gray), and likewise for red bits. Qin et al. [13] also introduce *green* bits, which can be expressed as a linear combination of blue and red degrees of freedom. The other bits are *white* (unknown). This coloring scheme produces figures like Figure 2.

A MITM path can lead to an attack if there exists green bits in the inner part after p^R, denoted *matching bits*. Indeed, such bits immediately translate into equations of the form: $h = f(x_R, x_G) \oplus g(x_B, x_G) \oplus c$ for some constant c. One can find solutions to these equations by enumerating $f(x_R, x_G)$ and $g(x_B, x_G)$ independently, the core of the MITM attack. The time and memory complexities of the attack depend on the number of red bits $\mathcal{D}_r$, of blue bits $\mathcal{D}_b$ and matching bits $\mathcal{D}_m$ as follows:

$$\begin{cases} T = 2^{63-\mathcal{D}_r-\mathcal{D}_b} \max\left(2^{\mathcal{D}_r}, 2^{\mathcal{D}_b}, 2^{\mathcal{D}_r+\mathcal{D}_b-\mathcal{D}_m}\right) \\ M = \min\left(2^{\mathcal{D}_r}, 2^{\mathcal{D}_b}\right) \end{cases} . \tag{1}$$

The attack runs in several *episodes* which fix the value of x_G first and find a list of candidates for M_3. (The true formula is also more complicated because it involves fixing constraints on the outer part before p^R, but this additional cost remains non-dominating in the attacks).

This first attack framework is improved by Qin et al. using *red bit cancellations*. Roughly speaking, a red degree of freedom can be removed from any bit within the propagation, by paying an additional cost. More precisely, we will select $\mathcal{A}_r$ bits in the path with a red component. For each such component $f(x_R, x_G)$, we equate it with a new bit-variable y, and we add y to the blue

degrees of freedom. This turns the red degree of freedom into a new internal matching point. The (simplified) complexities become:

$$\begin{cases} T = 2^{63-\mathcal{D}_r-\mathcal{D}_b} \max\left(2^{\mathcal{D}_r}, 2^{\mathcal{D}_b+\mathcal{A}_r}, 2^{\mathcal{D}_r+\mathcal{D}_b-\mathcal{D}_m-\mathcal{A}_r}\right) \\ M = \min\left(2^{\mathcal{D}_r}, 2^{\mathcal{D}_b+\mathcal{A}_r}\right) \end{cases} . \tag{2}$$

2.2 Automatic Search Strategies

The *configuration* of a MITM preimage attack is given by the coloring pattern of bits, the choice of cancellation points and conditions on the outer part. Considering the $\log_2$ of time and memory complexities, and the highest term, minimizing them is a linear optimization problem.

Qin et al. used a *Mixed-Integer Linear Programming* modeling, i.e., optimization under linear inequalities using real, integer and Boolean variables. Due to the large number of Boolean variables required for the entire Ascon state, the MILP solver cannot prove the optimality of the solutions that it finds. This is why the 3-round attack on Ascon-XOF of [13] could be improved in [14] with the heuristic method of "weak-diffusion structure", that fixes part of the configuration to accelerate the search of solutions.

2.3 New SAT Modeling

In our SAT modeling, the color of each bit x_i is encoded differently than before, using three Boolean variables (x_i^b, x_i^r, x_i^w). The variable x_i^b indicates whether the bit has a "blue part", x_i^r whether it has a "red part", x_i^w whether it is white (x_i^w dominates). The cancellation of red bits are also indicated by Boolean variables. Because they ultimately serve to avoid nonlinear effects, cancellations occur only in the state before p_S.

The propagation of colors from M_3 is deterministic. While the initial state requires a specific modeling, due to the additional constraints on the inner part in Qin et al.'s framework [13], the rest of the path is encoded as follows. • through p_L: each bit after p_L is the XOR of three bits located at different columns. If x_1, x_2, x_3 are the previous bits, the color of $y = x_1 \oplus x_2 \oplus x_3$ is determined by three implications. If one of the x_i is white, then y is white. If one of the x_i is blue, then y is blue. If one of the x_i is red and there is no cancellation, y is red. • through p_S: we simply look at the algebraic expression of each output bit as a function of the input bits. Blue (resp. red, white) colors propagate from the inputs to the outputs, and if a quadratic term of blue and red appears, then the output becomes also white.

Objective. To optimize the time complexity, we use Boolean cardinality constraints, which translate into a set of clauses an inequality of the form $\sum_{i=1}^{n} x_i \leq k$ where x_i are Boolean variables and k is a constant. To simulate the minimization of the time, we manually impose an upper bound on T and let the solver find a solution (or declare the problem "unsat"). To minimize the memory, we constrain the number of gray bits in M_3 (as the memory complexity primarily depends on them).

Table 1: Summary of results obtained. $\star$: optimal among MITM attacks of this form, under the symmetry constraint. $\star\star$: optimal when the time complexity is at 2^{124}. We consider only the dominating term in the complexity.

Number of Rounds	Authors	Time Complexity	Memory Complexity
3 rounds	[14]	2^{114}	2^{30}
	Ours	2^{114}	2^{24}
4 rounds	[13]	2^{124}	2^{54}
	Ours	2^{124} ($\star$)	2^{34} ($\star\star$)
5 rounds	**Ours**	2^{128} ($\star$)	—

Optimizations. Further optimizations of the model allowed to reduce the typical solving time down to less than an hour on a desktop computer, using the *Glucose4* solver within the Python library PySAT [10]. First, we hard-coded a specific cancellation strategy for the last S-Box in the 4-rounds attack, where all inputs to the last S-Box are green, and we cancel two bits to obtain a constraint. Second, we noticed that all solutions generated by our solver for the 4-round attack were circularly symmetric, meaning that the diffusion pattern repeated 32 positions later. In fact, the solutions found in [13,14] also exhibit a strong symmetry, and we have not found to date better solutions which would be non-symmetric. By imposing this symmetry, we reduced the number of variables by a factor 2. This was the key factor in allowing our model to run effortlessly.

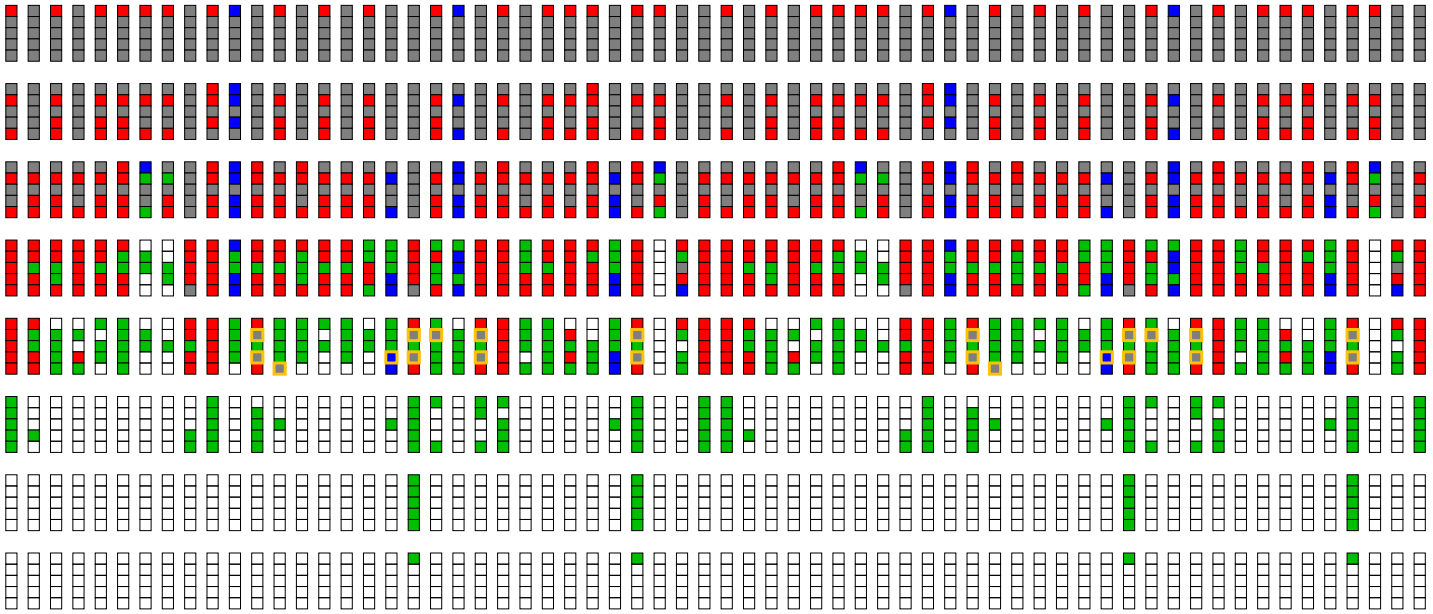
2.4 Results

While we did not improve the time complexities reported in [13] for 4-round Ascon (around 2^{124}, using 4 bits of matching) and in [14] for 3-round Ascon (around 2^{114}, using 14 bits of matching), we reduced the memory complexities as seen in Table 1. The path for the 4-round attack is shown in Figure 2, with $\mathcal{D}_b = 4$, $\mathcal{D}_r = 34$, $\mathcal{D}_m = 4$, $\mathcal{A}_r = 30$.

We also obtained some optimality results. For Ascon with 4 rounds, improving the time complexity would mean setting $\mathcal{D}_m \geq 5$. With about 3 hours of computation, we could show that such a path does not exist (assuming symmetry). Next, fixing the time complexity of the 4-round attack at 2^{124}, we find that it is impossible to increase the number of gray bits further, proving that our memory complexity is also optimal. Finally, we found that there is no valid configuration with $\mathcal{D}_m \geq 1$ for a 5-round path, i.e., this technique cannot reach 5 rounds. Our results are summarized in Table 1.

3 New MILP Model for Differential Bounds

Our second simple model aims at obtaining lower bounds on the probability of differential characteristics for the Ascon permutation. The bounds proven so



Optimal 4-round attack found with our SAT-based approach, improving the one of [13].

Fig. 2: Cancellations of red bits are represented by yellow squares. There are 52 bits of additional constraints which are imposed on the inner part of the first state to guarantee the first transition through p_S.

far are only tight up to 3 rounds despite years of investigation as summarized in Table 2, with different methods such as MILP [12], SAT and SMT [6]. The state of the art for lower bounds is given by a tree extension model from [9].

Table 2: Currently known differential bounds of the Ascon permutation restricted to R rounds.

	Upper bound			Lower Bound		
R	Bound	Method	Ref.	Bound	Method	Ref.
1	2^{-2}	DDT		2^{-2}	DDT	
2	2^{-8}	DDT + β		2^{-8}	DDT + β	
3	2^{-40}	ndltool	[3]	2^{-40}	MILP	[12]
4	2^{-107}	ndltool		2^{-86}	Tree extension	
5	2^{-190}	CP	[3],[7]	2^{-100}	Tree extension	
6	2^{-305}	CP	[7]	2^{-129}	Tree extension	
7				2^{-131}	Tree extension	
8				2^{-172}	Tree extension	[9]
9				2^{-186}	Tree extension	
10				2^{-215}	Tree extension	
11				2^{-229}	Tree extension	
12				2^{-258}	Tree extension	

One interesting remark is that many of the lower bounds on the probability are computed from the lower bounds of a lower number of rounds. For example, the current lower bound on 12 rounds is computed from the bound on 4 rounds: $2^{-258} = 2^{-86\times 3}$. As such, improving the lower bounds give us immediate results for higher rounds. This also means that the results for higher rounds are most likely far to be tight, which is especially visible for the bound on 7 rounds: $2^{-131} = 2^{-129} \times 2^{-2}$.

Our main objective was to reduce the gap between the lower and upper bounds. For now, the bounds for 4 rounds are **[86, 107]** (or to be more exact, since the bounds are probabilities, $[2^{-107}, 2^{-86}]$). However, [12] showed that there exists a 4 round trail with 43 active S-boxes. This mean that, if we want to reduce the gap by improving the lower bound on the weight, we cannot take the number of active S-boxes as objective, since we know that the minimum number of active S-boxes will be ≤ 43 and as such we will not be able to tell a better precision than $43 \times 2 = 86$ but have to model the transition weights.

3.1 A new MILP Approach: using the Hamming Weight

Our idea to improve the existing bounds is to consider a lossy model, as it was done in [2] regarding division property related problems. This would allow us to get calculations done much quicker, but at the cost of getting less accurate results. The main difficulty is to find a right balance between accuracy and time complexity. As such, we tried to model the internal state of Ascon using the Hamming weight of columns.

Table 3: hwDDT of Ascon's S-box: maximum number of solutions as a function of the Hamming weight of the input and output.

hw	0	1	2	3	4	5
0	32	0	0	0	0	0
1	0	0	8	8	8	4
2	0	8	8	8	8	4
3	0	8	8	4	4	4
4	0	4	4	4	4	2
5	0	4	4	4	0	0

The Hamming weight of a column is defined by the total number of active bits in this column. As such, instead of modeling 320 bits, we can manipulate 64 integers between 0 and 5. This approach works particularly well with MILP models, as they deal natively with integer variables.

3.2 Modeling p_S

To model the nonlinear layer, we need a way to represent the weight of each transition through the Ascon S-Box. Each transition has a weight of 2, 3, 4 or 0 (for the trivial $0 \rightarrow 0$ differential transition), and thus for each of them we create binary variables w_2, w_3, w_4 and w_0 with the constraint that $w_2[i] + w_3[i] + w_4[i] + w_0[i] = 1$.

Now that we have a way to represent the weights of the transition, we then have to find a way to model the DDT, then link both of them together. The classical way of modelling the DDT into inequalities is to use the convex hull operator. This operator takes a cloud of points in n dimensions and returns the convex hull associated with this cloud of points, which can take the form of inequalities, which is what we want. In our case, we want to know what is the minimum weight (the worst case) of a differential transition where the input difference has a Hamming weight of $\mathrm{hw}(i)$ and the output has a Hamming weight of $\mathrm{hw}(j)$. As such, our hwDDT can be expressed $(i, j) \in \{0, ..., 5\}$ as:

$$\mathrm{hwDDT}(i, j) = \max(\mathrm{DDT}(a, b) \mid \mathrm{hw}(a) = i, \mathrm{hw}(b) = j),$$

and is given in Table 3.

More precisely, this new DDT contains the best transitions between an input and an output of given Hamming weight. Hopefully, it is quite straightforward

to describe a transition $a \to b$ through this DDT:

$$\begin{array}{lll} a \geq 2w_4 & a \leq 5 - 5w_0 & a + b \geq 3w_2 \\ a \leq 5 - w_4 & b \leq 5 - 5w_0 & a \geq w_2 \\ b \geq w_4 & b \leq 5 - w_2 & b \geq w_2 \\ a \geq w_3 & a \leq 5 - 2w_2 & 2a + b \leq 15 - 7w_2 \\ b \geq w_3 & & \end{array}$$

Note that for the first and last non-linear layers, the constraints are simpler since we can assume that the best input or output will be selected. Furthermore, we only model Ascon from the output of the first non-linear layer to the input of the last one.

3.3 Modeling p_L

To model the linear layer, we need to express the relation between the bits of the state before and after this step. If we denote respectively by y and x these states, we have for all $(i, j) \in \{0, ..., 5\} \times \{0, ..., 64\}$:

$$x[i][j] = y[i][j] \oplus y[i][(j - d_1) \bmod 64] \oplus y[i][(j - d_2) \bmod 64],$$

where d_1 and d_2 depend on the row index. In [12], the authors proposed to use an extra binary variable per state bit to model the equation in a MILP-compliant form:

$$y[i][j] + y[i][(j - d_1) \bmod 64] + y[i][(j - d_2) \bmod 64] + x[i][j] = 2z[i][j].$$

However this modeling seems to be quite inefficient, making the model very slow to solve. Our idea is to describe as accurately as possible the possible transitions through the linear layer without going down to the bit level. To do so we introduce the variables x_{row}, x_{col}, y_{row} and y_{col} corresponding to the Hamming weight of the rows and columns of both states x and y (note that x_{col} and y_{col} are not new since they respectively correspond to the Hamming weight of the input and output of the S-boxes which are used in the modelization of the non-linear part).

First there are many straightforward relations between those states. For instance, $\sum_{i=0}^{4} x_{row}[i] = \sum_{i=0}^{6} 3x_{col}[i]$ and the same equality holds for y. It is also easy to add a constraint ensuring that an active column of y should at least activate the same column on x or one of the 10 associated columns of y. We also add the following constraint on rows:

$$3y_{row}[i] = x_{row}[i] + 2z,$$

where z is an extra integer variable, representing the number of cancellations occurring on the row.

3.4 Callbacks

Our model is very fast compared to previous ones, and in particular compared to [12]. However, the results are far from being accurate as many *false* trails are solutions of the model. To strengthen the model we use the callback functionality of the Gurobi MILP solver [8]. It allows to add an extra verification each time a solution of the model is found. First, for each linear layer, we check whether the pattern of active columns is possible using Gaussian elimination as it was done in [1]. If not we add an extra constraint to remove the pattern and the model continues to search for another solution. Finally, the whole trail is checked using an exact model. Note that the inequalities added to the model during the callback only involve the weight of the transitions as they all are binary variables.

3.5 Results

Our model is fast enough to retrieve the lower bound of the weight of differential characteristics up to 3 rounds. We also improve the lower bound for 7-round, showing that the minimal weight is at least 135 while the previous bound was 131. Note that these results were obtained on a laptop in few hours/days of computation, thus it makes no doubt that running the model on a bigger machine will lead to new results.

4 Conclusion

In this work we proposed several improvements for the modelization of important cryptanalysis problems related to the security of Ascon. We successfully decrease the running times required to search for some instances of both meet-in-the-middle preimage attacks on Ascon-XOF and lower bounds on the weight of differential characteristics on Ascon inner permutation, and obtained new results as well. The techniques we described show that it is sometimes more efficient to rely on simple modelizations, even though they are not exact, and we believe they could be used to improve models dedicated to other primitives.

Acknowledgments. This work has been partially supported by the French Agence Nationale de la Recherche through the OREO project under Contract ANR-22-CE39-0015, and through the France 2030 program under grant agreement No. ANR-22-PECY-0010.

References

1. Boura, C., Derbez, P., Funk, M.: Related-key differential analysis of the AES. IACR Trans. Symmetric Cryptol. **2023**(4), 215–243 (2023). https://doi.org/10.46586/TOSC.V2023.I4.215-243
2. Derbez, P., Lambin, B.: Fast MILP models for division property. IACR Trans. Symmetric Cryptol. **2022**(2), 289–321 (2022). https://doi.org/10.46586/TOSC.V2022.I2.289-321

3. Dobraunig, C., Eichlseder, M., Mendel, F., Schläffer, M.: Cryptanalysis of ASCON. In: Nyberg, K. (ed.) Topics in Cryptology - CT-RSA 2015, The Cryptographer's Track at the RSA Conference 2015, San Francisco, CA, USA, April 20-24, 2015. Proceedings. Lecture Notes in Computer Science, vol. 9048, pp. 371–387. Springer (2015). https://doi.org/10.1007/978-3-319-16715-2_20
4. Dobraunig, C., Eichlseder, M., Mendel, F., Schläffer, M.: Ascon v1.2: Lightweight authenticated encryption and hashing. J. Cryptol. **34**(3), 33 (2021). https://doi.org/10.1007/s00145-021-09398-9
5. Eichlseder, M.: TikZ libraries for crypto, https://extgit.iaik.tugraz.at/meichlseder/tikz
6. Erlacher, J., Mendel, F., Eichlseder, M.: Bounds for the security of Ascon against differential and linear cryptanalysis. IACR Trans. Symmetric Cryptol. **2022**(1), 64–87 (2022). https://doi.org/10.46586/TOSC.V2022.I1.64-87
7. Gérault, D., Peyrin, T., Tan, Q.Q.: Exploring differential-based distinguishers and forgeries for ASCON. IACR Trans. Symmetric Cryptol. **2021**(3), 102–136 (2021). https://doi.org/10.46586/TOSC.V2021.I3.102-136
8. Gurobi Optimization, LLC: Gurobi Optimizer Reference Manual (2023), https://www.gurobi.com
9. Hirch, S.E., Mella, S., Mehrdad, A., Daemen, J.: Improved differential and linear trail bounds for Ascon. IACR Trans. Symmetric Cryptol. **2022**(4), 145–178 (2022). https://doi.org/10.46586/TOSC.V2022.I4.145-178
10. Ignatiev, A., Morgado, A., Marques-Silva, J.: PySAT: A Python toolkit for prototyping with SAT oracles. In: SAT. pp. 428–437 (2018). https://doi.org/10.1007/978-3-319-94144-8_26
11. Li, H., He, L., Chen, S., Guo, J., Qiu, W.: Automatic preimage attack framework on Ascon using a linearize-and-guess approach. IACR Transactions on Symmetric Cryptology **2023**(3), 74–100 (Sep 2023). https://doi.org/10.46586/tosc.v2023.i3.74-100, https://tosc.iacr.org/index.php/ToSC/article/view/11185
12. Makarim, R.H., Rohit, R.: Towards tight differential bounds of Ascon: A hybrid usage of SMT and MILP. IACR Trans. Symmetric Cryptol. **2022**(3), 303–340 (2022). https://doi.org/10.46586/TOSC.V2022.I3.303-340
13. Qin, L., Hua, J., Dong, X., Yan, H., Wang, X.: Meet-in-the-middle preimage attacks on sponge-based hashing. In: EUROCRYPT (4). Lecture Notes in Computer Science, vol. 14007, pp. 158–188. Springer (2023). https://doi.org/10.1007/978-3-031-30634-1_6
14. Qin, L., Zhao, B., Hua, J., Dong, X., Wang, X.: Weak-diffusion structure: Meet-in-the-middle attacks on sponge-based hashing revisited. IACR Cryptol. ePrint Arch. p. 518 (2023), https://eprint.iacr.org/2023/518
15. Turan, M.S., McKay, K., Chang, D., Kang, J., Waller, N., Kelsey, J.M., Bassham, L.E., Hong, D.: Status report on the final round of the nist lightweight cryptography standardization process (2023)

Equivalence of Generalised Feistel Networks

Patrick Derbez[1] and Marie Euler[1,2]

[1] Univ Rennes, CNRS, IRISA, Rennes, France
firstname.lastname@irisa.fr
[2] DGA MI, Bruz, France

Abstract. This paper focuses on equivalences between Generalised Feistel Networks (GFN) of type-II. We introduce a new definition of equivalence which captures the concept that two GFNs are identical up to re-labelling of the inputs/outputs and are therefore cryptographically equivalent for several classes of attacks. It induces a reduction of the space of possible GFNs: the set of the $(k!)^2$ possible even-odd GFNs with $2k$ branches can be partitioned into $k!$ different classes.
From a designer perspective, it means that a much wider spectre of candidates can be explored to choose a good permutation. This leads to the suggestion of five 62-branches permutations performing better than WARP regarding the number of differentially/linearly active S-Boxes and to a new family of permutations with good diffusion properties.

1 Introduction

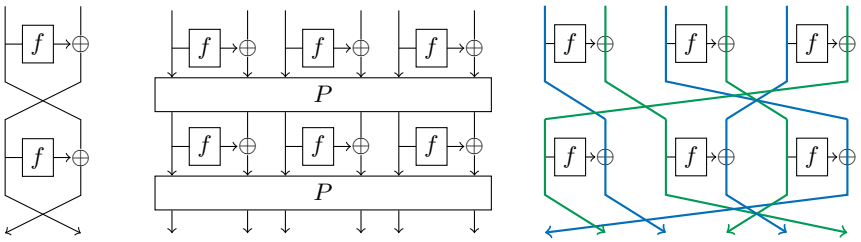
A Feistel network is a widely spread structure for symmetric cryptography primitives. Invented by Feistel and Coppersmith in 1973 for IBM's `Lucifer` cipher, it was later standardised in the block cipher `DES` in 1976 [S+77]. In a Feistel network, the internal state is divided into two parts of the same size: the left branch x and the right branch y. The round function of the i-th round of the Feistel network is the involutive operation $F_i := (x, y) \mapsto (x, y \oplus f_i(x))$, where f_i is a keyed function, followed by the swap of x and y as depicted in Figure 1a.

Later, Zheng et al. [ZMI90] generalised the original construction so that the state is now divided into $2k$ same-size parts $(x_0, x_1, \ldots, x_{2k-1})$. These parts are also called branches. Several generalisations – named type-I, type-II and type-III – were suggested, but in this paper we will focus on type-II, which seems to be the design favoured by the community. The round function of the i-th round of a type-II generalised Feistel network is

$$F_i^k := (x_0, x_1, \ldots, x_{2k-1}) \mapsto (x_0, x_1 \oplus f_i(x_0), \ldots, x_{2k-2}, x_{2k-1} \oplus f_i(x_{2k-2})),$$

followed by a circular shift of the $2k$ parts of the state, sending each branch to the next one: $(x_0, x_1, \ldots, x_{2k-1}) \mapsto (x_{2k-1}, x_0, x_1, \ldots, x_{2k-2})$.

At Asiacrypt '96, Nyberg proposed to replace the circular shift by another specific permutation [Nyb96]. Then, in [SM10], Suzaki and Minematsu proposed the Generalised Feistel Network (GFN) by replacing the circular shift by a general permutation P (see Figure 1b), aiming to improve the diffusion. Among



(a) A Feistel Network. (b) A type-II Generalised Feistel Network with 6 branches. (c) A GFN with an even-odd permutation.

Fig. 1: 2 rounds of some types of (Generalised) Feistel Networks.

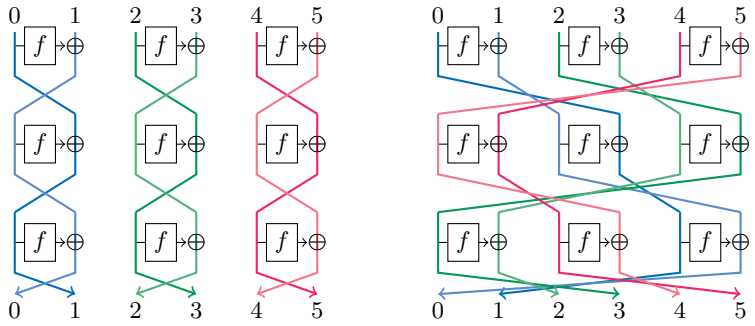
others, two type-II GFN with 16 branches (`LBlock` [WZ11], `TWINE` [SMMK13]) and one with 32 branches (`WARP` [BBI+20]) were later proposed.

The main problem in the search of optimal permutations for a certain property (diffusion, resistance against differential or linear attacks, etc.) comes from the huge search space: there are $(2k)!$ different permutations for a $2k$-branch GFN. In [CGT19], Cauchois *et al.* consider equivalence classes based on conjugacy: two GFNs relying on $2k$-permutations P and Q respectively are equivalent if and only if there exists a permutation of pairs A such that $P = AQA^{-1}$, where a permutation of pairs is a permutation A such that $A(2i+1) = A(2i)+1$ for all i between 0 and $k-1$. It allowed them to describe an efficient generation of even-odd Feistel permutations (*i.e.* permutations that map even indices to odd indices and reciprocally), based on the cycle type of the left-branches permutation.

Our contribution In this paper, we present a wider definition of equivalence of the underlying permutations of GFNs: we say that two permutations P and Q are expanded-equivalent if and only if there exists a permutation of pairs A such that for all positive integer i, Q^iAP^{-i} is a permutation of pairs. This definition takes into account several rounds of the GFN and therefore captures equivalence which are not visible on one round. Our motivation comes from the observation that equivalence notions introduced in previous works do not cover all the cases. For example, both the Feistel networks depicted in Figure 2 share the exact same properties while the inner permutations are not isomorphic (it would imply that the identity permutation and the rotation are conjugates).

Our new equivalence relation comes with two different characterisations. The first one reveals that the equivalence of GFNs can be observed as a cyclic behaviour on a finite number of rounds, which provides a new way to test whether two GFNs are equivalent. The second one, only valid for even-odd permutations, captures the structure of the equivalence classes and leads to the fact that the $k!^2$ GFNs associated with even-odd permutations on $2k$ branches can be grouped in exactly $k!$ equivalence classes, each of them containing $k!$ GFNs.

Moreover, we show how to enumerate only one element per expanded equivalence class. Exploring this smaller space of candidates, we obtain a new family of GFNs with good diffusion but also 5 good 32-branch permutations which



(a) With A being the identity map. (b) With A mapping i to $(i+1) \mod 3$.

Fig. 2: Trivial example of equivalent GFNs whose permutations are not conjugates. Both figures depict three rounds of a 6-branches GFN associated to $\Pi_{A,A}$: the i-th left-branch is mapped to the $A(i)$-th right-branch and similarly for right branches.

perform better than WARP for differential and linear properties without any degradation on the diffusion part. We also reduce the long list of permutations provided by previous works on GFN to more reasonable size using the expanded equivalence.

Organisation of the paper Section 2 is dedicated to notations, definitions, and properties useful for the following parts. Then in Section 3, we introduce the new definition of equivalence of GFNs, its characterisations, and equivalence testing. Section 4 presents several results obtained from the new equivalence relation.

2 Notations, definitions and previous works

2.1 Permutations

We will denote by $\mathcal{S}_k$ the symmetric group acting on a set with k elements. Any permutation will be described by its value table: for instance, writing $P = [0, 1, 3, 2]$ indicates that P is the permutation $P(0) = 0, P(1) = 1, P(2) = 3, P(3) = 2$. We denote by Id the identity permutation.

The centraliser of a permutation $P \in \mathcal{S}_k$ is the set of permutations that commute with P: $\texttt{Centr}(P) := \{Q \in \mathcal{S}_k, QP = PQ\}$. More generally, we will use the centraliser of a set of permutations $E \subset \mathcal{S}_k$: $\texttt{Centr}(E) := \bigcap_{P \in E} \texttt{Centr}(P)$.

2.2 Permutations used in Feistel networks

A $2k$-branch GFN is defined by its permutation of branches $P \in \mathcal{S}_{2k}$. We write $\mathcal{F}_P$ to describe the GFN whose i-th round function is $P \circ F_i^k$ and we will use the shorter product notation PF for this round function (which implies that we only

study the formal structure of the GFN, not its instantiation with functions f_i). Branches with an even (resp. odd) index are called left (resp. right) branches.

In the introduction, we have defined the even-odd permutations as the permutations of $\mathcal{S}_{2k}$ which map even numbers to odd numbers and reciprocally. Such a permutation P can be described by two smaller permutations L, R in $\mathcal{S}_k$ as follows: $L(i) := (P(2i) - 1)/2$ and $R(i) := P(2i + 1)/2$. L (resp. R) is called the left (resp. right)-branches permutation and we denote P by $\Pi_{L,R}$.

Let us also define the even permutations as the permutations which map even numbers to even numbers and odd numbers to odd numbers. Similarly, any even permutation P of $\mathcal{S}_{2k}$ can be described by two smaller permutations $L, R \in \mathcal{S}_k$: $L(i) := P(2i)/2$ and $R(i) := (P(2i + 1) - 1)/2$ and we denote P by $\Phi_{L,R}$.

Finally, we consider the group of permutations of pairs: $\mathcal{S}_k^p = \{\Phi_{A,A}, A \in \mathcal{S}_k\}$ Any permutation of pairs commutes with the Feistel step F, a property which will be very useful for the GFN isomorphism.

The diffusion round of a GFN is the minimal number of rounds needed for all output branches to depend on all the input branches (and conversely for the decryption). There exists a lower bound of the diffusion round for even-odd GFNs based on the Fibonacci sequence (ϕ_i): if $\phi_i \geq k > \phi_{i-1}$, then the diffusion round of any even-odd GFN with $2k$ branches is at least $i + 1$.

2.3 A first approach to GFN equivalence

Let us begin with a natural definition of equivalence of generalised Feistel networks: two GFNs are equivalent if, for any number of rounds, one is equal to the other up to a re-labelling of the inputs and outputs. More formally, this can be defined as follows:

Definition 1. *Let P and Q be two $2k$-permutations associated with two generalised Feistel networks $\mathcal{F}_P$ and $\mathcal{F}_Q$. $\mathcal{F}_P$ and $\mathcal{F}_Q$ are equivalent if and only if for all positive integer i, there exist two permutations A_i and B_i such that $(QF)^i = B_i(PF)^i A_i^{-1}$.*

This definition is interesting for cryptographers because it implies that both Feistel networks share some cryptographic properties: not only linear and differential characteristics but also diffusion, impossible differentials, etc. However, it is more convenient to have a property that directly links the underlying permutations. Hence [CGT19] suggested the following *natural* equivalence relations:

Definition 2. *Two $2k$-permutations P and Q are pair-equivalent if and only if there exists $A \in \mathcal{S}_k^p$ such that $Q = APA^{-1}$. P and Q are extended pair-equivalent if and only if P and Q are pair-equivalent or P and Q^{-1} are pair-equivalent.*

Indeed, in the first case, A commutes with F thus $(QF)^i = A(PF)^i A^{-1}$ for all i, and $\mathcal{F}_P$ and $\mathcal{F}_Q$ are equivalent. The second equivalence comes from the fact that $\mathcal{F}_{P^{-1}}$ corresponds to the decryption of $\mathcal{F}_P$ and thus both permutations are typically evaluated together. In the following, we will denote these equivalences as the (extended-)conjugacy-based equivalence.

3 Expanded Feistel Equivalence

In this section, we present the core of our work: a larger equivalence relation between the permutations used in GFNs. We also highlight some useful properties regarding this equivalence.

3.1 New definition

We propose the following widened equivalence relation of permutations which also implies the equivalence of associated GFNs.

Definition 3. *Two $2k$-permutations P and Q are called expanded-equivalent if and only if there exists $A \in \mathcal{S}_k^p$ such that for all i, $A_i := Q^i A P^{-i} \in \mathcal{S}_k^p$.*

Since F commutes with permutations of pairs, we can show that for any positive i, $(QF)^i = A_i(PF)^i A^{-1}$ and as a consequence, both $\mathcal{F}_P$ and $\mathcal{F}_Q$ are equivalent. Contrary to the previous equivalence notion, the output relabelling is now authorised to depend on the number of rounds. Note that P and Q are expanded-equivalent if and only if P^{-1} and Q^{-1} are expanded-equivalent. Furthermore, as for conjugacy-based equivalence, expanded equivalence can be extended to deal with the inverse permutations.

Definition 4. *Two $2k$-permutations P and Q are extended-expanded-equivalent if and only if P and Q are expanded-equivalent or P and Q^{-1} are expanded-equivalent.*

First example of a class Let us denote $\texttt{Cl}(\Pi_{L,R})$ the class of expanded equivalence of $\Pi_{L,R}$. The easiest class to compute is $\texttt{Cl}(\Pi_{\mathrm{Id},\mathrm{Id}}) = \{\Pi_{P,P}, P \in \mathcal{S}_k\}$. Indeed, if $Q \in \texttt{Cl}(\Pi_{\mathrm{Id},\mathrm{Id}})$, then there exist $A = \Phi_{a,a}, B = \Phi_{b,b} \in \mathcal{S}_k^p$ such that $QA\Pi_{\mathrm{Id},\mathrm{Id}}^{-1} = B$ *i.e.* $Q = \Pi_{ba^{-1},ba^{-1}} \in \{\Pi_{P,P}, P \in \mathcal{S}_k\}$. Conversely, if $Q = \Pi_{P,P}$ then $Q^i \Pi_{\mathrm{Id},\mathrm{Id}}^{-i} = \Phi_{P^i,P^i} \in \mathcal{S}_k^p$. Note that $\texttt{Cl}((\Pi_{\mathrm{Id},\mathrm{Id}}))$ is significantly larger than the conjugacy-based equivalence class of $\Pi_{\mathrm{Id},\mathrm{Id}}$: the former has $k!$ elements while the latter is reduced to one element.

3.2 Invariant cryptographic properties

Let us clarify here which cryptographic properties of GFNs are invariant under the equivalence relations. The conversion of the equivalence of permutations to an equivalence of GFNs requires that any permutation of pairs commutes with the function F. This implies that in each round, all the underlying Feistel functions (f_i at round i) are identical or can be considered as such. Indeed, many cryptanalysis techniques (minimal number of active S-boxes in a differential/linear trail, diffusion round and word-oriented Meet-in-the-Middle (MITM) distinguishers, etc.) do not rely on the exact specification of either the S-boxes or the key schedule. However, if the Feistel functions are not all identical (e.g. LBlock), instantiated differential/linear trails can no longer be transposed from

one GFN to another equivalent one. Similarly, related-key attacks are not invariant as the role of the round keys changes from one branch to another. More generally, key-recovery attacks are not invariant, as the behaviour of the key in the key-recovery rounds changes from one GFN to another.

Let us now introduce the main difference between conjugacy-based equivalence and expanded equivalence: invariant subspaces. An invariant subspace is a set S invariant by the round operation of the cipher. This property is preserved by conjugacy-based equivalence but not for expanded equivalence. In the latter case, only the more generic subspace trail attack framework is invariant.

3.3 Characterisation on a finite number of rounds

In practice, the former definition seems difficult to apply, as it relies on a property for all positive integers i. Yet, it can be reduced to a property verifiable on a finite number of i. This comes from the following smaller equivalence relation.

Definition 5. *Two permutations P and Q are r-cyclic equivalent if and only if for all $i \leq r$, there exists a permutation of pairs A_i such that $Q^i = A_i P^i A_0^{-1}$ and $A_r = A_0$.*

Conjugacy-based equivalence boils down to 1-cyclic equivalence. Moreover, if $Q = AP = PA^\alpha$ with A a permutation of pairs, then P and Q are r-cyclic equivalent for any r such that $A^{1+\alpha+\cdots+\alpha^{r-1}} = \mathrm{Id}$. For instance, if $\alpha = -1$, we get that P and Q are 2-cyclic equivalent. Moreover, this definition naturally leads to the following characterisation of expanded equivalence.

Property 1 (First characterisation). Two permutations P and Q are expanded-equivalent if and only if there exists a positive integer r such that they are r-cyclic equivalent.

We now suggest a procedure to test the r-cyclic equivalence between two even-odd permutations based on graph isomorphism : Two even-odd permutations P and Q are r-cyclic-equivalent if and only if their cyclic Feistel graph of length r ($G_F^c(P, r)$ and $G_F^c(Q, r)$) are isomorphic.

Definition 6. *Let P be an even-odd permutation of $2k$ elements. We call cyclic Feistel graph of length r associated to P the directed graph $G_F^c(P, r)$ such that its set of vertices is V and edges $E = E_P \bigcup E_F$ with $V = \{0, \ldots, 2k-1\} \times \{0, \ldots, r-1\}$,*

$$E_P = \{((i,j) \to (P(i), (j+1) \mod r)), (i,j) \in V\}$$
$$\text{and } E_F = \{((2i,j) \to (P(2i+1), (j+1) \mod r)), (2i,j) \in V\}.$$

3.4 More fundamental characterisation for even-odd permutations

The previous characterisation helps to understand what it means for two permutations to be equivalent. However, there can be three expanded-equivalent

permutations P, Q, R such that P and Q are r-cyclic-equivalent and Q and R are r'-cyclic-equivalent with $r \neq r'$. Thus, we introduce a more fundamental characterisation by using an alternative representation of the even-odd permutations: any even-odd permutation $\Pi_{L,R}$ can be uniquely defined by the two permutations R and $\alpha := R^{-1}L$. We denote this representation $\Psi_R^\alpha := \Pi_{R\alpha,R} = \Pi_{L,R}$.

Property 2 (Second characterisation). Two even-odd permutations $P = \Psi_R^\alpha$ and Q are expanded-equivalent if and only there exist two permutations of pairs $A = \Phi_{a,a}$ and $B = \Phi_{b,b}$ such that $Q = ABPA^{-1}$ and $b \in \texttt{Centr}(\left\{R^{-i}\alpha R^i\right\}_{i\geq 0})$.

This characterisation induces that expanded equivalence classes are the union of several conjugacy-based equivalence classes for some multiple of P.

Let us now discuss another easy example of equivalence classes: the case of $\texttt{Centr}(\left\{R^{-i}\alpha R^i\right\}_{i\geq 0}) = \{\mathrm{Id}\}$. In that case, by Property 2, $P := \Psi_R^\alpha$ is only equivalent to its conjugates via a permutation of pairs. Moreover, two conjugates via a permutation of pairs $\Phi_{a,a}$ are equal if and only if a commutes with α and with R. But $\texttt{Centr}(\{R, \alpha\})$ is a subset of $\texttt{Centr}(\left\{R^{-i}\alpha R^i\right\}_{i\geq 0})$ so there is no non-trivial element in this set. Thus, all the $k!$ conjugates of P are different and the expanded equivalence class of P has exactly $k!$ elements. Besides, it is true for all the expanded equivalence classes:

Theorem 1. *There exist $k!$ classes of expanded equivalence of even-odd GFNs with $2k$ branches. Each of these classes contains exactly $k!$ GFNs.*

4 Applications

4.1 Enumeration of expanded equivalence classes

Algorithm 1 Enumeration of expanded equivalence classes for $2k$-branch even-odd GFNs.

```
Initialise a set of representative of classes classes = {}.
for α ∈ A_k do
    Initialise a set S = S_k.
    while S is not empty do
        Pick R in S.
        Add Ψ_α^R to classes.
        for B in Centr({R^{-i} α R^i}_{i≥0}) do
            Remove {ABRA^{-1}, A ∈ Centr(α)} from S.
        end for
    end while
end for
Return classes.
```

We can use the characterisation from Property 2 to define an algorithm (see Algorithm 1) giving exactly one representative per expanded equivalence class.

We were able to run this algorithm in less than half an hour on a laptop for k up to 9 with a Python implementation.

4.2 A new family of GFNs with good diffusion properties

On small value of k, we observed that the following family of permutations leads to good diffusion properties.

Definition 7 (Pseudo-cyclic permutations). *Let k be a positive integer and j be a positive integer smaller than k. Let $i = k/\gcd(j,k)$. Let α be the cyclic permutation of order k defined as $\alpha : x \mapsto x+1 \mod k$. We call* pseudo-cyclic permutation *the permutation $\Psi_{R_j}^{\alpha}$ with R_j the permutation of $\mathcal{S}_k$ such that $R_j(x) = jx + \left\lfloor \frac{x}{i} \right\rfloor \mod k$.*

There are $k-1$ pseudo-cyclic permutations with $2k$ elements. Hence, it is easy to evaluate the diffusion round of all the pseudo-cyclic permutations for k relatively large. We reported the minimal diffusion round of pseudo-cyclic permutations for k up to 150 in Figure 3. In this figure, we distinguished the case where k is prime, since in this case, the diffusion round is really close to its lower bound. For $k \in \{11, 14, 29, 59, 61, 101, 145, 149\}$, the Fibonacci lower bound is even reached, which is a surprising result as this lower bound is not tight for smaller values of k: From [SM10,CGT19,DFLM19,DDGP22], we know that this lower bound cannot be reached for $k \in \{5, 6, 7, 8, 10, 12, 13\}$.

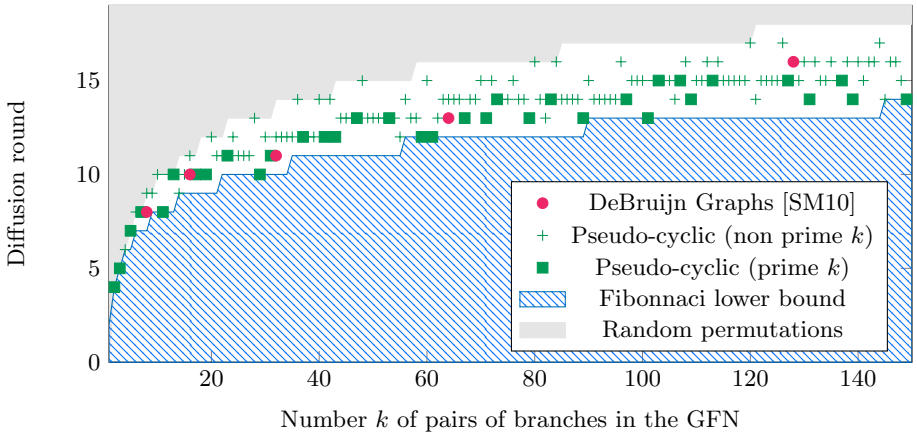


Fig. 3: Diffusion round of several families of GFNs.

In the Figure 3, we compare the pseudo-cyclic permutations with permutations obtained from De Bruijn graphs. Indeed in [SM10,CGT19], the authors showed that for k a power of 2, permutations obtained from colouring of De Bruijn graphs are good GFN candidates regarding diffusion. Experimental results about random even-odd permutations are also given for comparison.

4.3 Application to WARP

WARP [BBI+20] is a 128-bit block cipher based on a GFN with 32 branches targeting a minimalist hardware footprint. Its designers explored a space of $8! \times 16 \simeq 2^{19.3}$ candidates which led to 152 candidates with diffusion round 10. It took them 2 days on a computer with 44 cores to evaluate the number of differentially active S-boxes on 18 and 19 rounds with MILP for these 152 candidates. The goal is to find a permutation which achieves to have at least active 64 S-boxes in a minimal number of rounds. No permutation succeeds in 18 rounds but 8 permutations stand out as they have the best number of differentially active S-boxes on 19 rounds. However, the authors did not find any way to differentiate them on a large number of cryptographic properties. Indeed, using the graph isomorphism test, we can show that they are extended-expanded-equivalent.

Moreover, we can reproduce their search and regroup the 152 candidates in 7 extended expanded classes of equivalence. The MILP evaluation of the minimal number of differentially active S-boxes on 19 rounds for the 7 classes takes 50 minutes on a 12-cores laptop.

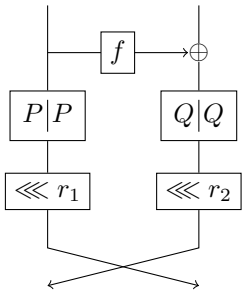


Fig. 4: Space explored by our program.

Since our search for good candidates was significantly faster, we explored a wider space of 32-branch permutations in the hope of finding one permutation with minimal AS of at least 64 after only 18 rounds and diffusion round less than or equal to 10.

We used the space described in Figure 4 (2 identical 8-branch permutations P followed by a rotation of an amount r_1 on the left branches, 2 identical 8-branch on the right branches Q followed by a rotation of an amount r_2 of the right branches) with the (P,Q) generated by Algorithm 1. Keeping only candidates with diffusion round less than or equal to 10, we obtained 184 candidates belonging to 68 extended-expanded classes. We found 5 classes of permutations[3] which need only 18 rounds (when WARP needs 19 rounds) to have the guarantee that at least 64 S-Boxes are active in any differential/linear trail. They also have the same diffusion round as WARP which enables to keep the security arguments against Impossible Differential cryptanalysis and Meet-in-the-Middle attacks.

4.4 Application to previous results

The literature on GFNs is full of long lists of good permutation candidates. Thus, we wanted to check whether these lists could be shortened by only considering one element per equivalence class. Our results are summarized in Table 1.

[3] For exemple, the GFN associated to [23, 28, 27, 0, 17, 4, 25, 26, 15, 2, 21, 24, 29, 30, 19, 6, 7, 12, 11, 16, 1, 20, 9, 10, 31, 18, 5, 8, 13, 14, 3, 22] has at least 66/64 differentially/linearly active S-boxes in 18 rounds.

The size reduction obviously leads to a more compact presentation of the results. But it also indicates that some computation could have been factorized by equivalence classes: for instance, in the case of the enumeration of [SSD+18] about alternative candidates of `TWINE`, the authors could have used the enumeration of Algorithm 1. In that case, they would have had only to test $8! = 40320$ candidates and the computation would have been 22 times faster, going from two hours to a few minutes. It would have even been approximately twice faster if the candidates were regrouped by extended expanded equivalence.

Source	Topic	Size of the list	Number of extended expanded equivalence classes
[CGT19]	Best-known permutations for GFNs with 32,64 or 128 branches regarding diffusion (regrouped by extended 1-cyclic equivalence)	32	10
[DFLM19]	Optimal permutations for even odd GFNs with 28 to 34 branches (regrouped by 1-cyclic equivalence)	19	9
[SSD+18]	Alternative permutations to improve the resistance of `LBlock` `TWINE` against Demirci-Selçuk Meet-in-the-Middle Attack.	64	2
[SSD+18]	Alternative permutations to improve the resistance of `TWINE` against Demirci-Selçuk Meet-in-the-Middle Attack.	12	1

Table 1: Reduction of the size of lists by using the expanded equivalence relation

5 Conclusion and perspectives

This paper brings new perspectives on GFNs and their permutations by considering bigger equivalence classes: many GFNs which were previously considered as different are actually cryptographically equivalent for a set of classical attacks. From a designer perspective, it reduces the space of GFN candidates and thus shrinks drastically the amount of time to compare their properties.

Finally, many open questions remain: Is there another representation of the permutations for which the expanded equivalence is an easy construction? Is there a more efficient way to compute the conjugacy classes? Can we characterise the classes which lead to good cryptographic properties? What is the size of the classes if we consider also non even-odd permutations? It may also have some implications for a cryptanalyst: for a given GFN, is there any equivalent GFN which is vulnerable to the attacks not taken into account here?

References

BBI+20. Subhadeep Banik, Zhenzhen Bao, Takanori Isobe, Hiroyasu Kubo, Fukang Liu, Kazuhiko Minematsu, Kosei Sakamoto, Nao Shibata, and Maki Shigeri. WARP : Revisiting GFN for lightweight 128-bit block cipher. In Orr Dunkelman, Michael J. Jacobson Jr., and Colin O'Flynn, editors, *SAC 2020*, volume 12804 of *LNCS*, pages 535–564. Springer, Heidelberg, October 2020.

CGT19. Victor Cauchois, Clément Gomez, and Gaël Thomas. General diffusion analysis: How to find optimal permutations for generalized type-II Feistel schemes. *IACR Trans. Symm. Cryptol.*, 2019(1):264–301, 2019.

DDGP22. Stéphanie Delaune, Patrick Derbez, Arthur Gontier, and Charles Prud'homme. New algorithm for exhausting optimal permutations for generalized feistel networks. In *International Conference on Cryptology in India*, pages 103–124. Springer, 2022.

DFLM19. Patrick Derbez, Pierre-Alain Fouque, Baptiste Lambin, and Victor Mollimard. Efficient search for optimal diffusion layers of generalized feistel networks. *IACR Trans. Symmetric Cryptol.*, 2019(2):218–240, 2019.

Nyb96. Kaisa Nyberg. Generalized feistel networks. In Kwangjo Kim and Tsutomu Matsumoto, editors, *Advances in Cryptology - ASIACRYPT '96, International Conference on the Theory and Applications of Cryptology and Information Security, Kyongju, Korea, November 3-7, 1996, Proceedings*, volume 1163 of *Lecture Notes in Computer Science*, pages 91–104. Springer, 1996.

S+77. Data Encryption Standard et al. Data encryption standard. *Federal Information Processing Standards Publication 46*, 1977.

SM10. Tomoyasu Suzaki and Kazuhiko Minematsu. Improving the generalized Feistel. In Seokhie Hong and Tetsu Iwata, editors, *FSE 2010*, volume 6147 of *LNCS*, pages 19–39. Springer, Heidelberg, February 2010.

SMMK13. Tomoyasu Suzaki, Kazuhiko Minematsu, Sumio Morioka, and Eita Kobayashi. TWINE : A lightweight block cipher for multiple platforms. In Lars R. Knudsen and Huapeng Wu, editors, *SAC 2012*, volume 7707 of *LNCS*, pages 339–354. Springer, Heidelberg, August 2013.

SSD+18. Danping Shi, Siwei Sun, Patrick Derbez, Yosuke Todo, Bing Sun, and Lei Hu. Programming the Demirci-Selçuk meet-in-the-middle attack with constraints. In Thomas Peyrin and Steven Galbraith, editors, *ASIACRYPT 2018, Part II*, volume 11273 of *LNCS*, pages 3–34. Springer, Heidelberg, December 2018.

WZ11. Wenling Wu and Lei Zhang. LBlock: A lightweight block cipher. In Javier Lopez and Gene Tsudik, editors, *ACNS 11*, volume 6715 of *LNCS*, pages 327–344. Springer, Heidelberg, June 2011.

ZMI90. Yuliang Zheng, Tsutomu Matsumoto, and Hideki Imai. On the construction of block ciphers provably secure and not relying on any unproved hypotheses. In Gilles Brassard, editor, *CRYPTO'89*, volume 435 of *LNCS*, pages 461–480. Springer, Heidelberg, August 1990.

Galois subcovers of the Hermitian curve in characteristic p with respect to subgroups of order dp with $d \neq p$ prime

Arianna Dionigi[1] and Barbara Gatti[2]

[1] Università di Firenze arianna.dionigi@unifi.it
[2] Università del Salento barbara.gatti@unisalento.it

Abstract. A problem of current interest, also motivated by applications to Coding theory, is to find explicit equations for *maximal* curves, that are projective, geometrically irreducible, non-singular curves defined over a finite field $\mathbb{F}_{q^2}$ whose number of $\mathbb{F}_{q^2}$-rational points attains the Hasse-Weil upper bound of $q^2+2\mathfrak{g}q+1$ where $\mathfrak{g}$ is the genus of the curve $\mathcal{X}$. For curves which are Galois covered of the Hermitian curve, this has been done so far ad hoc, in particular in the cases where the Galois group has prime order and also when has order the square of the characteristic. In this paper we obtain explicit equations of all Galois covers of the Hermitian curve with Galois group of order dp where p is the characteristic of $\mathbb{F}_{q^2}$ and d is prime other than p. We also compute the generators of the Weierstrass semigroup at a special $\mathbb{F}_{q^2}$-rational point of some of the curves, and discuss some possible positive impacts on the minimum distance problems of AG-codes.

Keywords: maximal curves, function fields, Galois cover, Weierstrass semigroup, AG-code

Subject classifications: 14H37, 14H05.

1 Introduction

Curves with many points over a finite field have intensively been investigated also by their connections to Coding theory, Cryptography, Finite geometry, and shift register sequences. In this context, the most important family consists of the maximal curves, that is, curves defined over the finite field $\mathbb{F}_{q^2}$, where $q = p^h$ and p is its charactreristic, which attain the famous Hasse-Weil upper bound. The Hermitian curve is the best known maximal curve and it is also the most useful for applications, especially in the study of algebraic geometry codes, shortly AG-codes. Actually, many other maximal curves derive from the Hermitian curve since any $\mathbb{F}_{q^2}$-subcover of a maximal curve is still maximal over the same field. If such a $\mathbb{F}_{q^2}$-subcover is a Galois subcover with Galois group G then the arising curve is named the quotient curve of the Hermitian curve with respect to G. Up to group isomorphism, the $\mathbb{F}_{q^2}$-automorphism group of the Hermitian curve is the 3-dimensional projective unitary group $\mathrm{PGU}(3, q)$ which has plenty of subgroups; see [9]. This motivated the systematic study of the quotients curves of the Hermitian curve which was eventually initiated in the seminal paper of Garcìa, Stichtenoth and Xing [6]. Ever since important progress has been made in the study of the spectrum of the possible genera of the quotients of the Hermitian curve over a given finite field; see [8, Chapter 10]. Nevertheless, the problem of determining explicit equations for such curves, which is a relevant issue for applications, remains largely open. In fact, this problem has so far been solved by ad hoc methods, apart form the cases where the Galois group has either prime order; see [2], or its order equals the square of

the characteristic; see [3].
In this paper we determine explicit equations for each quotient curve of the Hermitian curve whose Galois group has order dp where p is the characteristic of $\mathbb{F}_{q^2}$ and d is prime other than p. We also compute the Weierstrass semigroup at some $\mathbb{F}_{q^2}$-rational point of those curves, and discuss possible positive impacts on the minimum distance problems of AG-codes.

th1 **Theorem 1.** *In the $\mathbb{F}_{q^2}$-automorphism group $G \cong \mathrm{PGU}(3,q)$ of the Hermitian curve $\mathcal{H}_q$ defined over $\mathbb{F}_{q^2}$ with $q = p^h$ and $p \geq 5$, let H be a subgroup of order dp where $d \geq 5$ is a prime number other than p. Let $\bar{\mathcal{H}}_q = \mathcal{H}_q/H$ be the quotient curve of $\mathcal{H}_q$ with respect to the subgroup H. Then, up to an $\mathbb{F}_{q^2}$-isomorphism, one of the following cases occurs.*

(I) If $H = C_p \times C_d$ then $\bar{\mathcal{H}}_q$ has genus

$$\mathfrak{g} = \frac{1}{2d}(q-d+1)\left(\frac{q}{p}-1\right)$$

and equation

$$\sum_{i=0}^{h-1} Y^{p^i} + \omega X^{(q+1)/d} = 0 \quad \textit{with} \quad \omega^{q-1} = -1 \quad \textit{and} \quad d \mid (q+1). \tag{1}$$ eqthI

(II) If $H = C_p \rtimes C_d$ and C_p is in the center in a Sylow p-subgroup of G, then $\bar{\mathcal{H}}_q$ has genus

$$\mathfrak{g} = \frac{1}{2}\frac{q}{d}\left(\frac{q}{p}-1\right)$$

and equation

$$\omega X^{(q-1)/d} - A(X,Y) = 0 \quad \textit{with} \quad \omega^{q-1} = -1 \quad \textit{and} \quad d \mid (p-1). \tag{2}$$ eqthII

where

$$A(X,Y) = Y + X^{2(p-1)/d}Y^p + \cdots + X^{2(p^{h-1}-1)/d}Y^{q/p}.$$

(III) If $H = C_p \rtimes C_d$ but C_p is not in the center in a Sylow p-subgroup of G, then $\bar{\mathcal{H}}_q$ has genus

$$\mathfrak{g} = \frac{q}{2dp}(q-1)$$

and equation

$$\left(\frac{Y^2}{X^d}\right)^{(q-1)/d} + 1 - A(X,Y) = 0 \quad d \mid (p-1) \tag{3}$$ eqthIII

where

$$A(X,Y) = \sum_{i=0}^{h-1}\sum_{j=0}^{h-1}\left(\frac{Y^2}{X^d}\right)^{(p^i-1)/2d}\left(\frac{Y^2}{X^d}\right)^{(p^j-1)/2d} X^{(p^i+p^j)/2}.$$

Our Notation and terminology are standard; see [8,14,7]. In particular, q always stands for a power of p, namely $q = p^h$. We mostly use the language of function field theory rather than that of algebraic geometry.

2 Background

Let $\mathcal{X}$ be a projective, non-singular, geometrically irreducible, algebraic curve of genus $\mathfrak{g} \geq 2$ embedded in an r-dimensional projective space $\mathrm{PG}(r, \mathbb{F}_\ell)$ over a finite field of order ℓ of characteristic p. Let $\mathbb{F}_\ell(\mathcal{X})$ be the function field of $\mathcal{X}$ which is an algebraic function field of transcendency degree one with constant field $\mathbb{F}_\ell$. As it is customary, $\mathcal{X}$ is viewed as a curve defined over the algebraic closure $\mathbb{F}$ of $\mathbb{F}_\ell$. Then the function field $\mathbb{F}(\mathcal{X})$ is the constant field extension of $\mathbb{F}_\ell(\mathcal{X})$ with respect to field extension $\mathbb{F}|\mathbb{F}_\ell$. The automorphism group $\mathrm{Aut}(\mathcal{X})$ of $\mathcal{X}$ is defined to be the automorphism group of $\mathbb{F}(\mathcal{X})$ fixing every element of $\mathbb{F}$. It has a faithful permutation representation on the set of all points $\mathcal{X}$ (equivalently on the set of all places of $\mathbb{F}(\mathcal{X})$). The automorphism group $\mathrm{Aut}_{\mathbb{F}_\ell}(\mathcal{X})$ of $\mathbb{F}_\ell(\mathcal{X})$ is a subgroup of $\mathrm{Aut}(\mathcal{X})$. In particular, the action of $\mathrm{Aut}_{\mathbb{F}_\ell}(\mathcal{X})$ on the $\mathbb{F}_\ell$-rational points of $\mathcal{X}$ is the same as on the set of degree 1 places of $\mathbb{F}_\ell(\mathcal{X})$. Let G be a finite subgroup of $\mathrm{Aut}_{\mathbb{F}_\ell}(\mathcal{X})$. The *Galois subcover* of $\mathbb{F}_\ell(\mathcal{X})$ with respect to G is the fixed field of G, that is, the subfield $\mathbb{F}_\ell(\mathcal{X})^G$ consisting of all elements of $\mathbb{F}_\ell(\mathcal{X})$ fixed by every element in G. Let $\mathcal{Y}$ be a non-singular model of $\mathbb{F}_\ell(\mathcal{X})^G$, that is, a projective, non-singular, geometrically irreducible, algebraic curve with function field $\mathbb{F}_\ell(\mathcal{X})^G$. Then $\mathcal{Y}$ is the *quotient curve of* $\mathcal{X}$ *by* G and is denoted by $\mathcal{X}/G$. The covering $\mathcal{X} \mapsto \mathcal{Y}$ has degree equal to $\mid G \mid$ and the field extension $\mathbb{F}_\ell(\mathcal{X})|\mathbb{F}_\ell(\mathcal{X})^G$ is Galois. If P is a point of $\mathcal{X}$, the stabilizer G_P of P in G is the subgroup of G consisting of all elements fixing P.

resth11.49b **Result 2.** *[8, Theorem 11.49(b)] All p-elements of G_P together with the identity form a normal subgroup S_P of G_P so that $G_P = S_P \rtimes C$, the semidirect product of S_P with a cyclic complement C.*

resth11.129 **Result 3.** *[8, Theorem 11.129] If $\mathcal{X}$ has zero Hasse-Witt invariant then every non-trivial element of order p has a unique fixed point, and hence no non-trivial element in S_P fixes a point other than P.*

A useful corollary of Result 3 is the following.

lem15042023 **Result 4.** *Let $\mathcal{X}$ be an $\mathbb{F}_\ell$-rational curve whose number of $\mathbb{F}_\ell$-rational points is $N \geq 2$. If $\mathcal{X}$ has zero Hasse-Witt invariant and S is a p-subgroup of $\mathrm{Aut}_{\mathbb{F}_\ell}(\mathcal{X})$ then S fixes a unique point and $|S|$ divides $N - 1$.*

The following result is due to Stichtenoth [13].

sti1 **Result 5.** *[8, Theorem 11.78(i)] Let H be a p-subgroup of $\mathbb{F}(\mathcal{X})$ fixing a point. If $|S|$ is larger than the genus of $\mathbb{F}(\mathcal{X})$ then the Galois subcover of $\mathbb{F}(\mathcal{X})$ with respect to H is rational.*

From now on let $\ell = q^2$ with $q = p^h$ and assume that $\mathcal{X}$ is a $\mathbb{F}_{q^2}$-maximal curve.
The following result follows from [12, Lemma1].

zeroprank **Result 6.** *All $\mathbb{F}_{q^2}$-maximal curves have zero Hasse-Witt invariant.*

The following result is commonly attributed to Serre.

resth10.2 **Result 7.** *[8, Theorem 10.2] For every subgroup G of $\mathrm{Aut}_{\mathbb{F}_{q^2}}(\mathcal{X})$, the quotient curve $\mathcal{X}/G$ is also $\mathbb{F}_{q^2}$-maximal.*

We also use the classification of all groups whose order is the product of two distinct primes.

res07122023 **Result 8.** *Suppose u and v are distinct prime numbers with $u < v$. Then, there are two possibilities for groups G of order uv:*

(I) If $u \nmid (v-1)$ then G is a cyclic group.
(II) If $u \mid (v-1)$, then either G is a cyclic group, or G is a semidirect product $C_v \rtimes C_u$.

2.1 The function field of the Hermitian curve

In this subsection we collect some useful results about the function field of the Hermitian curve and its Galois subcovers. The affine equation $Y^q + Y = X^{q+1}$ of an $\mathbb{F}_{q^2}$-rational curve is the usual canonical form of the Hermitian curve $\mathcal{H}_q$ with function field is $\mathbb{F}_{q^2}(x, y)$ where $y^q+y-x^{q+1} = 0$. The equation $Y^q-Y+\omega X^{q+1} = 0$ with $\omega \in \mathbb{F}_{q^2}$ such that $\omega^{q-1} = -1$ is another useful equation of $\mathcal{H}_q$. We exploit numerous known results on the $\mathbb{F}$-automorphism group $\mathrm{Aut}(\mathbb{F}(\mathcal{H}_q))$ of $\mathcal{H}_q$. For more details, the Reader is referred to [9,7].

sect12.3

Result 9. *[8, Theorem 12.24(iv), Proposition 11.30]* $\mathrm{Aut}(\mathbb{F}(\mathcal{H}_q)) = \mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q)) \cong \mathrm{PGU}(3, q)$. *Moreover,* $\mathrm{Aut}(\mathbb{F}(\mathcal{H}_q))$ *acts on the set of all* $\mathbb{F}_{q^2}$*-rational points of* $\mathcal{H}_q$ *as* $\mathrm{PGU}(3, q)$ *in its natural doubly transitive permutation representation of degree* $q^3 + 1$ *on the isotropic points of the unitary polarity of the projective plane* $PG(2, \mathbb{F}_{q^2})$.

The maximal subgroups of $\mathrm{PGU}(3, q)$ were determined by Mitchell in 1911, see Hoffer [9]. Let S_p be a Sylow p-subgroup of $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q)) = \mathrm{PGU}(3, q)$. From Results 9, it may be assumed up to conjugacy that the unique fixed point of S_p is the point at infinity Y_∞ of $\mathcal{H}_q$. The following result describes the structure of the stabiliser of Y_∞ in $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q))$.

struct

Result 10. *Let the Hermitian function field be given by its canonical form* $\mathbb{F}_{q^2}(x, y)$ *with* $y^q + y - x^{q+1} = 0$. *Then the stabiliser* G *of* Y_∞ *in* $\mathrm{Aut}(\mathbb{F}(\mathcal{H}_q))$ *consists of all maps*

$$\psi_{a,b,\lambda} : (x, y) \mapsto (\lambda x + a, a^q\lambda x + \lambda^{q+1}y + b) \tag{4}$$

eq250523

where

$$a \in \mathbb{F}_{q^2},\ \lambda \in \mathbb{F}_{q^2}^*,\ b^q + b = a^{q+1}. \tag{5}$$

eqA250523

In particular, $G = S_p \rtimes C$ *where* $S_p = \{\psi_{a,b,1} | b^q + b = a^{q+1}, a, b \in \mathbb{F}_{q^2}\}$ *and* $C = \{\psi_{0,0,\lambda} | \lambda \in \mathbb{F}_{q^2}^*\}$.

A direct computation by induction on i shows that for $1 \le i \le p$

$$\psi_{a,b,1}^i = \psi_{ia,a^{q+1}(i^2-i)/2+ib,1}. \tag{6}$$

eq0812202

For more about Result 10 see [6], Section 4.

rem250523

Remark 11. Changes of the generators x, y of the Hermitian function field $\mathbb{F}_{q^2}(x, y), y^q + y - x^{q+1} = 0$ provide another canonical form. For our purpose, a useful change is $\tau : (x, y) \to (\omega x, -\omega y)$ where $\omega^{q-1} = -1$, and the arising canonical form is $y^q - y + \omega x^{q+1} = 0$. *Then the elements in the stabilizer* G *of* Y_∞ *in* $\mathrm{Aut}(\mathbb{F}(\mathcal{H}_q))$ *are of the form:*

$$\varphi_{a,b,\lambda} : (x, y) \mapsto (\lambda x + a, a^q\lambda\omega x + \lambda^{q+1}y + b) \tag{7}$$

eaQ081220

where (5) is replaced by

$$a \in \mathbb{F}_{q^2},\ \lambda \in \mathbb{F}_{q^2}^*,\ b^q - b = -\omega a^{q+1}.$$

A direct computation by induction on i *shows that for* $1 \le i \le p$

$$\varphi_{a,b,1}^i = \varphi_{ia,a^{q+1}\omega(i^2-i)/2+ib,1}. \tag{8}$$

eq0812202

For more about Result 11 see [6], Section 4.

From Result 10, S_p is the (unique) Sylow p-subgroup of the stabiliser of Y_∞ in $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q))$.

conjcl **Result 12.** *S_p has the following properties.*

(I) The center $Z(S_p)$ of S_p has order q and it consists of all maps $\psi_{0,b,1}$ with $b^q+b=0, b\in\mathbb{F}_{q^2}$. Also, $Z(S_p)$ is an elementary abelian group of order p.

(II) The non-trivial elements of S_p form two conjugacy classes in the stabiliser of Y_∞ in $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q))$, one comprises all non-trivial elements of $Z(S_p)$, the other does the remaining q^3-q elements.

(III) The elements of G other than those in $Z(S_p)$ have order p, or $p^2=4$ according as $p>2$ or $p=2$.

For completeness, we provide a proof for the classification of subgroups of $\mathrm{PGU}(3,q)$ of order dp. We use the canonical form $y^q-y+\omega x^{q+1}=0$ with $\omega^{q-1}=-1$. The Galois subcovers of $\mathbb{F}_{q^2}(\mathcal{H}_q)$ with respect to a subgroup H of prime order or when its order equals the square of the characteristic were thoroughly classified in [2] and [3] respectively. For the case $|H|=dp$, the classification is reported in the following result.

dp **Result 13.** *Let p and d two distinct prime numbers both larger than 3. Then, up to conjugacy in $\mathrm{PGU}(3,q)$, there exist at most three subgroups of order dp in $\mathrm{PGU}(3,q)$, one is cyclic and the other two are semidirect products of $C_p\rtimes C_d$ with $p<d$. They are subgroups of the stabiliser of Y_∞ in $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q))$ where*

(I) *$G=\Sigma_p\times\Sigma_d$ with $\Sigma_p=\langle\varphi_{0,1,1}\rangle$ and $\Sigma_d=\langle\varphi_{0,0,\lambda}\rangle$ with $\lambda^d=1$, $d|(q+1)$;*

(II) *$G=\Sigma_p\rtimes\Sigma_d$ with $\Sigma_p=\langle\varphi_{0,1,1}\rangle$ and $\Sigma_d=\langle\varphi_{0,0,\lambda}\rangle$ with $\lambda^d=1$, $d|(p-1)$.*

(III) *$G=\Sigma_p\rtimes\Sigma_d$ with $\Sigma_p=\langle\varphi_{1,\omega/2,1}\rangle$ and $\Sigma_d=\langle\varphi_{0,0,\lambda}\rangle$ with $\lambda^d=1$, $d|(p-1)$.*

Proof. Let G be a subgroup of order pd in $\mathrm{PGU}(3,q)$. Two cases are treated separately according as $p>d$ or $p<d$. Assume first $p>d$. Then Result 8 shows that G has a unique Sylow p-subgroup Σ_p. Moreover, Σ_p is a normal subgroup of G, and hence $G=\Sigma_p\rtimes\Sigma_d$ where Σ_d is a Sylow d-subgroup of G. Since any non-trivial element of $\mathrm{PGU}(3,q)$ of order p has exactly one fixed point on $\mathcal{H}_q(\mathbb{F}_{q^2})$ whereas $\mathrm{PGU}(3,q)$ acts transitively on $\mathcal{H}_q(\mathbb{F}_{q^2})$, we may assume, up to conjugacy in $\mathrm{PGU}(3,q)$, that Y_∞ is the unique fixed point of S_p. As S_p is a normal subgroup of G, the point Y_∞ is also fixed by Σ_d. From $|\mathcal{H}_q(\mathbb{F}_{q^2})|-1=q^3$, Σ_d must have a fixed point $O\in\mathcal{H}_q(\mathbb{F}_{q^2})$ other than Y_∞. Since $\mathrm{PGU}(3,q)$ is doubly transitive on $\mathcal{H}_q(\mathbb{F}_{q^2})$ we may assume, up to conjugacy, that $O=(0:0:1)$. Then Σ_d is generated by $t=\varphi_{0,0,\lambda}$ with $\lambda^d=1$ where $d|(q^2-1)$. Furthermore, as Σ_p is a subgroup of the Sylow subgroup S_p of $\mathrm{PGU}(3,q)$ fixing Y_∞, two cases arise according as Σ_p is in the center $Z(S_p)$ of S_p or not. Let s be a generator of Σ_p. If $s\in Z(S_p)$ then $s=\varphi_{0,b,1}$ with $b^q-b=0$. Take $\mu\in\mathbb{F}_{q^2}^*$ such that $\mu^{q+1}=b^{-1}$. Then the conjugate of s by $\varphi_{0,0,\mu}$ is $\varphi_{0,0,1}$ while t and $\varphi_{0,0,\mu}$ commute. Therefore, up to conjugacy, $G=\Sigma_p\rtimes\Sigma_d$ with $\Sigma_p=\langle s\rangle$ and $s=\varphi_{0,0,1}$ whereas Σ_d and t are as before. Since Σ_p is a normal subgroup of G, there exists i with $1\le i\le p-1$ such that $st=ts^i$. A straightforward computation shows that this occurs if and only if $i=1/\lambda^{q+1}$. For $d|(q+1)$, this implies $i=1$, thus G is cyclic and Case (I) occurs. For $d|(q-1)$, we have $i\neq 1$ and hence G is not abelian. From Result 8, $d|(p-1)$. Thus Case (II) occurs. If $s\notin Z(S_p)$ then $s=\varphi_{a,b,1}$. For $\mu=a^{-1}$, the conjugate of s by $\varphi_{0,0,\mu}$ is $\varphi_{1,b/a^{q+1},1}$ while t and $\varphi_{0,0,\mu}$ commute. Therefore, up to conjugacy, we may assume $G=\Sigma_p\rtimes\Sigma_d$ where $\Sigma_p=\langle s\rangle$ and $s=\varphi_{1,b/a^{q+1},1}$ while Σ_d and t are not changed. Then $st=\phi_{1,b/a^{q+1},\lambda}$ and, from (6), $ts^i=\phi_{\lambda i,\lambda^{q+1}(\omega(i^2-i)/2+ib/a^{q+1}),\lambda}$. Therefore, $st=ts^i$ if and only if $\lambda i=1$ and $\lambda^{q+1}(\frac{1}{2}\omega(i^2-i)+ib/a^{q+1})=b/a^{q+1}$. The latter condition can also be written as $\frac{1}{2}\omega(i^2-i)=(i^2-i)b/a^{q+1}$, that is, $b=\frac{1}{2}\omega a^{q+1}$ as $\lambda\neq 1$. Therefore, $st=ts^i$ if and only if $s=\varphi_{1,\omega/2,1}$ and $i\lambda=1$. In particular, G is not abelian, and $d|(p-1)$. This gives Case (III). Now, assume $p<d$. Then a Sylow d-subgroup Σ_d of G is a normal subgroup of G, and hence Σ_d is the unique d-subgroup of G. As d divides the order of $\mathrm{PGU}(3,q)$, either $d|(q-1)$, or $d|(q+1)$, or $d|(q^2-q+1)$. Assume that Σ_d fixes a point on $\mathcal{H}_q(\mathbb{F}_{q^2})$. Then Σ_d has at least two fixed points, as $|\mathcal{H}_q(\mathbb{F}_{q^2})|-1$ equals q^3. Up to conjugacy, we may assume that Σ_d fixes Y_∞ and O. Then $\Sigma_d=\langle\varphi_{0,0,\lambda}\rangle$ with $\lambda^d=1$. If $\lambda^{q+1}\neq 1$ then Σ_d has no

any further fixed point, and hence G preserves the pair $\{Y_\infty, O\}$. Since $p > 2$ this yields that elements of G of order p fix two points on $\mathcal{H}_q(\mathbb{F}_{q^2})$ which is not possible. Therefore, $\lambda^{q+1} = 1$ and $d|(q+1)$. This yields that Σ_d fixes all points $P = (0, \eta)$ with $\eta^q - \eta = 0$, i.e. with $\eta \in \mathbb{F}_q$. Since Σ_d is a normal subgroup of G, this yields that a generator s of Σ_p takes O to a point $P = (0, \eta)$ with $\eta \in \mathbb{F}_q$. But then $s = \varphi_{0,b,1}$ with $b \in \mathbb{F}_q^*$. For $\mu = b^{-1}$, the conjugate of s by $\varphi_{0,0,\nu}$ with $\nu^{q+1} = \mu$ is $\varphi_{0,1,1}$ while a generator t of Σ_d and $\varphi_{0,0,\mu}$ commute. Therefore, up to conjugacy, we may assume $s = \varphi_{0,1,1}$. Also, $st = ts$ and Case (I) occurs. We are left with the case where Σ_d fixes no point on $\mathcal{H}_q(\mathbb{F}_{q^2})$. Then either $d|(q+1)$, or $d|(q^2-q+1)$. We look at the action of $\mathrm{PGU}(3,q)$ as a projective group of the plane $\mathrm{PG}(2,\mathbb{K})$ where $\mathbb{K}$ is an algebraic closure of $\mathbb{F}_{q^2}$. Then the Hermitian curve $\mathcal{H}_q$ is left invariant by $\mathrm{PGU}(3,q)$. In particular, $\mathrm{PGU}(3,q)$ preserves both $\mathcal{H}_q(\mathbb{F}_{q^2})$ and its complementary set in $\mathrm{PG}(2,\mathbb{F}_{q^2})$ whose size equals $q^4+q^2+1-(q^3+1) = q^2(q^2-q+1)$. Furthermore, $\mathcal{H}_q(\mathbb{F}_{q^2})$ can also be viewed as the set of all isotropic points of a unitary polarity π of $\mathrm{PG}(2,\mathbb{F}_{q^2})$. If $d|(q+1)$ then Σ_d fixes a point $R \in \mathrm{PG}(2,\mathbb{F}_{q^2})$ outside $\mathcal{H}_q(\mathbb{F}_{q^2})$. Let r be the polar line of R w.r.t. π. Then r is a chord of $\mathcal{H}_q(\mathbb{F}_{q^2})$. Since r has as many as $q(q-1)$ points other than those on $\mathcal{H}_q(\mathbb{F}_{q^2})$, there are at least two fixed points on r outside $\mathcal{H}_q(\mathbb{F}_{q^2})$ under the action of Σ_d. Since Σ_d does not fix r pointwise, these two points, say R_1, R_2 are the only fixed points of Σ_d on r. In particular, Σ_d fixes the vertices of the triangle RR_1R_2. We show that no more point in $\mathrm{PG}(2,\mathbb{F}_{q^2})$ is fixed by Σ_d. In fact, such a further fixed point T of Σ_d should lie on a side of the triangle, and that side would be fixed pointwise by Σ_d. But this is impossible in our case, since the sides of RR_1R_2 are chords of $\mathcal{H}_q(\mathbb{F}_{q^2})$ whereas Σ_d is supposed not to fix points on $\mathcal{H}_q(\mathbb{F}_{q^2})$. Since Σ_d is a normal subgroup of G, the triangle RR_1R_2 is left invariant by G. But then G is a contained in a maximal subgroup of $\mathrm{PGU}(3,q)$ whose order equals $6(q+1)^2$. Since $p > 3$, this is impossible. A similar geometric approach is used to rule out the other possibility, i.e. $d|(q^2-q+1)$. Look at the action of $\mathrm{PGU}(3,q)$ on $\mathrm{PG}(2,\mathbb{F}_{q^6})$. From $|\mathcal{H}_q(\mathbb{F}_{q^6})| = q^6+1+q^4(q-1)$ and $|\mathcal{H}_q(\mathbb{F}_{q^3})| = q^3+1$, the Hermitian curve $\mathcal{H}_q$ has as many as $q^3(q+1)^2(q-1)$ points in $\mathrm{PG}(2,\mathbb{F}_{q^6})$ but not in $\mathrm{PG}(2,\mathbb{F}_{q^2})$. From $d|(q^2-q+1)$ and $d > 3$, Σ_d fixes a point $R \in \mathcal{H}_q(\mathbb{F}_{q^6})$ not lying in $PG(2,\mathbb{F}_{q^2})$. The Frobenius collineation $\mathfrak{f}$ which sends the point $P = (a_1 : a_2 : a_3)$ to the point $P_{q^2} = \left(a_1^{q^2} : a_2^{q^2} : a_3^{q^2}\right)$ leaves $\mathcal{H}_q(\mathbb{F}_{q^6})$ invariant. Since $\mathfrak{f}$ and Σ_d commute, Σ_d also fixes the points R_{q^2} and R_{q^3}. Actually, Σ_d does not fix another point, otherwise one of the sides, say ℓ, of the triangle $RR_{q^2}R_{q^4}$ would be fixed by Σ_d pointwise. Since $\mathfrak{f}$ takes ℓ to another side r of $RR_{q^2}R_{q^4}$ and $\mathfrak{f}$ and Σ_d commute, this would yield that r is also fixed pointwise by Σ_d, which is impossible. As before, this implies that G leaves the triangle invariant $RR_{q^2}R_{q^4}$ invariant. Therefore G is a contained in a maximal subgroup of $\mathrm{PGU}(3,q)$ whose order equals $3(q^2-q+1)$. Since $p > 3$, this is impossible.

ckt

Result 14. *[8, Theorem 5.74] Let H be a subgroup of $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q))$ of order p. The Galois subcover $\mathbb{F}_{q^2}(\mathcal{F}'))$ of $\mathbb{F}_{q^2}(\mathcal{H}_q)$ with respect to H is $\mathbb{F}_{q^2}$-isomorphic to the function field $\mathbb{F}_{q^2}(\xi,\eta)$ where either* (I) or (II) hold:

(I) $\sum_{i=1}^{h} \eta^{q/p^i} + \omega\xi^{q+1} = 0$ with $\omega^{q-1} = -1$, $\mathfrak{g}(\mathbb{F}_{q^2}(\mathcal{F}')) = \frac{1}{2}q\left(\frac{q}{p}-1\right)$, and H is in the center of a Sylow p-subgroup of $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q))$;

(II) $\eta^q + \eta - (\sum_{i=1}^{h} \xi^{q/p^i})^2 = 0$ for $p > 2$, $\mathfrak{g}(\mathbb{F}_{q^2}(\mathcal{F}')) = \frac{1}{2}\frac{q}{p}(q-1)$, and H is not in the center of a Sylow p-subgroup of $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q))$.

The following result is a corollary of [6, Section 4].

generi

Result 15. *Let $\mathfrak{g}$ be the genus of the Galois cover of $\mathbb{F}_{q^2}(\mathcal{H}_q)$ with respect to a subgroup G of $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}_q))$ of order dp. Let S_p is a Sylow p-subgroup of $\mathrm{Aut}(\mathbb{F}_{q^2}(\mathcal{H}))$ containing a subgroup H of G of order p. Then either*

$$\mathfrak{g} = \frac{1}{2}\frac{q}{d}\left(\frac{q}{p}-1\right) \qquad \textit{for} \qquad (d, q+1) = 1,$$

or

$$\mathfrak{g} = \frac{1}{2d}(q-d+1)\left(\frac{q}{p}-1\right) \qquad \textit{for} \qquad (d, q+1) = d.$$

3 Galois subcovers of $\mathbb{F}_{q^2}(\mathcal{H})$ of type (I) of Result (13)

typeI

As in Remark 11, take $\mathbb{F}_{q^2}(\mathcal{H}_q)$ in its canonical form $\mathbb{F}_{q^2}(x,y)$ with $y^q - y + \omega x^{q+1} = 0$ and $\omega^{q-1} = -1$. The group $\Phi = \langle \varphi_{0,1,1} \rangle$ has order p, and it is contained in $Z(S_p)$. Let $\eta = y^p - y$ and $\xi = x$. Then $\varphi_{0,1,1}(\eta) = \varphi_{0,1,1}(y^p - y) = \varphi_{0,1,1}(y)^p - \varphi_{0,1,1}(y) = (y+1)^p - (y+1) = y^p - y = \eta$. Moreover, $y^q - y = Tr(\eta)$. Since $\varphi_{0,1,1}$ fixes ξ, this shows that the Galois subcover $\mathbb{F}_{q^2}(\mathcal{F}')$ of $\mathbb{F}_{q^2}(\mathcal{H}_q)$ with respect to Φ is as in (i) of Result 14. That equation can also be written as

$$\sum_{i=0}^{h-1} \eta^{p^i} + \omega \xi^{q+1} = 0. \tag{9}$$

eqiia

Take an element $r \in \mathbb{F}_{q^2}$ with $r^d = 1$. Then $\varphi_{0,0,r}$ commutes with $\varphi_{0,1,1}$. Therefore, if $d|(q+1)$ then $\varphi_{0,0,r}$ induces an automorphism φ of $\mathbb{F}_{q^2}(\mathcal{F}')$. More precisely, a straightforward computation shows that φ is the map $\varphi : (\xi, \eta) \mapsto (r\xi, \eta)$. Let Φ_r be the $\mathbb{F}_q$-automorphism group of $\mathbb{F}_{q^2}(\mathcal{F}')$ generated by φ. Then the Galois subcover of $\mathbb{F}_{q^2}(\mathcal{H}_q)$ with respect to G of Result (13) of type (I) is the same as the Galois subcover G_r of $\mathbb{F}_{q^2}(\mathcal{F}')$ with respect to Φ_r.

propiia

Theorem 16. *The Galois subcover $G_r = \mathbb{F}_{q^2}(\zeta, \tau)$ of $\mathbb{F}_{q^2}(\mathcal{F}')$ with respect to Φ_r has genus*

$$\mathfrak{g} = \frac{1}{2d}(q-d+1)(\frac{q}{p}-1)$$

and is given by

$$\sum_{i=0}^{h-1} \tau^{p^i} + \omega \zeta^{q+1/d} = 0, \quad d \mid (q+1). \tag{10}$$

eq1

Proof. We show first that the fixed field F of Φ_r is generated by $\tau = \eta$ together with

$$\zeta = \xi^d. \tag{11}$$

eq1iia

Since $\varphi(\tau) = \tau$ and

$$\varphi(\zeta) = \varphi(\xi^d) = \varphi(\xi)^d = r^d \xi^d = \xi^d = \zeta,$$

we have $\mathbb{F}_{q^2}(\zeta, \tau) \subseteq F$. Furthermore, $[\mathbb{F}_{q^2}(\mathcal{F}') : \mathbb{F}_{q^2}(\zeta, \tau)] = d$. Since d is prime, this yields either $\mathbb{F}_{q^2}(\zeta, \tau) = F$ or $F = \mathbb{F}_{q^2}(\mathcal{F}')$. The latter case cannot actually occur, and hence $F = \mathbb{F}_{q^2}(\zeta, \tau)$. Therefore $F = G_r$. Now, eliminate ξ from Equations (9) and (11). Since d divides $q+1$, replacing ξ^{q+1} with $(\xi^{q+1/d})^d$ and $\tau = \eta$ in (9) gives equation in (10). The formula for the genus follows from [8, Lemma 12.1(iii)(b)].

4 Galois subcovers of $\mathbb{F}_{q^2}(\mathcal{H})$ of type (II) of Result (13)

We keep our notation up from Section 3. Assume that d divides $p-1$, and take $r \in \mathbb{F}_p^*$ with $r^d = 1$. Then $\varphi_{0,0,r}^{-1} \circ \varphi_{0,1,1} \circ \varphi_{0,0,r} = \varphi_{0,r^2,1} \in \langle \varphi_{0,1,1} \rangle$, and hence $\varphi_{0,0,r}$ induces an automorphism φ of $\mathbb{F}_{q^2}(\mathcal{F}')$. Here, φ is the map $\varphi : (\xi, \eta) \mapsto (r\xi, r^2\eta)$. Let Φ_r be the $\mathbb{F}_q$-automorphism group of $\mathbb{F}_{q^2}(\mathcal{F}')$ generated by φ. Then the Galois subcover of $\mathbb{F}_{q^2}(\mathcal{H}_q)$ with respect to G of Result (13) of type (II) is the Galois subcover G_r of $\mathbb{F}_{q^2}(\mathcal{F}')$ with respect to Φ_r.

Theorem 17. *The Galois subcover $G_r = \mathbb{F}_{q^2}(\epsilon, \rho)$ of $\mathbb{F}_{q^2}(\mathcal{F}')$ with respect to Φ_r has equation*

$$\omega\epsilon^{(q-1)/d} - A(\epsilon, \rho) = 0, \quad d \mid (p-1) \tag{12}$$ eq2

where

$$A(\epsilon, \rho) = \rho + \epsilon^{2(p-1)/d}\rho^p + \cdots + \epsilon^{2(p^{h-1}-1)/d}\rho^{q/p}.$$

Proof. We show first that the fixed field F of Φ_r is generated by

$$\epsilon = \xi^d \tag{13}$$ inv1

together with

$$\rho = \frac{\eta}{\xi^2}. \tag{14}$$ inv2

Since

$$\varphi(\epsilon) = \varphi(\xi^d) = \varphi(\xi)^d = r^d\xi^d = \xi^d = \epsilon$$

and

$$\varphi(\rho) = \frac{\varphi(\eta)}{\varphi(\xi^2)} = \frac{\varphi(\eta)}{\varphi(\xi)^2} = \frac{r^2\eta}{r^2\xi^2} = \frac{\eta}{\xi^2} = \rho,$$

we have $\mathbb{F}_{q^2}(\epsilon, \rho) \subseteq F$. Furthermore, $[\mathbb{F}_{q^2}(\mathcal{F}') : \mathbb{F}_{q^2}(\epsilon, \rho)] = d$. Since d is prime, this yields either $\mathbb{F}_{q^2}(\epsilon, \rho) = F$ or $F = \mathbb{F}_{q^2}(\mathcal{F}')$. The latter case cannot actually occur, and hence $F = \mathbb{F}_{q^2}(\epsilon, \rho)$. Therefore $F = G_r$. We have to eliminate ξ and η from equations (9), (13) and (14). From (14) we have $\eta = \rho\xi^2$ then $Tr(\eta) = Tr(\rho\xi^2)$. This yields that

$$Tr(\eta) = \xi^2\rho + \xi^{2p}\rho^p + \cdots + \xi^{2q/p}\rho^{q/p}, \tag{15}$$ pass1

whence

$$Tr(\eta) = \xi^2\big(\rho + \xi^{2(p-1)}\rho^p + \cdots + \xi^{2(q/p-1)}\rho^{q/p}\big). \tag{16}$$ pass2

Since d divides $p-1$, $Tr(\eta)$ in (15) can also be written as

$$Tr(\eta) = \xi^2\big(\rho + (\xi^d)^{2(p-1)/d}\rho^p + \cdots + (\xi^d)^{2(p^{h-1}-1)/d}\rho^{q/p}\big). \tag{17}$$ pass3

Therefore

$$Tr(\eta) = \xi^2\big(\rho + \epsilon^{2(p-1)/d}\rho^p + \cdots + \epsilon^{2(p^{h-1}-1)/d}\rho^{q/p}\big) = \xi^2 A(\epsilon, \rho). \tag{18}$$ pass4

This, together with (9), give

$$\omega\xi^{q+1} = \xi^2 A(\epsilon, \eta). \tag{19}$$ pass5

Since $d \mid (p-1)$ the number $\frac{q-1}{p-1}$ is an integer. Thus Equation (12) follows from (19).

5 Galois subcovers of $\mathbb{F}_{q^2}(\mathcal{H})$ of type (III) of Result (13)

This time, take $\mathbb{F}_{q^2}(\mathcal{H}_q)$ in its canonical form $\mathbb{F}_{q^2}(x, y)$ with $y^q + y - x^{q+1} = 0$. The group $\Psi = \langle\psi_{1,1/2,1}\rangle$ has order p, and it is not contained in $Z(S_p)$. Let $\xi = x^p - x$ and $\eta = y - \frac{1}{2}x^2$. A straightforward computation shows that $\psi_{1,1/2,1}(\xi) = \xi$ and $\psi_{1,1/2,1}(\eta) = \eta$. Moreover,

$$y^q + y - x^{q+1} = \eta^q + \tfrac{1}{2}x^{2q} + \eta + \tfrac{1}{2}x^2 - x^{q+1} = \eta^q + \eta + \tfrac{1}{2}x^{2q} + \tfrac{1}{2}x^2 - x^{q+1} = \eta^q + \eta + \tfrac{1}{2}(x^q - x)^2.$$

Since $Tr(\xi) = x^q - x$, this gives

$$\eta^q + \eta + \tfrac{1}{2}(x^q - x)^2 = \eta^q + \eta + \tfrac{1}{2}Tr(\xi)^2.$$

Therefore, the Galois subcover $\mathbb{F}_{q^2}(\mathcal{F}')$ of $\mathbb{F}_{q^2}(\mathcal{H}_q)$ with respect to Ψ is $\mathbb{F}_{q^2}(\xi,\eta)$ with

$$\eta^q + \eta + \tfrac{1}{2}\left(\sum_{i=1}^{h} \xi^{p^{i-1}}\right)^2 = 0. \tag{20}$$ eqiib

In particular, $\mathbb{F}_{q^2}(\mathcal{F}')$ is $\mathbb{F}_{q^2}$-isomorphic to (II) of Result 14. Assume that d divides $p-1$, and take $r \in \mathbb{F}_p^*$ with $r^d = 1$. Then $\psi_{0,0,r}^{-1} \circ \psi_{1,1/2,1} \circ \psi_{0,0,r} = \varphi_{r,1/2r^2,1} \in \langle \varphi_{1,1/2,1} \rangle$, and hence $\varphi_{0,0,r}$ induces an automorphism φ of $\mathbb{F}_{q^2}(\mathcal{F}')$. Moreover, ψ is the map $\psi : (\xi,\eta) \mapsto (r\xi, r^2\eta)$. Let Ψ_r be the $\mathbb{F}_q$-automorphism group of $\mathbb{F}_{q^2}(\mathcal{F}')$ generated by ψ. Then the Galois subcover of $\mathbb{F}_{q^2}(\mathcal{H}_q)$ with respect to G of Result (13) of type (III) is the Galois subcover G_r of $\mathbb{F}_{q^2}(\mathcal{F}')$ with respect to Ψ_r.

propiiaA **Theorem 18.** *The Galois subcover $G_r = \mathbb{F}_{q^2}(\iota,\nu)$ of $\mathbb{F}_{q^2}(\mathcal{F}')$ with respect to Ψ_r has genus*

$$\mathfrak{g} = \frac{q}{2dp}(q-1)$$

and is given by

$$\left(\frac{\tau^2}{\iota^d}\right)^{(q-1)/d} + 1 - A(\iota,\tau) = 0 \tag{21}$$ eq3

where

$$A(\iota,\tau) = \sum_{i=0}^{h-1}\sum_{j=0}^{h-1} \left(\frac{\tau^2}{\iota^d}\right)^{(p^i-1)/2d} \left(\frac{\tau^2}{\iota^d}\right)^{(p^j-1)/2d} \iota^{(p^i+p^j)/2}.$$

Proof. We show first that the fixed field F of Ψ_r is generated by

$$\nu = \eta^d, \tag{22}$$ inv3

together with

$$\iota = \frac{\xi^2}{\eta}, \tag{23}$$ inv4

and

$$\tau = \xi^d. \tag{24}$$ inv5

Since

$$\varphi(\nu) = \varphi(\eta)^d = r^{2d}\eta^d = \eta^d = \nu, \quad \varphi(\tau) = \varphi(\xi)^d = r^d\xi^d = \xi^d = \tau$$

and

$$\varphi(\iota) = \varphi\left(\frac{\xi^2}{\eta}\right) = \frac{\varphi(\xi^2)}{\varphi(\eta)} = \frac{\varphi(\xi)^2}{\varphi(\eta)} = \frac{(r\xi)^2}{r^2\eta} = \frac{r^2\xi^2}{r^2\eta} = \frac{\xi^2}{\eta},$$

we have $\mathbb{F}_{q^2}(\iota,\nu,\tau) \subseteq F$. Furthermore, $[\mathbb{F}_{q^2}(\iota,\nu,\tau)(\xi) : \mathbb{F}_{q^2}(\iota,\nu,\tau)] = d$ and $\eta \in \mathbb{F}_{q^2}(\iota,\nu,\tau)(\xi)$. Therefore, $[\mathbb{F}_{q^2}(\mathcal{F}') : \mathbb{F}_{q^2}(\iota,\nu,\tau)] \le d$. Since d is prime, this yields either $\mathbb{F}_{q^2}(\iota,\mu,\nu) = F$ or $F = \mathbb{F}_{q^2}(\mathcal{F}')$. The latter case cannot actually occur, and hence $F = \mathbb{F}_{q^2}(\iota,\mu,\nu)$. Therefore $F = G_r$. We go on by eliminating ξ and

η from Equations (20), (22), (23) and (24). From the definition of the trace of ξ, $Tr(\xi)^2 = (\xi + \cdots + \xi^{q/p})^2$. By a straightforward computation,

$$Tr(\xi)^2 = \sum_{i=0}^{h-1}\sum_{j=0}^{h-1} \xi^{p^i+p^j}.$$

This can also be written as

$$Tr(\xi)^2 = \sum_{i=0}^{h-1}\sum_{j=0}^{h-1} (\xi^2)^{(p^i+p^j)/2}. \tag{25}$$ pass6

From (23), $\xi^2 = \eta\iota$. Therefore, in (25) the square trace of ξ is equal to

$$\sum_{i=0}^{h-1}\sum_{j=0}^{h-1} (\eta\iota)^{(p^i+p^j)/2}. \tag{26}$$ pass7

Since $\frac{1}{2}(p^i + p^j) - 1 = \frac{1}{2}(p^i - 1) + \frac{1}{2}(p^j - 1)$, the sum in (26) turns out to be equal to

$$\eta \sum_{i=0}^{h-1}\sum_{j=0}^{h-1} \eta^{(p^i-1)/2}\eta^{(p^j-1)/2}\iota^{(p^i+p^j)/2}. \tag{27}$$ pass7bis

As d divides $\frac{1}{2}(p^i + p^j - 2)$ and 2 divides both $p^i - 1$ and $p^j - 1$, the sum in (26) equals

$$\eta \sum_{i=0}^{h-1}\sum_{j=0}^{h-1} (\eta^d)^{(p^i-1)/2d}(\eta^d)^{(p^j-1)/2d}\iota^{(p^i+p^j)/2} \tag{28}$$ pass8

by replacing η^d with ν

$$\eta \sum_{i=0}^{h-1}\sum_{j=0}^{h-1} \nu^{(p^i-1)/2d}\nu^{(p^j-1)/2d}\iota^{(p^i+p^j)/2} \tag{29}$$ pass9

Let

$$A(\iota, \nu) = \sum_{i=0}^{h-1}\sum_{j=0}^{h-1} \nu^{(p^i+p^j-2)/2d}\iota^{(p^i+p^j)/2}. \tag{30}$$ pass10

Therefore, $\eta^q + \eta = \eta A(\iota, \nu)$, and dividing both sides by η gives $\eta^{q-1} + 1 = A(\iota, \nu)$. Since d divides $q - 1$, replacing η^d by ν shows

$$\nu^{(q-1)/d} + 1 - A(\iota, \nu) = 0. \tag{31}$$ nuiota

From (22), (23), and (24), $\nu = \frac{\tau^2}{\iota^d}$. Now the claim follows from (31).

6 Weierstrass semigroups and application to AG-codes

We compute the Weierstrass semigroup at the unique place centred at the point at infinity of some of the maximal curves considered in the present paper.

WS **Proposition 19.** *Let P_∞ be the unique point at infinity of the following two curves*

$$\sum_{i=1}^{h} Y^{q/p^i} + \omega X^{q+1} = 0, \quad \omega^{q-1} = -1, h \geq 2; \qquad \sum_{i=1}^{h} Y^{q/p^i} + \omega X^{(q+1)/d} = 0, \omega^{q-1} = -1, d \mid (q+1). \quad (32)$$ INTERM1

Then the Weierstrass semigroup at P_∞ is generated by $\frac{q}{p}$, and $q+1$, respectively by $\frac{q}{p}$, and $\frac{q+1}{d}$.

Proof. Fore the first equation the claim follows from the remark after the proof of Lemma 12.2 in [8] applied for $n = h-1$ and $m = q+1$.
For the second equation the claim follows from the remark after the proof of Lemma 12.2 in [8] applied for $n = h-1$ and $m = \frac{q+1}{d}$.

Let S be a numerical semigroup. The gaps of S are the elements in $\mathbb{N} \setminus S$. The number $g(S)$ of gaps of S is the *genus* of S. If S is the Weierstrass semigroup of a curve at a point then $g(S)$ coincides with the genus of the curve. Let $(a_1, \ldots, a_k)$ be a sequence of positive integers such that their greatest common divisor is 1. Let $d_0 = 0$, $d_i = g.c.d.(a_1, \ldots, a_i)$ and $A_i = \left\{ \frac{a_1}{d_i}, \ldots, \frac{a_i}{d_i} \right\}$ for $i = 1, \ldots, k$. Let S_i be the semigroup generated by A_i. The sequence $(a_1, \ldots, a_k)$ is *telescopic* whenever $\frac{a_i}{d_i} \in S_{i-1}$ for $i = 2, \ldots, k$. A *telescopic semigroup* is a numerical semigroup generated by a telescopic sequence.

heresemigroup **Result 20.** *[10, Lemma 6.5] For the semigroup generated by a telescopic sequence $(a_1, \ldots, a_k)$, let*

$$l_g(S_k) := \sum_{i=1}^{k} \left(\frac{d_{i-1}}{d_i} - 1 \right) a_1, \quad g(S_k) := \frac{l_g(S_k) + 1}{2}.$$

WS1 **Theorem 21.** *Let P_∞ be the unique point of infinity of the curves $\bar{\mathcal{H}}_q$ in Theorem 1. Then the Weierstrass semigroup $H(P_\infty)$ has the following properties:*

(I) $H(P_\infty) = \langle \frac{q}{p}, q+1 \rangle$, *for the curve of Equation (I);*
(II) $\frac{q}{p}, \frac{q-1}{d} \in H(P_\infty)$, *for the curve of Equation (II);*
(III) $\frac{2(q-1)}{d}, q-1 \in H(P_\infty)$, *for the curve of Equation (III).*

Proof. Case (i). The pole numbers of x and y at P_∞ are q and $2q/p$, respectively. Since the curve is $\mathbb{F}_{q^2}$-maximal and P_∞ is an $\mathbb{F}_{q^2}$-rational point, $q+1 \in H(P_\infty)$; see [8, Theorem 10.6]. Let $d_0 = 0$, $d_1 = 2\frac{q}{p}$, $d_2 = \frac{2}{q}$ and $d_3 = 1$, and $A_1 = \{1\}$, $A_2 = \{2, p\}$, $A_3 = \{2\frac{q}{p}, q, q+1\}$. Then $p \in S_1$ and $q+1 \in S_2$. Thus the sequence $\{2\frac{q}{p}, q, q+1\}$ is telescopic. Furthermore,

$$l_g(S_3) = -\frac{2q}{p} + q + (\frac{q}{p} - 1)(q+1) = \frac{q^2}{p} - \frac{q}{p} - 1,$$

whence the claim follows by Result (20).

Case (ii). From Equation (II), $\left[\mathbb{F}_{q^2}(\bar{\mathcal{H}}_q) : \mathbb{F}_{q^2}(x)\right] = \frac{q}{p}$ and $\left[\mathbb{F}_{q^2}(\bar{\mathcal{H}}_q) : \mathbb{F}_{q^2}(y)\right] = \frac{q-1}{d}$. Therefore, $\frac{q}{p}$ and $\frac{q-1}{d}$ are non-gaps at P_∞.

Case (iii). The above argument applied to the curve $\bar{\mathcal{H}}_q$ of Equation (III), shows that $\frac{q-1}{d}$ and $\frac{q}{p}$ are non-gaps of G_r at P_∞.

Let C denote any $\mathbb{F}_{q^2}$-maximal curve equipped with an $\mathbb{F}_{q^2}$-rational point P. Let D be a set of $\mathbb{F}_{q^2}$-rational points of C other than P. From previous work by Janwa [11] and Garcìa-Kim-Lax [4], if the divisor G is taken as multiple of P then knowledge of the gaps at P_∞ may allow one to show that the minimum distance of the resulting evaluation code $C_L(G,D)$ or differential code $C_\Omega(G,D)$ may be better than the designed minimum distance of that code. In particular, it is shown in [5] that t consecutive gaps at P (under some conditions on the order sequence at P) gives a minimum distance d of the code at least t greater than the designed minimum distance. This motivates to investigate large intervals of gaps at the point P_∞ of the $\mathbb{F}_{q^2}$-maximal curves considered in the present paper. Here we limit ourselves to show a couple of experimental results. We use Janwa's result as stated in [5, Theorem 2] together with [5, Theorem 3] for the zero divisor $B = 0$.

Example 1. Take the curve of equation (1) for $\mathcal{C}$, and let $p = 7, d = 5, h = 2$. Then $d \mid (q+1) = 7^2 + 1$. Form Proposition 19, the non-gaps at P_∞ are $q/p = 7$ and $(q+1)/d = 10$. The gap sequence at P_∞ is $1, 2, 3, 4, 5, 6, 8, 9, 11, 12, 13, 15, 16, 18, 19, 22, 23, 25, 26, 29, 32, 33, 36, 39, 43, 46, 53$. Each of the integers $11 = \gamma - 2 = \gamma - t$, $12 = \gamma - 1$ and $13 = \gamma$ is a gap at P_∞. From [5, Theorem 3], the minimum distance of the code $C_L(\gamma P_\infty, D)$ is at least $d^* = |D| - \gamma + t + 1 = 5037$ whereas the designed minimum distance is $d' = |D| - \gamma = 5034$.

Example 2. Take the curve of equation (1) for $\mathcal{C}$, and let $p = 5, d = 3, h = 3$. Then $d \mid (q+1) = 5^3 + 1$. Form Proposition 19, the non-gaps at P_∞ are $q/p = 25$ and $(q+1)/d = 42$. The gap sequence at P_∞ is $1, \ldots, 24, 26, \ldots, 41, 43, \ldots, 66, 68, \ldots\ldots, 920, 922, \ldots, 962, 964, \ldots, 981, 983$. Each of the integers $1022 = \alpha, \ldots, 1030 = \alpha + 8 = \alpha + t$ and $1072 = \beta, \ldots, 1063 = \beta - t = \beta - (t-1)$ is a gap at P_∞. From [5, Theorem 4], the minimum distance of the differential code $C_\Omega(\gamma P_\infty, D)$ with $\gamma = \alpha + \beta - 1$ is at least $d^* = \alpha+\beta-1-(2\mathfrak{g}-2)+(t+1) = 1120$ whereas the designed minimum distance is $d' = \alpha+\beta-1-(2\mathfrak{g}-2) = 1112$.

It may be noticed that the curve of equation (1) is a C_{ab}-curve with $a = q/p$ and $b = (q+1)/d$. Evaluation codes defined over a C_{ab}-curve have been the subject of recent papers where both encoding and decoding problems are also treated; see [1].

References

BRS 1. P. Beelen, J. Rosenkilde, G. Solomatov, Fast decoding of AG codes, *IEEE Trans. Inform. Theory* **68** (2022), 7215-7232.

s-torres2000 2. A. Cossidente, G. Korchmáros and F. Torres, Curves of large genus covered by the Hermitian curve, *Comm. Algebra* **28** (2000), 4707–4728.

GK 3. B. Gatti, G. Korchmáros, Galois subcovers of the Hermitian curve in characteristic p with respect to subgroups of order p^2, arXiv:2307.15192, (2023).

GL 4. A. Garcia, R.F. Lax, Goppa codes and Weierstrass gaps, *Coding Theory and Algebraic Geometry, Proceedings, Luminy.* (1991), *Lecture Notes in Mathematics*, **1518** (Springer, Berlin, (1992), 33-42.

GKL 5. A. Garcia, S. J. Kim, R. F. Lax, Consecutive Weierstrass gaps and minimum distance of Goppa codes, *J. Pure Appl. Algebra* **84** (1993), 199-207.

GSX 6. A. García, H. Stichtenoth, and C.P. Xing, On Subfields of the Hermitian Function Field, *Comp. Math* **120** (2000), 137-170.

huppertI1967 7. B. Huppert, *Endliche Gruppen. I*, Grundlehren der Mathematischen Wissenschaften **134**, Springer, Berlin, 1967, xii+793 pp.

HKT 8. J.W.P. Hirschfeld, G. Korchmáros, and F. Torres, *Algebraic Curves over a Finite Field*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, 2008. xx+696 pp.

hoffer1972 9. A.R. Hoffer, On unitary collineation groups, *J. Algebra* **22** (1972), 211–218.

Pellikaan1995 10. C. Kirfel, R. Pellikaan, The minimum distance of codes in an array coming from telescope semigroups, *IEEE Trans. Information Theeory*, **41** (1995), 1720-1732.

jan 11. H. Janwa, On the parameters of algebraic geometric codes, in *Applied algebra, algebraic algorithms and error-correcting codes* (New Orleans, LA, 1991), 19–28, Lecture Notes in Comput. Sci., **539**, Springer, Berlin, 1991.

RS 12. H.G. Rück, H. Stichtenoth, A characterization of Hermitian function fields over finite fields, *J. Reine Angew. Math.* **457**, (1994), 185–188.

chtenoth1973I 13. H. Stichtenoth, Über die Automorphismengruppe eines algebraischen Funktionenkörpers von Primzahlcharakteristik. I. Eine Abschätzung der Ordnung der Automorphismengruppe, *Arch. Math.* **24** (1973), 527–544.

stich1993 14. H. Stichtenoth, *Algebraic Function Fields and Codes*, Springer, Berlin, 1993, x+260 pp.

A geometric construction of a class of non-linear MRD codes

Nicola Durante, Giovanni Giuseppe Grimaldi, and Giovanni Longobardi

Università degli Studi di Napoli Federico II
Dipartimento di Matematica e Applicazioni "Renato Caccioppoli"
{ndurante,giovannigiuseppe.grimaldi,giovanni.longobardi}@unina.it

Abstract. In the finite projective space $\mathrm{PG}(n-1,q^n)$, let $\mathcal{X}$ be a C_F^σ-set of an $(n-k+1)$-dimensional subspace Λ with vertices A and B and $\Lambda^\star$ be a $(k-3)$-dimensional subspace skew with Λ. In [8], it is shown that $\mathcal{C}$ is a union of $\{A,B\}$ and $q-1$ pairwise disjoint scattered $\mathbb{F}_q$-linear sets of rank n, say $\mathcal{X}_a$ for any $a\in\mathbb{F}_q^*$. Moreover, the line AB can be partitioned in $\{A,B\}$ and $q-1$ scattered $\mathbb{F}_q$-linear sets J_a of rank n, for any $a\in\mathbb{F}_q^*$. Denote by $\mathcal{K}(\Lambda^\star,\mathcal{E})$ the cone with vertex $\Lambda^\star$ and base the set

$$\mathcal{E}=\left(\mathcal{X}\setminus\bigcup_{a\in T}\mathcal{X}_a\right)\cup\bigcup_{a\in T}J_a,$$

with $1\in T\subset\mathbb{F}_q^*$. Then $\mathcal{K}(\Lambda^\star,\mathcal{E})$ gives rise to a new family of non-linear $(n,n,q;d)$-MRD codes for any $n\geq 3$, $2\leq d\leq n-1$ and $d=n-k+1$. By choosing the parameters or by puncturing appropriately a code in this class, the codes constructed in [3, 9] and in [8] are re-obtained. Finally, an element in this family, if not equivalent to a generalized Gabidulin, is not equivalent to any non-linear MRD codes constructed by Otal and Özbudak in [21].

Keywords: Rank distance code · Linearized polynomial · Linear set · Finite field

1 Rank distance codes and σ-linearized polynomials

Let $\mathbb{F}_q^{m\times n}$ be the set of $m\times n$ matrices with entries over the finite field $\mathbb{F}_q$ of q elements, q a prime power. This set can be equipped with *rank distance* defined as

$$d(A,B)=\mathrm{rk}(A-B)$$

with $A,B\in\mathbb{F}_q^{m\times n}$. A subset of $\mathbb{F}_q^{m\times n}$, including at least two elements, is called a *rank distance code*. The *minimum distance* $d(\mathcal{C})$ of a code $\mathcal{C}$ is defined as

$$d(\mathcal{C})=\min_{\substack{A,B\in\mathcal{C}\\ A\neq B}}d(A,B).$$

If $d:=d(\mathcal{C})$, we will say that $\mathcal{C}\subseteq\mathbb{F}_q^{m\times n}$ is an $(m,n,q;d)$-*rank distance code*. A rank distance code is called *additive* if it is an additive subgroup of $\mathbb{F}_q^{m\times n}$, it is

called $\mathbb{F}_q$-*linear* if it is a subspace of $\mathbb{F}_q^{m\times n}$ seen as a vector space over $\mathbb{F}_q$. A code that is not an $\mathbb{F}_q$-subspace is called a *non-linear* code.

The size of an $(m,n,q;d)$-rank distance code $\mathcal{C}$ satisfies the *Singleton-like bound*, precisely:

$$|\mathcal{C}| \leq q^{\max\{m,n\}(\min\{m,n\}-d+1)},$$

see e.g. [5, Theorem 5.4]. When this bound is achieved, $\mathcal{C}$ is called an $(m,n,q;d)$-*maximum rank distance* code, or shortly $(m,n,q;d)$-*MRD* code. The *adjoint code* of $\mathcal{C}$ is defined as

$$\mathcal{C}^t = \{X^t : X \in \mathcal{C}\}$$

where the superscript t stands for the matrix transposition. Two rank distance codes $\mathcal{C},\mathcal{C}' \subseteq \mathbb{F}_q^{m\times n}$, $m,n \geq 2$, are called *equivalent* if there exist $P \in \mathrm{GL}(m,q)$, $Q \in \mathrm{GL}(n,q)$, $R \in \mathbb{F}_q^{m\times n}$ and a field automorphism $\rho \in \mathrm{Aut}(\mathbb{F}_q)$ such that

$$\mathcal{C}' = P\mathcal{C}^\rho Q + R = \{PX^\rho Q + R : X \in \mathcal{C}\}.$$

When $m = n$, in addition to being equivalent, two codes are said *adjointly equivalent* if $\mathcal{C}'$ and $\mathcal{C}^t$ are equivalent. If both $\mathcal{C}$ and $\mathcal{C}'$ are additive, then one may assume that R is the zero matrix.
The code $\mathcal{C}^{[u]} \subseteq \mathbb{F}_q^{(m-u)\times n}$ obtained from $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ by deleting the last u rows, $1 \leq u \leq m-1$, is called a *punctured code* of $\mathcal{C}$. In [1, Corollary 7.3], it is showed that if $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$, $m \leq n$, is an MRD code then $\mathcal{C}^{[u]}$ is MRD as well.
Although the rank distance codes are subsets of matrices, they can be represented in the setting of *σ-linearized polynomials*.
From now on, suppose $m = n$ and let $\sigma : x \in \mathbb{F}_{q^n} \longrightarrow x^{q^s} \in \mathbb{F}_{q^n}$ be a field automorphism of $\mathbb{F}_{q^n}$ with $\gcd(s,n) = 1$. A *σ-linearized polynomial with coefficients over* $\mathbb{F}_{q^n}$ is a polynomial with the shape $\alpha = \sum_{i=0}^{\ell} \alpha_i X^{\sigma^i}$, $\ell \in \mathbb{N}$. If $\alpha_\ell \neq 0$, the integer ℓ is called the *σ-degree* of α. The set

$$\tilde{\mathcal{L}}_{n,q,\sigma}[X] = \left\{ \sum_{i=0}^{n-1} \alpha_i X^{\sigma^i} : \alpha_i \in \mathbb{F}_{q^n} \right\} \tag{1}$$

of σ-linearized polynomials with σ-degree at most $n-1$, endowed with the usual sum, the scalar multiplication by an element of $\mathbb{F}_q$ and the map composition modulo $X^{q^{sn}} - X$ is an algebra. This isomorphic to $\mathbb{E} = \mathrm{End}(\mathbb{F}_{q^n}, \mathbb{F}_q)$, the algebra of the endomophisms of $\mathbb{F}_{q^n}$ seen as vector space over $\mathbb{F}_q$. Indeed, every $\alpha = \sum_{i=0}^{n-1} \alpha_i X^{\sigma^i} \in \tilde{\mathcal{L}}_{n,q,\sigma}[X]$ corresponds naturally to the endomorphism $\alpha(x) : x \in \mathbb{F}_{q^n} \longrightarrow \sum_{i=0}^{n-1} \alpha_i x^{\sigma^i} \in \mathbb{F}_{q^n}$ and vice versa, every endormophism can be represented uniquely via a σ-linearized polynomial in $\tilde{\mathcal{L}}_{n,q,\sigma}[X]$, see [15, Chapter 3].
Therefore, for any σ-linearized polynomial $\alpha = \sum_{i=0}^{n-1} \alpha_i X^{\sigma^i} \in \tilde{\mathcal{L}}_{n,q,\sigma}[X]$, the *rank* of α is the integer $\mathrm{rk}\alpha := \dim_{\mathbb{F}_q} \mathrm{im}\,\alpha(x)$. Also, α has rank r if and only if

the σ-*Dickson matrix*

$$D_{\sigma,\alpha} = \begin{pmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_{n-1} \\ \alpha_{n-1}^{\sigma} & \alpha_0^{\sigma} & \dots & \alpha_{n-2}^{\sigma} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{\sigma^{n-1}} & \alpha_2^{\sigma^{n-1}} & \dots & \alpha_0^{\sigma^{n-1}} \end{pmatrix}$$

has rank r, see for more details [15].
Then any rank distance code $\mathcal{C}$, $|\mathcal{C}| \geq 2$, can be seen as a suitable subset of $\tilde{\mathcal{L}}_{n,q,\sigma}[X]$ and the definitions of transpose matrix, adjoint code and equivalence between codes can be reformulated in this setting. Indeed, let $\alpha = \sum_{i=0}^{n-1} \alpha_i X^{\sigma^i} \in \tilde{\mathcal{L}}_{n,q,\sigma}[X]$, then the *adjoint polynomial* of α is defined as

$$\hat{\alpha} = \sum_{i=0}^{n-1} \alpha_i^{\sigma^{n-i}} X^{\sigma^{n-i}}$$

and if $\mathcal{C} \subseteq \tilde{\mathcal{L}}_{n,q,\sigma}[X]$ is the rank distance code, $\mathcal{C}^t = \{\hat{\alpha} : \alpha \in \mathcal{C}\}$. Moreover, two rank codes $\mathcal{C}$ and $\mathcal{C}'$ are equivalent or adjointly equivalent if and only if there exist (f, ρ, g, h) such that $f, g, h \in \tilde{\mathcal{L}}_{n,q,\sigma}[X]$, with $f(x)$ and $g(x)$ permutation maps, and $\rho \in \mathrm{Aut}(\mathbb{F}_q)$ such that

$$\mathcal{C}' = \{f \circ \alpha^{\rho} \circ g + h \colon \alpha \in \mathcal{C}\} \quad \text{or} \quad \mathcal{C}' = \{f \circ \alpha^{\rho} \circ g + h \colon \alpha \in \mathcal{C}^t\},$$

respectively, where the automophism ρ acts only over the coefficients of a polynomial α in $\mathcal{C}$ or $\mathcal{C}^t$, respectively.
The first class of linear MRD codes has been discovered independently by Delsarte [5] and Gabidulin [10]. These codes are known in literature as *Delsarte-Gabidulin codes*. Later in [11], Gabidulin and Kshevetskiy provided a generalization of them, called *generalized Gabidulin codes*. Since these codes will be used in the last section, we recall their shape: let $1 \leq k \leq n$ be an integer, a generalized Gabidulin code is equivalent to the set of σ-linearized polynomials

$$\mathcal{G}_{k,\sigma} = \left\{ \sum_{i=0}^{k-1} \alpha_i X^{\sigma^i} : \alpha_0, \dots, \alpha_{k-1} \in \mathbb{F}_{q^n} \right\}$$

and it is an $(n, n, q; d)$-MRD code where $d = n - k + 1$. The code $\mathcal{G}_{k,\sigma}$ is a k-dimensional $\mathbb{F}_{q^n}$-linear subspace of $\tilde{\mathcal{L}}_{n,q,\sigma}[X]$.

In [24], Sheekey exhibited a wider class of linear MRD codes, called *twisted Gabidulin codes* and later generalized in [20] by Lunardon, Trombetti and Zhou. In [23, 25], further generalizations of generalized twisted Gabidulin codes are obtained. Finally, a family of linear maximum rank distance codes in $\tilde{\mathcal{L}}_{n,q,\sigma}[X]$, $n = 2t$, $2 \leq d \leq n$ and q odd is discovered by Trombetti and Zhou in 2019 and described in [26].
Examples of maximal codes with parameters $(n, n, q; n)$, both linear and non-linear exist and are known as *spread sets*, see e.g. [6]. In [3], Cossidente *et al.* exhibited a family of non-linear $(3, 3, q; 2)$-MRD codes and, hence, different from

spread sets. Then, Durante and Siciliano generalized these codes to non-linear $(n, n, q; n-1)$-MRD codes for any $n \geq 3$. Both families have been generalized by Donati and Durante in [8] to a family of non-linear $(d+1, n, q; d)$-MRD codes for any $q > 2$, $n \geq 3$ and $2 \leq d \leq n-1$. This will be described in detail in Section 3. Another family of non-additive MRD codes for all n, d has been constructed by Otal and Özbudak in [21]: let I be a subset of $\mathbb{F}_q$, $1 \leq k \leq n-1$ and consider the set of σ-linearized polynomials

$$\mathcal{C}^{(1)}_{n,k,\sigma,I} = \left\{ \sum_{i=0}^{k-1} \alpha_i X^{\sigma^i} : \alpha_0, \ldots, \alpha_{k-1} \in \mathbb{F}_{q^n}, \mathrm{N}_{q^n/q}(\alpha_0) \in I \right\}, \tag{2}$$

$$\mathcal{C}^{(2)}_{n,k,\sigma,I} = \left\{ \sum_{i=1}^{k} \beta_i X^{\sigma^i} : \beta_1, \ldots, \beta_k \in \mathbb{F}_{q^n}, \mathrm{N}_{q^n/q}(\beta_k) \notin (-1)^{n(k+1)} I \right\}.$$

Then $\mathcal{C}_{n,k,\sigma,I} = \mathcal{C}^{(1)}_{n,k,\sigma,I} \cup \mathcal{C}^{(2)}_{n,k,\sigma,I} \subset \tilde{\mathcal{L}}_{n,q,\sigma}$ is an $(n, n, q; n-k+1)$-MRD code. In [21, Corollary 2.1], the authors proved that

1. if $q = 2$ or $I \in \{\emptyset, \{0\}, \mathbb{F}_q^*, \mathbb{F}_q\}$ then $\mathcal{C}_{n,k,\sigma,I}$ is equivalent to a generalized Gabidulin code
2. if $q > 2$ and $I \notin \{\emptyset, \{0\}, \mathbb{F}_q^*, \mathbb{F}_q\}$, then $\mathcal{C}_{n,k,\sigma,I}$ is not an affine code (i.e. not a translated version of an additive code).

In the following sections, we provide a geometric construction of a class of non-linear $(n, n, q; d)$-MRD $\mathcal{C}_{\sigma,T}$, $1 \in T \subseteq \mathbb{F}_q^*$, $2 \leq d \leq n-1$, and puncturing properly a code in this relevant class, one gets a code described in [8, 9]. We shall show that this class is effectively new, i.e. any code, if not equivalent to a Gabidulin code, is not equivalent to a code constructed as in [21].

2 The geometric setting

Let V be a v-dimensional vector space over the field $\mathbb{F}_{q^n}$ and let $\mathrm{PG}(v-1, q^n) = \mathrm{PG}(V, \mathbb{F}_{q^n})$. Let $U \subset V$ be an $\mathbb{F}_q$-linear vector space such that $\dim_{\mathbb{F}_q} U = u$. The set

$$\Theta = \left\{ \langle \mathbf{u} \rangle_{\mathbb{F}_{q^n}} : \mathbf{u} \in U \setminus \{\mathbf{0}\} \right\}$$

is called $\mathbb{F}_q$*-linear set of rank* u. The size of Θ can be at most $\frac{q^u-1}{q-1}$ and if it is attained, Θ is said to be *scattered*. If $u = v$ and $\langle \Theta \rangle = \mathrm{PG}(v-1, q^n)$, then Θ is a *(canonical) subgeometry* of $\mathrm{PG}(v-1, q^n)$. It follows that Θ is a canonical subgeometry if and only if any its frame is also a frame of $\mathrm{PG}(v-1, q^n)$.

Denote by $(X_0, X_1, \ldots, X_{n-1})$ the homogeneous projective coordinates of $\mathrm{PG}(n-1, q^n)$ and let $\hat{\sigma}$ be the collineation of $\mathrm{PG}(n-1, q^n)$ defined by

$$(X_0, X_1, \ldots, X_{n-1})^{\hat{\sigma}} = (X_{n-1}^{\sigma}, X_0^{\sigma}, \ldots, X_{n-2}^{\sigma}).$$

Then, the collineation $\hat{\sigma}$ fixes pointwise the canonical subgeometry

$$\Sigma = \Sigma_{n,n} = \{(x, x^{\sigma}, \ldots, x^{\sigma^{n-1}}) : x \in \mathbb{F}_{q^n}^*\} \cong \mathrm{PG}(n-1, q). \tag{3}$$

Let $\overline{S} = \mathrm{PG}(W, q^n)$ be a subspace of $\mathrm{PG}(n-1, q^n)$, the integer $\dim_{\mathbb{F}_{q^n}} W = w$ will be called *rank* of $\overline{S}$. Then $S = \overline{S} \cap \Sigma$ is a subspace of Σ of rank at most w. We will say that $\overline{S}$ is a *subspace of* Σ if S and $\overline{S}$ have the same rank. In particular, this holds if and only if $\overline{S}$ is fixed by the collineation $\hat{\sigma}$ (see e.g. [16]).

Let $\mathrm{PG}(\mathbb{F}_q^{m\times n}, \mathbb{F}_q) = \mathrm{PG}(mn-1, q)$, $m \leq n$, and let $\mathcal{S}_{m,n}$ be the *Segre variety* of $\mathrm{PG}(\mathbb{F}_q^{m\times n}, \mathbb{F}_q)$, i.e., $\mathcal{S}_{m,n}$ is the set of all points $\langle M\rangle_{\mathbb{F}_q}$ in $\mathrm{PG}(\mathbb{F}_q^{m\times n}, \mathbb{F}_q)$ such that $\mathrm{rk} M = 1$, see [13, Section 4.5]. This can be seen as the $\mathbb{F}_q$-field reduction of the set of points

$$\Sigma_{m,n} = \{(x, x^\sigma, \ldots, x^{\sigma^{m-1}}) \colon x \in \mathbb{F}_{q^n}^*\} \cong \mathrm{PG}(n-1, q)$$

of $\mathrm{PG}(m-1, q^n)$ in $\mathrm{PG}(mn-1, q)$, see [14].

Let $\mathcal{A}$ be a subset of $\mathrm{PG}(n-1, q)$ and denote by $\Omega_h(\mathcal{A})$ the *h-secant variety* of $\mathcal{A}$, i.e. the union of the ℓ-dimensional projective subspaces spanned by points of $\mathcal{A}$ for any $0 \leq \ell \leq h$, [12]. Note that $\Omega_0(\mathcal{A}) = \mathcal{A}$ and for any $1 \leq h \leq n-1$, $\Omega_{h-1}(\mathcal{A}) \subseteq \Omega_h(\mathcal{A})$. Moreover, if $\mathcal{A} \subset \mathrm{PG}(n-1, q)$ such that $\langle \mathcal{A}\rangle = \mathrm{PG}(t-1, q) \subset \mathrm{PG}(n-1, q)$, then $\Omega_h(\mathcal{A}) = \Omega_{t-1}(\mathcal{A})$ for any $t \leq h \leq n-1$. A set of points $\mathcal{E} \subset \mathrm{PG}(n-1, q)$ is called an *exterior set* with respect to $\Omega_h(\mathcal{A})$ if any line joining two points of $\mathcal{E}$ is disjoint from $\Omega_h(\mathcal{A})$. Applying a similar argument as in [4], we get the following results regarding the size of an exterior set.

Theorem 1. *Let $\mathcal{A} \subset \mathrm{PG}(n-1, q)$ such that $\langle \mathcal{A}\rangle = \mathrm{PG}(n-1, q)$. Let $\mathcal{E} \subset \mathrm{PG}(n-1, q)$ be an exterior set with respect to $\Omega_h(\mathcal{A})$, then*

$$|\mathcal{E}| \leq \frac{q^{n-h-1}-1}{q-1}$$

for any $0 \leq h \leq n-1$.

Given M, N two sets of points of $\mathrm{PG}(n-1, q)$, with $M \cap N = \emptyset$, we will denote by $\mathcal{K}(M, N)$ the *cone with vertex M and base N*, i.e. $\mathcal{K}(M, N)$ is the set of all points belonging to a line joining a point of M and a point of N.

Corollary 1. *Let $\mathcal{A} \subset \mathrm{PG}(n-1, q)$ such that $\langle \mathcal{A}\rangle = \mathrm{PG}(t-1, q)$, $1 \leq t < n$, and let $\mathcal{E} \subset \mathrm{PG}(n-1, q)$ be an exterior set with respect to $\Omega_h(\mathcal{A})$. Then $\mathcal{E}$ is contained in a cone $\mathcal{K} = \mathcal{K}(S_{n-t-1}, \bar{\mathcal{E}})$, with base $\bar{\mathcal{E}} = \mathcal{E} \cap \langle \mathcal{A}\rangle$ and vertex an $(n-t-1)$-dimensional subspace S_{n-t-1} complementary with $\langle \mathcal{A}\rangle$. Moreover,*

$$|\mathcal{E}| \leq \begin{cases} \frac{q^{n-h-1}-1}{q-1} & \text{if} \quad 0 \leq h \leq t-1, \\ \frac{q^{n-t}-1}{q-1} & \text{otherwise.} \end{cases} \tag{4}$$

An exterior set $\mathcal{E} \subset \mathrm{PG}(n-1, q)$ with respect to $\Omega_h(\mathcal{A})$, $\langle \mathcal{A}\rangle = \mathrm{PG}(t-1, q)$, $1 \leq h+1 \leq t \leq n$, is called *maximum* if $|\mathcal{E}| = \frac{q^{n-h-1}-1}{q-1}$. Note that the image of $\Omega_h(\Sigma_{m,n}) \subset \mathrm{PG}(m-1, q^n)$ under the $\mathbb{F}_q$-field reduction is the h-secant variety $\Omega_h(\mathcal{S}_{m,n})$ of the points whose the representative matrices in $\mathbb{F}_q^{m\times n}$ have rank at most $h+1$.

The (maximum) exterior sets with respect to $\Omega_h(\Sigma_{m,n})$ are related to (maximum) rank distance codes. More precisely,

Theorem 2. *Let $\mathcal{E}$ be an exterior set with respect to $\Omega_h(\Sigma_{m,n})$ of $\mathrm{PG}(m-1,q^n)$ and denote by $\mathcal{E}'$ the image of $\mathcal{E}$ under the $\mathbb{F}_q$-field reduction. Then, the set*

$$\mathcal{C} = \left\{\rho M : \langle M\rangle_{\mathbb{F}_q} \in \mathcal{E}', \rho \in \mathbb{F}_q\right\} \tag{5}$$

is an $(m,n,q;h+2)$-RD code. In addition, if $\mathcal{E}$ is maximum then $\mathcal{C}$ is an MRD code.

See [2, 3, 8] and the references therein for more details on exterior sets.

3 Non-linear MRD codes arising from C_F^σ-set

In this section, we recall the construction of the family of non-linear MRD codes exhibited by Donati and Durante in [8].
Let A and B be two distinct points of a projective space $\mathrm{PG}(d,q^n)$, $d \geq 2$, and let $\mathcal{S}_A$ and $\mathcal{S}_B$ be the stars of lines (pencils of lines if $d=2$) through A and B, respectively. Thus, let σ be the Frobenius automorphism of $\mathbb{F}_{q^n}$ defined by $\sigma : x \mapsto x^{q^s}$ with $\gcd(n,s)=1$ and consider Φ a σ-collineation between $\mathcal{S}_A$ and $\mathcal{S}_B$ which does not map the line AB into itself and such that the subspace spanned by the lines $\Phi^{-1}(AB), AB, \Phi(AB)$ has dimension $\min\{3,d\}$. The set $\mathcal{X}$ of points of intersection of corresponding lines under the collineation Φ is called *C_F^σ-set* of $\mathrm{PG}(d,q^n)$ and the points A and B are called the *vertices* of $\mathcal{X}$, see [7, 8]. Let $N_a = \{y \in \mathbb{F}_{q^n} : \mathrm{N}_{q^n/q}(y) = a\}$ for any $a \in \mathbb{F}_q^*$. In [7, 8], it is proved that every C_F^σ-set $\mathcal{X}$ of $\mathrm{PG}(d,q^n)$ with vertices A and B is projectively equivalent to the set

$$\{A,B\} \cup \bigcup_{a\in\mathbb{F}_q^*} \mathcal{X}_a,$$

where $A=(0,\ldots,0,1)$, $B=(1,0,\ldots,0)$ and

$$\mathcal{X}_a = \left\{(1,t,t^{\sigma+1},\ldots,t^{\sigma^{d-1}+\ldots+\sigma+1}) : t \in N_a\right\},$$

for any $a \in \mathbb{F}_q^*$. The $q-1$ sets $\mathcal{X}_a$, called the *components* of $\mathcal{X}$, are pairwise disjoint and any of them is a scattered $\mathbb{F}_q$-linear set of rank n, for more details on linear sets see e.g. [22].

Since $|\mathcal{X}_a| = (q^n-1)/(q-1)$, then $|\mathcal{X}| = q^n+1$. In particular, every $\mathcal{X}_a$ is isomorphic to $\mathrm{PG}(n-1,q)$, see [8, Remark 3.5]. Moreover, for any $a \in \mathbb{F}_q^*$, the line AB of $\mathrm{PG}(d,q^n)$ can be partitioned into the set $\{A,B\}$ and the $q-1$ sets

$$J_a = \left\{(1,0,\ldots,0,(-1)^{d+1}t) : t \in N_a\right\},$$

where any J_a is an $\mathbb{F}_q$-scattered linear set (of pseudoregulus type with transversal points A and B, see [7, Remark 2.2] and [18]) and note that $\mathcal{X}_1 = \Sigma_{d+1,n}$. Now, let Π be a subgeometry of $\Sigma_{d+1,n}$ isomorphic to $\mathrm{PG}(d,q)$. In [8], the following have been proved.

Theorem 3. *[8, Theorem 5.1] For any $T \subset \mathbb{F}_q^*$, $1 \in T$, the set*

$$\mathcal{E} = \left(\mathcal{X} \setminus \bigcup_{a \in T} \mathcal{X}_a \right) \cup \bigcup_{a \in T} J_a$$

is a maximum exterior set with respect to $\Omega_{d-2}(\Pi)$.

Corollary 2. *[8, Corollary 5.2] For all $q > 2$, $n \geq 3$, and $2 \leq d \leq n-1$, to the set $\mathcal{E}$ corresponds to a $(d+1, n, q; d)$-MRD code.*

4 The class of non-linear MRD codes $\mathcal{C}_{\sigma,T}$

Let $\Sigma \cong \mathrm{PG}(n-1,q)$ be a canonical subgeometry of $\mathrm{PG}(n-1,q^n)$ and consider a subspace $\Lambda^\star$ of rank k disjoint from Σ and Λ a subspace of $\mathrm{PG}(n-1,q^n)$ of rank $n-k$ disjoint from $\Lambda^\star$. Let Γ be the projection of Σ from $\Lambda^\star$ to Λ, i.e.,

$$\Gamma = \mathrm{p}_{\Lambda^\star,\Lambda}(\Sigma) = \{\langle \Lambda^\star, P\rangle \cap \Lambda : P \in \Sigma\}\,.$$

By [19, Theorem 1], this is an $\mathbb{F}_q$-linear set of rank n. We recall the following definition given in [17].

Definition 1. Let $\Gamma = \mathrm{p}_{\Lambda^\star,\Lambda}(\Sigma)$ be the projection of a canonical subgeometry $\Sigma \cong \mathrm{PG}(n-1,q)$ from the subspace $\Lambda^\star$ of rank k to the subspace Λ of rank $(n-k)$ in $\mathrm{PG}(n-1,q^n)$. The set Γ is called $(n-k-1)$-*embedding of* Σ if any subspace of Σ of rank $n-k$ is disjoint from $\Lambda^\star$.

Note that $\Gamma = \mathrm{p}_{\Lambda^\star,\Lambda}(\Sigma)$ is an $(n-k-1)$-embedding if and only if for any choice of $n-k$ independent points $R_1, R_2, \ldots, R_{n-k}$ of Σ,

$$\Lambda = \langle R_1', R_2', \ldots, R_{n-k}'\rangle$$

where $R_i' = \mathrm{p}_{\Lambda^\star,\Lambda}(R_i)$, $i = 1, 2, \ldots, n-k$, also this is equivalent to say that

$$\mathrm{p}_{\Lambda^\star,\Lambda} : P \in \Sigma \longrightarrow \langle P, \Lambda^\star\rangle \cap \Lambda \in \Lambda$$

induces an injective map from the set of all subspaces of Σ of rank $\ell \leq n-k-1$ to the set of all subspaces of Λ of the same rank ℓ and, hence, if $R \in \Omega_{n-k-1}(\Sigma)$ then $\langle R, \Lambda^\star\rangle \cap \Lambda \in \Omega_{n-k-1}(\Gamma)$.

Theorem 4. *Let $\Sigma \cong \mathrm{PG}(n-1,q)$ be a canonical subgeometry of $\mathrm{PG}(n-1,q^n)$ and let $\Lambda^\star$ and Λ be subspaces of $\mathrm{PG}(n-1,q^n)$ of rank $k-2$ and $n-k+2$, respectively, such that $\Lambda^\star \cap \Sigma = \emptyset = \Lambda^\star \cap \Lambda$. Let $\Gamma = \mathrm{p}_{\Lambda^\star,\Lambda}(\Sigma)$ be an $(n-k+1)$-embedding of Σ and let $\mathcal{E}$ be a (maximum) exterior set with respect to $\Omega_{n-k-1}(\Gamma)$. Then, $\mathcal{K} = \mathcal{K}(\Lambda^\star, \mathcal{E})$ is a (maximum) exterior set with respect to $\Omega_{n-k-1}(\Sigma)$.*

Proof. Let P, Q be two points in $\mathcal{K} = \mathcal{K}(\Lambda^\star, \mathcal{E})$. We shall distinguish some cases:

- $P, Q \in \Lambda^\star$. Since Γ is an $(n-k+1)$-embedding the line PQ is disjoint from the subspaces of rank $(n-k)$ of Σ and so from $\Omega_{n-k-1}(\Sigma)$.

- $P, Q \in \mathcal{E}$. Suppose that the line PQ meets a subspace S of Σ of rank $(n-k)$ and consider

$$\emptyset \neq \langle \Lambda^\star, S \cap PQ \rangle \cap \Lambda \subseteq \langle \Lambda^\star, S \rangle \cap \Lambda \cap PQ. \tag{6}$$

  This is a contradiction by the hypothesis done over $\mathcal{E}$ and since $\langle \Lambda^\star, S \rangle \cap \Lambda$ is a subspace of rank $n-k$.
- the line $PQ \subseteq \mathcal{K} \setminus (\Lambda^\star \cup \mathcal{E})$ and joins a point of $\Lambda^\star$ and a point of $\mathcal{E}$. Then, without loss of generality, we may suppose that $P \in \Lambda^\star$ and $Q \in \mathcal{E}$. If the line PQ meets a subspace S of Σ of rank $(n-k)$, then

$$Q \in \langle \Lambda^\star, S \cap PQ \rangle \cap \Lambda \subseteq \langle \Lambda^\star, S \rangle \cap \Lambda. \tag{7}$$

  Then Q belongs to a space spanned by points of Γ with rank $(n-k)$, a contradiction.
- the line $PQ \subseteq \mathcal{K} \setminus (\Lambda^\star \cup \mathcal{E})$ and does not joint a point of $\Lambda^\star$ and a point of $\mathcal{E}$ or $PQ \not\subseteq \mathcal{K}$. If PQ meets a subspace S of Σ of rank $(n-k)$, then

$$\emptyset \neq \langle \Lambda^\star, S \cap PQ \rangle \cap \Lambda \subseteq \langle \Lambda^\star, S \rangle \cap \langle \Lambda^\star, PQ \rangle \cap \Lambda. \tag{8}$$

  Since the projection from $\Lambda^\star$ to Λ of the line PQ is a line through two points P', Q' of $\mathcal{E}$, we get that this line meets the space $\langle \Lambda^\star, S \rangle \cap \Lambda$ spanned by $(n-k)$ points of Γ, a contradiction.

Then we have showed that any line through two points of $\mathcal{K}$ is disjoint from $\Omega_{n-k-1}(\Sigma)$. Now, since $\langle \Gamma \rangle = \Lambda$, if $\mathcal{E}$ is a maximum exterior set with respect to $\Omega_{n-k-1}(\Gamma)$, we get

$$|\mathcal{K}(\Lambda^\star, \mathcal{E})| = |\Lambda^\star| + |\mathcal{E}| + |\Lambda^\star||\mathcal{E}|(q^n - 1) = \frac{q^{nk}-1}{q^n-1}.$$

As application of Theorem 4, by an appropriate choice of $\Lambda^\star$, Λ, an $(n-k+1)$-embedding Γ of $\Sigma = \Sigma_{n,n}$, $2 \leq k \leq n-1$, and a maximum exterior set with respect to $\Omega_{n-k-1}(\Gamma)$ in $\mathrm{PG}(n-1, q^n)$, one will get a class of non-linear MRD codes in $\tilde{\mathcal{L}}_{n,q,\sigma}[X]$. Precisely, let

$$\Lambda^\star : X_0 = X_1 = \ldots = X_{n-k+1} = 0$$

and

$$\Lambda : X_{n-k+2} = X_{n-k+3} = \ldots = X_{n-1} = 0$$

be disjoint subspaces of rank $(k-2)$ and $(n-k+2)$ in $\mathrm{PG}(n-1, q^n)$, respectively. Consider the linear set of rank n

$$\Gamma = \mathrm{p}_{\Lambda^\star, \Lambda}(\Sigma) = \{(\alpha, \alpha^\sigma, \ldots, \alpha^{\sigma^{n-k+1}}, 0, \ldots, 0) : \alpha \in \mathbb{F}_{q^n}^*\}, \tag{9}$$

this set is an $(n-k+1)$-embedding of Σ. Finally, let

$$A = (\underbrace{0, \ldots, 0}_{n-k+1}, 1, 0, \ldots, 0) \text{ and } B = (1, 0, \ldots, 0)$$

be points in Λ and consider

$$\mathcal{X} = \bigcup_{a\in\mathbb{F}_q^*} \mathcal{X}_a \cup \{A, B\},$$

where

$$\mathcal{X}_a = \left\{(1, t, t^{\sigma+1}, \ldots, t^{\sigma^{n-k}+\ldots+\sigma+1}, 0, \ldots, 0) : t \in N_a\right\}.$$

The set $\mathcal{X}$ is the C_F^σ-set with vertices A and B generated by a σ-collineation Φ between the star of lines through A and B contained in Λ. Let $T \subseteq \mathbb{F}_q^*$, $1 \in T$, then the set

$$\mathcal{E} = \left(\mathcal{X} \setminus \bigcup_{a\in T} \mathcal{X}_a\right) \cup \bigcup_{a\in T} J_a \tag{10}$$

with

$$J_a = \left\{ (1, \underbrace{0, \ldots, 0}_{n-k}, (-1)^{n-k}t, 0, \ldots, 0) : t \in N_a \right\}, \quad a \in \mathbb{F}_q^*.$$

is a maximum exterior set with respect to $\Omega_{n-k-1}(\Gamma)$ as proved in [8, Theorem 5.1] of size $q^n + 1$. Therefore, the hypothesis of Theorem 4 are satisfied and the cone $\mathcal{K}(\Lambda^\star, \mathcal{E})$ is a maximum exterior set with respect to $\Omega_{n-k-1}(\Sigma)$ and by Theorem 2, the set $\mathcal{C}_{\sigma,T} \subset \mathbb{F}_q^{n\times n}$, as in (5), is a non linear-MRD with minimum distance $d = n - k + 1$ which cannot be a translated version of an additive code. Note that the punctured code $\mathcal{C}_{\sigma,T}^{[k-2]} \subset \mathbb{F}_q^{(n-k+2)\times n}$ obtained by $\mathcal{C}_{\sigma,T}$ deleting the last $(k-2)$ rows is exactly the code constructed in [8, Theorem 5.1] or for $k = 2$ the code appeared in [9] and for $n = 3, k = 2$ is code in [3].

By Theorem 2 and Theorem 4, the $\mathcal{C}_{\sigma,T}$ is a subset of $n \times n$ matrices and so, it can be seen as a subset of $\tilde{\mathcal{L}}_{n,q,\sigma}[X]$. Indeed, fixing $T \subseteq \mathbb{F}_q^*$, $1 \in T$, the homogeneous coordinates of the points belonging to $\mathcal{E}$ have one of the following shape

$$\begin{aligned}
&(\alpha, \alpha^\sigma\xi, \alpha^{\sigma^2}\xi^{\sigma+1}, \ldots, \alpha^{\sigma^{n-k+1}}\xi^{\sigma^{n-k}+\ldots+1}, 0, \ldots, 0)\\
&(\alpha, \underbrace{0, \ldots, 0}_{n-k}, (-1)^{n-k}\alpha^\sigma\eta, 0, \ldots, 0)\\
&A = (\underbrace{0, \ldots, 0}_{n-k+1}, \alpha, 0, \ldots, 0) \text{ and } B = (\alpha, 0, \ldots, 0)
\end{aligned} \tag{11}$$

with $\mathrm{N}_{q^n/q}(\xi) \in \mathbb{F}_q^* \setminus T$ and $\mathrm{N}_{q^n/q}(\eta) \in T$. So the non-linear $(n, n, q; d)$, $d = n - k + 1$, MRD-code $\mathcal{C}_{\sigma,T}$ is the set of σ-linearized polynomials with σ-degree at most $n - 1$ whose coefficients are the homogeneous coordinates of a point in

$\mathcal{K}(\Lambda^\star, \mathcal{E})$ with the zero map, i.e.

$$\mathcal{C}_{\sigma,T} = \left\{ \sum_{i=0}^{d} \lambda\alpha^{\sigma^i} \xi^{\frac{\sigma^i-1}{\sigma-1}} X^{\sigma^i} + \sum_{i=d+1}^{n-1} \beta_i X^{\sigma^i} : \lambda, \alpha, \beta_i \in \mathbb{F}_{q^n}, \mathrm{N}_{q^n/q}(\xi) \in \mathbb{F}_q^* \setminus T \right\}$$
$$\cup \left\{ \lambda\alpha X + (-1)^{d+1}\lambda\alpha^{\sigma}\eta X^{\sigma^d} + \sum_{i=d+1}^{n-1} \beta_i X^{\sigma^i} : \lambda, \alpha, \beta_i \in \mathbb{F}_{q^n}, \mathrm{N}_{q^n/q}(\eta) \in T \right\}$$
$$\cup \left\{ \alpha X^{\sigma^d} + \sum_{i=d+1}^{n-1} \beta_i X^{\sigma^i} : \alpha, \beta_i \in \mathbb{F}_{q^n} \right\} \cup \left\{ \alpha X + \sum_{i=d+1}^{n-1} \beta_i X^{\sigma^i} : \alpha, \beta_i \in \mathbb{F}_{q^n} \right\}.$$

5 The equivalence issue for $\mathcal{C}_{\sigma,T}$

Let σ and τ be generators of the group $\mathrm{Gal}(\mathbb{F}_{q^n}|\mathbb{F}_q)$. In order to state the novelty of the class of non-linear codes obtained, we have to compare a code of type $\mathcal{C}_{\sigma,T}$, $1 \in T \subseteq \mathbb{F}_q^*$, with the code of type $\mathcal{C}_{n,k,\tau,I}$ with $I \subseteq \mathbb{F}_q$, cf. (2).
Let $d = n - k + 1$, note that the sets

$$\mathcal{U} = \left\{ \alpha X^{\sigma^d} + \sum_{i=d+1}^{n-1} \beta_i X^{\sigma^i} : \alpha, \beta_i \in \mathbb{F}_{q^n} \right\} = \left\{ f \circ X^{\sigma^d} : f \in \mathcal{G}_{k-1,\sigma} \right\} \tag{12}$$

and

$$\mathcal{V} = \left\{ \alpha X + \sum_{i=d+1}^{n-1} \beta_i X^{\sigma^i} : \alpha, \beta_i \in \mathbb{F}_{q^n} \right\} = \left\{ f \circ X^{\sigma^{d+1}} : f \in \mathcal{G}_{k-1,\sigma} \right\} \tag{13}$$

contained in $\mathcal{C}_{\sigma,T}$ are equivalent to a generalized Gabidulin code $\mathcal{G}_{k-1,\sigma}$ and their intersection

$$\left\{ \sum_{i=d+1}^{n-1} \beta_i X^{\sigma^i} : \beta_i \in \mathbb{F}_{q^n} \right\} = \left\{ f \circ X^{\sigma^{d+1}} : f \in \mathcal{G}_{k-2,\sigma} \right\}$$

is equivalent to a generalized Gabidulin code $\mathcal{G}_{k-2,\sigma}$.
Clearly, if $q = 2$ or $T = \mathbb{F}_q^*$, $\mathcal{C}_{\sigma,T}$ is equivalent to the generalized Gabidulin code $\mathcal{G}_{k,\sigma}$. While, for the non-linear $(n, n, q; d)$-MRD code $\mathcal{C}_{n,k,\tau,I}$, the following holds.

Lemma 1. *Let $I \notin \{\emptyset, \{0\}, \mathbb{F}_q^*, \mathbb{F}_q\}$, the non-linear $(n, n, q; d)$-MRD code $\mathcal{C}_{n,k,\tau,I}$ contains a unique subspace equivalent to $\mathcal{G}_{k-1,\tau}$ given by*

$$\left\{ \sum_{i=1}^{k-1} \gamma_i X^{\tau^i} : \gamma_i \in \mathbb{F}_{q^n} \right\}.$$

This and [20, Theorem 4.4] allow us to state the following.

Theorem 5. *If $q = 2$ or $T = \mathbb{F}_q^*$ and $I \in \{\emptyset, \{0\}, \mathbb{F}_q^*, \mathbb{F}_q\}$, then the codes $\mathcal{C}_{\sigma,T}$ and $\mathcal{C}_{n,k,\tau,I}$ are equivalent if and only if $\tau \in \{\sigma, \sigma^{-1}\}$. Otherwise, they are neither equivalent nor adjointly equivalent.*

References

1. Byrne, E., Ravagnani, A.: Covering radius of matrix codes endowed with the rank metric. SIAM Journal on Discrete Mathematics **31**(2), 927–944 (2017), ISSN 1095-7146, https://doi.org/10.1137/16M1091769
2. Cooperstein, B.: External flats to varieties in PG($M_{n,n}(GF(q))$). Linear Algebra and its Applications **267**, 175–186 (1997), https://doi.org/10.1016/S0024-3795(97)80049-3
3. Cossidente, A., Marino, G., Pavese, F.: Non-linear maximum rank distance codes. Designs, Codes and Cryptography **79**(2), 597–609 (2016), https://doi.org/10.1007/s10623-015-0108-0
4. De Clerck, F., Thas, J.A.: Exterior sets with respect to the hyperbolic quadric in PG($2n-1, q$). In "Finite Geometry", Lectures Notes in Pure and Applied Mathematics **103**, 83–91 (1985)
5. Delsarte, P.: Bilinear forms over a finite field, with applications to coding theory. Journal of Combinatorial Theory, Series A **25**(3), 226–241 (1978), https://doi.org/10.1016/0097-3165(78)90015-8
6. Dembowski, P.: Finite Geometries. Springer, New York (1968)
7. Donati, G., Durante, N.: Scattered linear sets generated by collineations between pencils of lines. Journal of Algebraic Combinatorics **40**, 1121–1134 (2014), https://doi.org/10.1007/s10801-014-0521-x
8. Donati, G., Durante, N.: A generalization of the normal rational curve in PG(d, q^n) and its associated non-linear MRD codes. Designs, Codes and Cryptography **86**, 1175–1184 (2018), https://doi.org/10.1007/s10623-017-0388-7
9. Durante, N., Siciliano, A.: Non-linear maximum rank distance codes in the cyclic model for the field reduction of finite geometries. The Electronic Journal of Combinatorics **24**(2) (2017), https://doi.org/10.37236/6106
10. Gabidulin, E.: The theory with maximal rank metric distance. Problems if Information Transmission **21**(1), 1–12 (1985)
11. Gabidulin, E., Kshevetskiy, A.: The new construction of rank codes. In: Proceedings. International Symposium on Information Theory, 2005. ISIT 2005, pp. 2105–2108 (2005), https://doi.org/10.1109/ISIT.2005.1523717
12. Harris, J.: Algebraic Geometry, A First Course. Springer-Verlag, New York (1992)
13. Hirschfeld, J., Thas, J.A.: General Galois Geometries, 2nd edn. Springer, New York (1991)
14. Lavrauw, M., Van De Voorde, G.: Field reduction and linear sets in finite geometry. In: Contemporary Mathematics, vol. 632, pp. 271–293 (2015), https://doi.org/10.1090/conm/632/12633
15. Lidl, R., Niederreiter, H.: Finite Fields, 2nd edition. Cambridge University Press, Cambridge (1997)
16. Lunardon, G.: Normal spread. Geometria Dedicata **75**, 245–261 (May 1999), https://doi.org/10.1023/A:1005052007006
17. Lunardon, G.: MRD-codes and linear sets. Journal of Combinatorial Theory, Series A **149**, 1–20 (2017), https://doi.org/0.1016/j.jcta.2017.01.002
18. Lunardon, G., Marino, G., Polverino, O., Trombetti, R.: Maximum scattered linear sets of pseudoregulus type and the Segre Variety $\mathcal{S}_{n,n}$. Journal of Algebraic Combinatorics **39**, 807–831 (2014), https://doi.org/10.1007/s10801-013-0468-3
19. Lunardon, G., Polverino, O.: Translation ovoids of orthogonal polar spaces. Forum Mathematicum **16**(5), 663–669 (2004), https://doi.org/doi:10.1515/form.2004.029

20. Lunardon, G., Trombetti, R., Zhou, Y.: Generalized twisted Gabidulin codes. Journal of Combinatorial Theory, Series A **159**, 79–106 (2018), https://doi.org/10.1016/j.jcta.2018.05.004
21. Otal, K., Özbudak, F.: Some new non-additive maximum rank distance codes. Finite Fields and Their Applications **50**, 293–303 (2018), https://doi.org/10.1016/j.ffa.2017.12.003
22. Polverino, O.: Linear sets in finite projective spaces. Discrete Mathematics **310**(22), 3096–3107 (2010), https://doi.org/10.1016/j.disc.2009.04.007
23. Puchinger, S., Rosenkilde né Nielsen, J., Sheekey, J.: Further Generalisations of Twisted Gabidulin Codes. In: In Proceedings of International Workshop on Coding and Cryptography (2017)
24. Sheekey, J.: A new family of linear maximum rank distance codes. Advances in Mathematics of Communications **10**(3), 475–488 (2016), https://doi.org/10.3934/amc.2016019
25. Sheekey, J.: New semifields and new mrd codes from skew polynomial rings. Journal of the London Mathematical Society **101**(1), 432–456 (2020), https://doi.org/https://doi.org/10.1112/jlms.12281
26. Trombetti, R., Zhou, Y.: A New Family of MRD Codes in $\mathbb{F}_q^{2n\times 2n}$. IEEE Transactions on Information Theory **65**(2), 1054–1062 (2019), https://doi.org/10.1109/TIT.2018.2853184

On the non-existence of 3-dimensional MRD codes of type $\langle X^{q^t}, X + \delta X^{q^{2t}}, G(X)\rangle$

Francesco Ghiandoni

Department of Mathematics and Informatics, University of Florence, Florence, Italy
francesco.ghiandoni@unifi.it

Abstract. In this work we present recent results on the classification of $\mathbb{F}_{q^n}$-linear MRD codes of dimension three. In particular, we provide non-existence results for MRD codes $\mathcal{C} = \langle X^{q^t}, F(X), G(X)\rangle \subseteq \mathcal{L}_{n,q}$ of exceptional type, i.e. such that $\mathcal{C}$ is MRD over infinite many extensions of the field $\mathbb{F}_{q^n}$. These results partially address a conjecture in [6].

Keywords: Scattered polynomials · MRD codes · algebraic curves · finite fields

1 Introduction

Let q be a prime power, n be a positive integer, and denote by $\mathbb{F}_{q^n}$ the finite field with q^n elements and by $\mathbb{P}^N(\mathbb{K})$ (resp. $\mathbb{A}^N(\mathbb{K})$) the N-dimensional projective (resp. affine) space over the field $\mathbb{K}$.
Let $\mathcal{L}_{n,q} = \{\sum_{i=0}^{n-1} a_i X^{q^i} : a_i \in \mathbb{F}_{q^n}\}$ denote the $\mathbb{F}_q$-algebra of the $\mathbb{F}_q$-linearized polynomials (or q-polynomials) of q-degree smaller than n. For any $f(X) = \sum_{i=0}^{n-1} a_i X^{q^i}$, we define $\deg_q(f(X)) = \max\{i : a_i \neq 0\}$ and $\min\deg_q(f(X)) = \min\{i : a_i \neq 0\}$. We identify a polynomial $g(X) \in \mathcal{L}_{n,q}$ with the $\mathbb{F}_q$-linear map $x \mapsto g(x)$ over $\mathbb{F}_{q^n}$; in this way, $\mathbb{F}_q$-linearized polynomials over $\mathbb{F}_{q^n}$ are in one-to-one correspondence with $\mathbb{F}_q$-linear maps over $\mathbb{F}_{q^n}$.
The *rank metric* on the $\mathbb{F}_q$-vector space $\mathbb{F}_q^{m\times n}$ is defined by

$$d(A, B) := \operatorname{rank}(A - B) \quad \text{for } A, B \in \mathbb{F}_q^{m\times n}.$$

We call a subset of $\mathbb{F}_q^{m\times n}$ equipped with the rank metric a *rank-metric code*. For a rank-metric code $\mathcal{C}$ containing at last two elements, its *minimum distance* is given by

$$d(\mathcal{C}) := \min_{A,B\in\mathcal{C}, A\neq B} d(A, B).$$

When $\mathcal{C}$ is an $\mathbb{F}_q$-subspace of $\mathbb{F}_q^{m\times n}$, we say that $\mathcal{C}$ is an $\mathbb{F}_q$-*linear* code of dimension $\dim_{\mathbb{F}_q}(\mathcal{C})$. Under the assumption that $m \leq n$, it is well known (and easily verified) that every rank-metric code $\mathcal{C}$ in $\mathbb{F}_q^{m\times n}$ with minimum distance d satisfies the Singleton-like bound

$$|\mathcal{C}| \leq q^{n(m-d+1)}.$$

In case of equality, $\mathcal{C}$ is called a *maximum rank-metric* code, or MRD code for short. MRD codes have been studied since the 1970s by Delsarte [13] and Gabidulin [14] and have seen much interest in recent years due to an important application in network coding and cryptography [22].
From a different perspective, rank-metric codes can also be seen as sets of (restrictions of) $\mathbb{F}_q$-linear homomorphisms from $(\mathbb{F}_{q^n})^m$ to $\mathbb{F}_{q^n}$ equipped with the rank metric; see [2, Sections 2.2 and

2.3]. With this second point of view, it is evident that multivariate linearized polynomials can be seen as the natural algebraic counterpart of rank-metric codes. In particular, when $m = n$, a rank metric code $\mathcal{C}$ can be seen as set of $\mathbb{F}_q$-linear endomorphisms of $\mathbb{F}_{q^n}$, i.e. $\mathcal{C} \subseteq \mathcal{L}_{n,q}$. From now on we will consider $m = n$ and $d < n$. In the case of univariate linearized polynomials, Sheekey pointed out in [27] the following connection between $\mathbb{F}_{q^n}$-linear MRD codes (i.e MRD codes with left idealizer containing a subring isomorphic to $\mathbb{F}_{q^n}$, see [24]) and the so called *scattered* polynomials: $\mathcal{C}_{f,t} = \langle X^{q^t}, f(X)\rangle_{\mathbb{F}_{q^n}}$ is an MRD code with $\dim_{\mathbb{F}_{q^n}}(\mathcal{C}) = 2$ if and only if

$$\dim_{\mathbb{F}_q} \ker(f(X) - mX) \leq 1$$

for every $m \in \mathbb{F}_{q^n}$. The concept of scattered polynomial introduced in [27] has been slightly generalized in [4].

Definition 1. *[4] An $\mathbb{F}_q$-linearized polynomial $f(X) \in \mathbb{F}_{q^n}[X]$ is called a **scattered** polynomial of index $t \in \{0, \dots, n-1\}$ if*

$$\dim_{\mathbb{F}_q} \ker(f(X) - mX^{q^t}) \leq 1,$$

*for every $m \in \mathbb{F}_{q^n}$. Also, a scattered polynomial of index t is **exceptional** if it is scattered of index t over infinitely many extensions $\mathbb{F}_{q^{nm}}$ of $\mathbb{F}_{q^n}$.*

While several families of scattered polynomials have been constructed in recent years [3, 9, 12, 16, 17, 25, 27, 28], only two families of exceptional ones are known:

(Ps) $f(X) = X^{q^t}$ of index 0, with $\gcd(t, n) = 1$ (polynomials of so-called pseudoregulus type);
(LP) $f(X) = X + \delta X^{q^{2t}}$ of index t, with $\gcd(t, n) = 1$ and $N_{q^n/q}(\delta) = \delta^{(q^n-1)/(q-1)} \neq 1$ (so called LP polynomials).

From a coding theory point of view, if f is exceptional scattered of index t, the corresponding rank distance code $\mathcal{C}^m_{f,t} = \langle X^{q^t}, f(X)\rangle_{\mathbb{F}_{q^{mn}}} \subseteq \mathcal{L}_{nm,q}$ turns out to be an MRD code for infinitely many m; codes of this kind are called exceptional $\mathbb{F}_{q^n}$-linear MRD codes (see [6]). Moreover, in [2] the authors introduce the notions of h-scattered sequences and exceptional h-scattered sequences which constitute the right environment for exceptional MRD codes. Only two families of exceptional $\mathbb{F}_{q^n}$-linear MRD codes are known so far:

(G) $\mathcal{G}_{r,s} = \langle X, X^{q^s}, \dots, X^{q^{s(r-1)}}\rangle_{\mathbb{F}_{q^n}}$, with $\gcd(s, n) = 1$, see [13, 14];
(T) $\mathcal{H}_{r,s}(\delta) = \langle X^{q^s}, \dots, X^{q^{s(r-1)}}, X + \delta X^{q^{sr}}\rangle_{\mathbb{F}_{q^n}}$, with $\gcd(s, n) = 1$ and $N_{q^n/q}(\delta) \neq (-1)^{nr}$ see [17, 27].

The first family is known as generalized Gabidulin codes and the second one as generalized twisted Gabidulin codes.
In [5] it has been shown that the only exceptional $\mathbb{F}_{q^n}$-linear MRD codes spanned by monomials are the codes (G), in connection with so-called Moore exponent sets, while in [6] the authors investigated exceptional $\mathbb{F}_{q^n}$-linear MRD codes not generated by monomials and proved that an exceptional r-dimensional $\mathbb{F}_{q^n}$-linear MRD code contains an exceptional scattered polynomial (see Theorem 4).
Our contribution Motivated by this last necessary condition on MRD codes of exceptional type, we considered codes of type $\mathcal{C} = \langle X^{q^t}, F(X), G(X)\rangle_{\mathbb{F}_{q^n}}$ and we proved a conjecture in [6] for $r = 3$ and F a LP polynomial. Our main result can be summarized as follows (see Theorem 6).

Theorem 1. *If $(t, q) \notin \{(1,3); (1,4); (1,5); (2,3); (2,4); (2,5), (4,3)\}$, then there are no exceptional 3-dimensional $\mathbb{F}_{q^n}$-linear MRD codes of type $\mathcal{C} = \langle X^{q^t}, X + \delta X^{q^{2t}}, G(X)\rangle \subseteq \mathcal{L}_{n,q}$, with $\deg_q(G(X)) > 2t$.*

2 Preliminaries on algebraic curves and varieties

Let $F(X,Y) \in \mathbb{K}[X,Y]$, $\mathbb{K}$ a field, be a polynomial defining an affine plane curve $\mathcal{C} : F(X,Y) = 0$. A plane curve is absolutely irreducible if there are no non-trivial factorizations of its defining polynomial $F(X,Y)$ in $\overline{\mathbb{K}}[X,Y]$, where $\overline{\mathbb{K}}$ is the algebraic closure of $\mathbb{K}$. If $F(X,Y) = \prod_i F^{(i)}(X,Y)$, with $F^{(i)}(X,Y) \in \overline{\mathbb{K}}[X,Y]$ of positive degree, then $\mathcal{C}_i : F^{(i)}(X,Y) = 0$ are called components of $\mathcal{C}$. A component is $\mathbb{F}_q$-rational if it is fixed by the Frobenius morphism φ or equivalently $\lambda F^{(i)}(X,Y) \in \mathbb{K}[X,Y]$ for some $\lambda \in \overline{\mathbb{K}}$.

Let $P = (u,v) \in \mathbb{A}^2(\mathbb{K})$ be a point in the plane, and write

$$F(X+u, Y+v) = F_0(X,Y) + F_1(X,Y) + F_2(X,Y) + \cdots,$$

where F_i is either zero or homogeneous of degree i. The *multiplicity* of $P \in \mathcal{C}$, written as $m_P(\mathcal{C})$ or $m_P(F)$, is the smallest integer m such that $F_m \neq 0$ and $F_i = 0$ for $i < m$; $F_m = 0$ is the *tangent cone* of $\mathcal{C}$ at P. A linear component of the tangent cone is called a *tangent* of $\mathcal{C}$ at P. The point P is on the curve $\mathcal{C}$ if and only if $m_P(\mathcal{C}) \geq 1$. If P is on $\mathcal{C}$, then P is a *simple* point of $\mathcal{C}$ if $m_P(\mathcal{C}) = 1$, otherwise P is a *singular* point of $\mathcal{C}$. It is possible to define in a similar way the multiplicity of an ideal point of $\mathcal{C}$, that is a point of the curve lying on the line at infinity. We denote by $Sing(\mathcal{C})$ the set of singular points of the curve $\mathcal{C}$.

Given two plane curves $\mathcal{A}$ and $\mathcal{B}$ and a point P on the plane, the *intersection number* (or *intersection multiplicity*) $I(P, \mathcal{A} \cap \mathcal{B})$ of $\mathcal{A}$ and $\mathcal{B}$ at the point P can be defined by seven axioms. We do not include its precise and long definition here. For more details, we refer to [18] and [20] where the intersection number is defined equivalently in terms of local rings and in terms of resultants, respectively.

For a given plane curve $\mathcal{C}$ and a point $P \in \mathcal{C}$, we denote by $I_{P,max}(\mathcal{C})$ the maximum possible intersection multiplicity of two components of $\mathcal{C}$ at $P \in Sing(\mathcal{C})$. We list here two useful results in this direction.

Lemma 1. *[18, Section 3.3] [26, Lemma 4.3] [4, Lemma 2.5] Let q be a prime power and $F(X,Y) \in \mathbb{F}_q[X,Y]$. Let $P = (\alpha, \beta) \in \mathbb{F}_q^2$ and write*

$$F(X+\alpha, Y+\beta) = F_m(X,Y) + F_{m+1}(X,Y) + \ldots,$$

where $F_i \in \mathbb{F}_q[X,Y]$ is zero or homogeneous of degree i and $F_m \neq 0$. The following properties hold.

(i) If $F_m(X,Y)$ is separable, then $I_{P,max}(\mathcal{C}) \leq \lfloor m^2/2 \rfloor$.
(ii) Suppose that $F_m = L^m$ with L a linear form;
- *if $L \nmid F_{m+1}$ then $I_{P,max}(\mathcal{C}) = 0$,*
- *if $L^2 \nmid F_{m+1}$ then $I_{P,max}(\mathcal{C}) \leq m$.*

Lemma 2. *[21, Lemma 11] Let $\mathcal{C} : h(X,Y) = 0$ be a curve of degree n defined over $\mathbb{F}_q$.*
If $\sum_{P \in Sing(h)} I_{P,max}(\mathcal{C}) < \frac{2}{9}\deg^2(h)$ then $\mathcal{C}$ possesses at least one absolutely irreducible component defined over $\mathbb{F}_q$.

Consider the set $\overline{\mathbb{F}_q}[[t]]$ of the formal power series on t. Let $(x_0, y_0) \in \overline{\mathbb{F}_q}^2$ be an affine point of $\mathcal{C} : F(X,Y) = 0$. A *branch* of center (x_0, y_0) of $\mathcal{C}$ is a point $(x(t), y(t)) \in (\overline{\mathbb{F}_q}[[t]])^2$ such that

$F(x(t), y(t)) = 0$, where

$$x(t) = x_0 + u_1 t + u_2 t^2 + \dots,$$
$$y(t) = y_0 + v_1 t + v_2 t^2 + \dots.$$

See [20, Chapter 4] for more details on branches. There exists a unique branch centered at a simple point of $\mathcal{C}$. If there exists only a branch centered in a point $P \in \mathcal{C}$ then $I_{P,max}(\mathcal{C}) = 0$.
An algebraic hypersurface is an algebraic variety that may be defined by a single polynomial equation. An algebraic hypersurface defined over a field $\mathbb{K}$ is *absolutely irreducible* if the associated polynomial is irreducible over every algebraic extension of $\mathbb{K}$. An absolutely irreducible $\mathbb{K}$-rational component of a hypersurface $\mathcal{V}$, defined by the polynomial F, is simply an absolutely irreducible hypersurface which is associated to a non-costant factor of F defined over $\mathbb{K}$.

Lemma 3. *[1, Lemma 2.1] Let $\mathcal{X} \subseteq \mathbb{A}^N(\mathbb{F}_q)$ be an affine hypersurface and let $H \subseteq \mathbb{P}^N(\mathbb{F}_q)$ be a projective hypersurface. If its projective closure $\overline{\mathcal{X}} \cap H$ has a non-repeated absolutely irreducible component defined over $\mathbb{F}_q$, then $\overline{\mathcal{X}}$ has an absolutely irreducible component defined over $\mathbb{F}_q$.*

In our investigation we will need bounds on the number of $\mathbb{F}_q$-rational points of algebraic varieties and we will make use of the following result a number of times.

Theorem 2. *[8, Theorem 7.1] Let $\mathcal{W}$ be an absolutely irreducible variety defined over $\mathbb{F}_q$ of dimension n and degree d. If $q > 2(n+1)d^2$, then*

$$\#(\mathcal{W} \cap \mathbb{A}^N(\mathbb{F}_q)) \geq q^n - (d-1)(d-2)q^{n-1/2} + 5d^{13/3}q^{n-1}.$$

3 Scattered sequences of order 1 and MRD codes

In this section we recall the notion of scatteredness of subspaces and sequences in $\mathbb{F}_{q^n}^r$, and how they are related to rank-metric codes.

Definition 2. *[11] Let h, r, n be positive integers, such that $h < r$. An $\mathbb{F}_q$-subspace $U \subseteq \mathbb{F}_{q^n}^r$ is said to be h-**scattered** if for every h-dimensional $\mathbb{F}_{q^n}$-subspace $H \subseteq \mathbb{F}_{q^n}^r$, it holds $\dim_{\mathbb{F}_q}(U \cap H) \leq h$. When $h = 1$, a 1-scattered subspace is simply called **scattered**.*

For what concerns h-scattered subspaces, there is a well-known bound on their $\mathbb{F}_q$-dimension. Namely, an h-scattered subspace $U \subseteq \mathbb{F}_{q^n}^r$ which does not define a subgeometry satisfies

$$\dim_{\mathbb{F}_q}(U) \leq \frac{rn}{h+1}; \tag{1}$$

see [7]. An h-scattered subspace meeting (1) with equality is called a **maximum h-scattered subspace**.

Let $\mathcal{G} = \{g_1, \dots, g_k\} \subseteq \mathcal{L}_{n,q}[X_1, \dots, X_m]$, where $\mathcal{L}_{n,q}[X_1, \dots, X_m] := \left\{ \sum_{i=1}^{m} \sum_{j=0}^{n-1} \gamma_{i,j} X^{q^j} : \gamma_{i,j} \in \mathbb{F}_{q^n} \right\}$.
Let us consider the $\mathbb{F}_q$-space

$$U_{\mathcal{G}} := \{(g_1(x_1, \dots, x_m), \dots, g_r(x_1, \dots, x_m)) : x_1, \dots, x_m \in \mathbb{F}_{q^n}\} \subseteq \mathbb{F}_{q^n}^r. \tag{2}$$

Definition 3. *[2] Let $\mathcal{I} := (i_1, i_2, \dots, i_m) \in (\mathbb{Z}/n\mathbb{Z})^m$ and consider $f_1, \dots, f_r \in \mathcal{L}_{n,q}[X_1, \dots, X_m]$. Let $U_{\mathcal{I},\mathcal{F}} := U_{\mathcal{F}'}$, where $\mathcal{F}' = (X_1^{q^{i_1}}, \dots, X_m^{q^{i_m}}, f_1, \dots, f_s) \subseteq \mathcal{L}_{n,q}[X_1, \dots, X_m]$.*
*The s-tuple $\mathcal{F} := (f_1, \dots, f_s)$ is said to be an $(\mathcal{I}; h)_{q^n}$-**scattered sequence** of order m if $U_{\mathcal{I},\mathcal{F}}$ is maximum h-scattered in $\mathbb{F}_{q^n}^{m+s}$.*
*An $(\mathcal{I}; h)_{q^n}$-scattered sequence $\mathcal{F} := (f_1, \dots, f_s)$ of order m is said to be **exceptional** if it is h-scattered over infinitely many extensions $\mathbb{F}_{q^n}^{\ell}$ of $\mathbb{F}_{q^n}$.*

As the following remark shows, $(\mathcal{I}; h)_{q^n}$-scattered sequences, with $|\mathcal{I}| = 1$, have been considered also in [6], though with slightly different terminology.

Remark 1. It is not difficult to see that, for $m = 1$ and $\mathcal{I} = \{t\}$, an $(r-1)$-tuple $(f_2, \dots, f_r) \subseteq \mathcal{L}_{n,q}$ is a $(\mathcal{I}; r-1)_{q^n}$-scattered (or simply $(t, r-1)_{q^n}$-scattered) sequence of order 1 if and only if, for any $\alpha_1, \dots, \alpha_r \in \mathbb{F}_{q^n}$,

$$\det \begin{pmatrix} \alpha_1^{q^t} & f_2(\alpha_1) & \cdots & f_r(\alpha_1) \\ \alpha_2^{q^t} & f_2(\alpha_2) & \cdots & f_r(\alpha_2) \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_r^{q^t} & f_2(\alpha_r) & \cdots & f_r(\alpha_r) \end{pmatrix} = 0 \quad \Longrightarrow \quad \dim_{\mathbb{F}_q} \langle \alpha_1, \dots, \alpha_r \rangle_{\mathbb{F}_q} < r,$$

see for istance [5, 6, 28] for an explicit link between scattered spaces and Moore matrices.
If the previous property holds, $\underline{f} = (X^{q^t}, f_2, \dots, f_r)$ is said to be a **Moore polynomial set** for q and n (see [6, Definition 9]).

Moore polynomial sets can be characterized in terms of MRD codes as follows.

Theorem 3. *[6] Let k and n be positive integers with $k \le n$, and let $\underline{f} = (X^{q^t}, f_2(X), \dots, f_r(X))$, where $X^{q^t}, f_2(X), \dots, f_r(X) \in \mathcal{L}_{n,q}$ are $\mathbb{F}_{q^n}$-linearly independent. The $\mathbb{F}_{q^n}$-linear rank metric code*

$$\mathcal{C}_{\underline{f}} = \langle X^{q^t}, f_2(X), \dots, f_r(X) \rangle_{\mathbb{F}_{q^n}}$$

is an MRD code if and only if $\underline{f}$ is a Moore polynomial set for q and n.

Now we focus on the exceptionality of $\mathbb{F}_{q^n}$-linear MRD codes $\mathcal{C} \subseteq \mathcal{L}_{n,q}$ of dimension r, or equivalently by Theorem 3 on the exceptionality of scattered sequences of order 1. We can assume infact, without restrictions, the following properties on the polynomials generating a non-degenerate $\mathbb{F}_{q^n}$-linear code $\mathcal{C}$ (for details see [6] and [2, Definition 2.8]). In particular, from [2, Proposition 2.11] we can assume that $\mathcal{C}$ contains a monomial.

Remark 2. [6, Properties 13] Given a non-degenerate $\mathbb{F}_{q^n}$-linear code $\mathcal{C}$ of dimension r, there exist $f_1(X), \dots, f_r(X) \in \mathcal{C}$ such that the following properties hold:

(1) $f_1(X) = X^{q^t}$;
(2) $f_1(X), \dots, f_r(X)$ are $\mathbb{F}_{q^n}$-linearly independent;
(3) $M_1 := \deg_q(f_1(X)), \dots, M_r := \deg_q(f_r(X))$ are all distinct;
(4) $m_1 := \min \deg_q(f_1(X)), \dots, m_r := \min \deg_q(f_r(X))$ are all distinct, and $m_i = 0$ for some i;
(5) $f_1(X), \dots, f_r(X)$ are monic;
(6) for any i, if $f_i(X)$ is a monomial then $m_i = M_i \ge t$.

A Moore polynomial set $\underline{f} = (f_1(X), \dots, f_r(X)) \subseteq \mathcal{L}_{n,q}$ satisfying the previous six properties is said to be a Moore polynomial set for q and n of index t.

4 Scattered sequences and algebraic varieties

In this section, we consider varieties introduced in [6] to traslate the determination of scattered sequences of order 1 into an algebraic geometry problem.

- $\mathcal{U} := \mathcal{U}_{\underline{f}} \subset \mathbb{P}^r(\overline{\mathbb{F}_{q^n}})$, $\quad \mathcal{U} : F_{\underline{f}}(X_1, \dots, X_r) := \det(M_{\underline{f}}(X_1, \dots, X_r)) = 0$, where
$$M_{\underline{f}}(X_1, \dots, X_r) = \begin{pmatrix} f_1(X_1) & f_2(X_1) & \cdots & f_r(X_1) \\ f_1(X_2) & f_2(X_2) & \cdots & f_r(X_2) \\ \vdots & \vdots & \cdots & \vdots \\ f_1(X_r) & f_2(X_r) & \cdots & f_r(X_r) \end{pmatrix};$$
- $\mathcal{V} := \mathcal{U}_{(x, x^q, \dots, x^{q^{r-1}})} \subset \mathbb{P}^r(\overline{\mathbb{F}_{q^n}})$, $\quad \mathcal{V} : F_{(x, x^q, \dots, x^{q^{r-1}})}(X_1, \dots, X_r) = 0$ where
$$F_{(x, x^q, \dots, x^{q^{r-1}})}(X_1, \dots, X_r) = \prod_{(a_1, \dots, a_r) \in \mathbb{P}^{r-1}(\mathbb{F}_q)} (a_1 X_1 + \cdots + a_r X_r); \tag{3}$$
- $\mathcal{W} \subset \mathbb{P}^r(\overline{\mathbb{F}_{q^n}})$, with affine equation
$$\mathcal{W} : \frac{F_{\underline{f}}(X_1, \dots, X_r)}{F_{(x, x^q, \dots, x^{q^{r-1}})}(X_1, \dots, X_r)} = 0. \tag{4}$$

The link between scattered sequence of order 1 and algebraic hypersurfaces is straightforward.

Proposition 1. *[6] The $(r-1)$-tuple $(f_2, \dots, f_r)$ is a $(\{t\}, r-1)_{q^n}$-scattered sequence of order 1 if and only if all the affine $\mathbb{F}_{q^n}$-rational points of $\mathcal{W}$ lie on $\mathcal{V}$.*

Theorem 4. *[6, Main Theorem] Let $\mathcal{C} \subseteq \mathcal{L}_{n,q}$ be an exceptional r-dimensional $\mathbb{F}_{q^n}$-linear MRD code containing at least a separable polynomial $f(x)$ and a monomial. If $r > 3$, assume also that $q > 5$. Let t be the minimum integer such that $X^{q^t} \in \mathcal{C}$. If $t > 0$ and $\mathcal{C} = \langle X^{q^t}, f(X), g_3(X), \dots, g_r(X) \rangle_{\mathbb{F}_{q^n}}$, with $\deg(g_i(X)) > \max\{q^t, \deg(f(X))\}$ for each $i = 3, \dots, r$, then $f(X)$ is exceptional scattered of index t.*

As previously stated in the introduction, until today, the only known non-monomial example of exceptional scattered polynomials, for arbitrary t, is given by the LP polynomials; therefore it is natural to check for exceptional MRD codes where $f(X)$ is of such type. In the following we will focus on RD codes of dimension 3.

5 Moore polynomial sets of type $\underline{f} = (X^{q^t}, X + \delta X^{q^{2t}}, G(X))$

In this section we investigate curves arising from Moore polynomial sets for q and n, of index t, of type $\underline{f} = (X^{q^t}, X + \delta X^{q^{2t}}, G(X))$. Let q be a prime power and t, n integers greater than zero.

Proposition 2. *[6, Proposition 23] If $(X + \delta X^{q^{2t}}, G(X)) \subseteq \mathcal{L}_{n,q}$ is a $(\{t\}, 2)_{q^n}$-scattered sequence of order 1 and $n > 4 \deg_q(G) + 2$, then $\min \deg_q(G) = 2t$ or $\min \deg_q(G) = t/2$.*

Set $m := \min \deg_q(G) = 2t, t/2$, and

$$F := X + \delta X^{q^{2t}}, \qquad N_{q^n|q}(\delta) \neq 1, \tag{5}$$

$$G := X^{q^m} + \cdots + CX^{q^k}, \qquad C \neq 0, \tag{6}$$

where $F, G \in \mathbb{F}_{q^n}[X]$, and $0 < 2t < k < n$.
Fix an element $\lambda \in \mathbb{F}_{q^n} \setminus \mathbb{F}_q$ such that $F(\lambda) \neq 0 \neq G(\lambda)$ and

$$L_\xi(\lambda) := (\xi - \xi^{q^t})^{q^{k+t}}(\xi^{q^{k-t}} - \xi)^{q^t} F(\lambda)^{q^t+1} + \delta(\xi^{q^k} - \xi)^{q^t(q^t+1)} \lambda^{q^t(q^t+1)} \neq 0, \tag{7}$$

for each $\xi \in \mathbb{F}_{q^{k-2t}} \setminus \mathbb{F}_{q^{k-t}}$. Such an element exists in any field $\mathbb{F}_{q^n}$, with $n \geq k+t+1$. Indeed the polynomial L_ξ is not zero and of degree $q^{3t} + q^{2t}$, for each $\xi \in \mathbb{F}_{q^{k-2t}} \setminus \mathbb{F}_{q^{k-t}}$, so

$$\#\left(\cup_{\xi \in \mathbb{F}_{q^{k-2t}}} \{\eta : L_\xi(\eta) = 0\}\right) \leq (q^{3t} + q^{2t}) q^{k-2t} = q^{k+t} + q^k < q^n$$

for $n \geq k+t+1$. This technical assumption on λ is necessary for the proof of Lemma 5.
Consider the curves

$$\mathcal{C} : \begin{vmatrix} X^{q^t} & F(X) & G(X) \\ Y^{q^t} & F(Y) & G(Y) \\ \lambda^{q^t} & F(\lambda) & G(\lambda) \end{vmatrix} = 0, \tag{8}$$

$$\mathcal{A} : \frac{\begin{vmatrix} X^{q^t} & F(X) & G(X) \\ Y^{q^t} & F(Y) & G(Y) \\ \lambda^{q^t} & F(\lambda) & G(\lambda) \end{vmatrix}}{\begin{vmatrix} X & X^q & X^{q^2} \\ Y & Y^q & Y^{q^2} \\ \lambda & \lambda^q & \lambda^{q^2} \end{vmatrix}} = 0. \tag{9}$$

Lemma 4. *[6, Lemma 18] Let* $\mathcal{V} : F_{(x,x^q,x^{q^2})}(X,Y,Z) = 0$ *and* $\mathcal{W} : \dfrac{F_{(x^{q^t}, x+\delta x^{q^{2t}}, G(x))}(X,Y,Z)}{F_{(x,x^q,x^{q^2})}(X,Y,Z)} = 0$ *as in Equations 3,4. If* $\mathcal{A}$ *has a non-repeated* $\mathbb{F}_{q^n}$*-rational absolutely irreducible component not contained in the curve defined by* $F_{(x,x^q,x^{q^2})}(X,Y,\lambda) = 0$*, then* $\mathcal{W}$ *has a non-repeated* $\mathbb{F}_{q^n}$*-rational absolutely irreducible component not contained in* $\mathcal{V}$*.*

In order to apply Lemma 2 on the curve $\mathcal{A}$, we will investigate the singular points of $\mathcal{C}$. In fact, as it can be easily seen, the set of its singular points contains also the singular points of $\mathcal{A}$.

5.1 Singularities of $\mathcal{C}$ at infinity

Consider the line at infinity $\ell_\infty : Z = 0$. A homogeneous equation of $\mathcal{C}$ is given by

$$\begin{vmatrix} X^{q^t} & XZ^{q^{2t}-1} + \delta X^{q^{2t}} & X^{q^m} Z^{q^k - q^m} + \cdots + CX^{q^k} \\ Y^{q^t} & YZ^{q^{2t}-1} + \delta Y^{q^{2t}} & Y^{q^m} Z^{q^k - q^m} + \cdots + CY^{q^k} \\ \lambda^{q^t} & F(\lambda) Z^{q^{2t} - q^t} & G(\lambda) Z^{q^k - q^t} \end{vmatrix} = 0. \tag{10}$$

Thus the curve $\mathcal{C} \cap \ell_\infty$ is defined by

$$\begin{vmatrix} \delta X^{q^{2t}} & CX^{q^k} \\ \delta Y^{q^{2t}} & CY^{q^k} \end{vmatrix} = \delta C(X^{q^{2t}}Y^{q^k} - X^{q^k}Y^{q^{2t}}) = \delta C X^{q^{2t}} Y^{q^{2t}} \prod_{\xi \in \mathbb{F}_{q^{k-2t}} \setminus \{0\}} (Y - \xi X)^{q^{2t}} = 0, \tag{11}$$

so points at infinity of $\mathcal{C}$ are of type $P = (1 : \xi : 0)$, $\xi \in \mathbb{F}_{q^{k-2t}}$, or $P = (0 : 1 : 0)$.
Let consider the projectivity $\Psi : (x : y : z) \mapsto (x : y - \xi x : z)$, that maps $(1 : \xi : 0)$ into $(1 : 0 : 0)$. An affine equation of $\Psi(\mathcal{C})$ (dehomogenizing with respect to X) is given by

$$f(Y, Z) := \begin{vmatrix} 1 & Z^{q^{2t}-1} + \delta & Z^{q^k - q^m} + \cdots + C \\ (Y+\xi)^{q^t} & F^*(Y+\xi, Z) & G^*(Y+\xi, Z) \\ \lambda^{q^t} & F(\lambda) Z^{q^{2t}-q^t} & G(\lambda) Z^{q^k - q^t} \end{vmatrix} = 0, \tag{12}$$

where

$$F^*(Y+\xi, Z) = YZ^{q^{2t}-1} + \delta Y^{q^{2t}} + \xi Z^{q^{2t}-1} + \delta \xi^{q^{2t}}; \tag{13}$$

$$G^*(Y+\xi, Z) = Y^{q^{2t}} Z^{q^k - q^{2t}} + \cdots + CY^{q^k} + \xi^{q^{2t}} Z^{q^k - q^{2t}} + \cdots + C\xi^{q^k}. \tag{14}$$

It is not hard to see that

$$\begin{vmatrix} 1 & Z^{q^{2t}-1} + \delta & Z^{q^k - q^{2t}} + \cdots + C \\ (Y+\xi)^{q^t} & F^*(Y+\xi, Z) & G^*(Y+\xi, Z) \\ \lambda^{q^t} & F(\lambda) Z^{q^{2t}-q^t} & G(\lambda) Z^{q^k - q^t} \end{vmatrix} = B_{q^{2t}-q^t} + B_{q^{2t}-1} + B_{q^{2t}} + L(Y, Z),$$

where

$$B_{q^{2t}-q^t} = -CF(\lambda)(\xi^{q^{k-t}} - \xi)^{q^t} Z^{q^{2t}-q^t}; \tag{15}$$

$$B_{q^{2t}-1} = C\lambda^{q^t}(\xi^{q^k} - \xi) Z^{q^{2t}-1}; \tag{16}$$

$$B_{q^{2t}} = C[F(\lambda) Y^{q^t} Z^{q^{2t}-q^t} - \lambda^{q^t}(YZ^{q^{2t}-1} + \delta Y^{q^{2t}})], \tag{17}$$

and

$$L(Y, Z) = \sum_{i,j} \alpha^{(i,j)} Y^i Z^j, \qquad i + j \geq q^k - q^{k-1} = q^{k-1}(q-1) \geq q^{2t}(q-1). \tag{18}$$

Proposition 3. *Let $P_\xi = (1 : \xi : 0)$, $\xi \in \mathbb{F}_{q^{k-2t}}$. Then $m_{P_\xi}(\mathcal{C}) = q^{2t} - q^t$ or $m_{P_\xi}(\mathcal{C}) = q^{2t}$. Also $m_{P_\xi}(\mathcal{C}) = q^{2t}$ if and only if $\xi \in \mathbb{F}_{q^{\gcd(k,t)}}$.*

Proposition 4. *Let $P = (1 : \xi : 0)$, where $\xi \in \mathbb{F}_{q^{\gcd(k,t)}}$, or $P = (0 : 1 : 0)$. Then $I_{P,max}(\mathcal{C}) \leq \frac{q^{4t}}{4}$.*

Lemma 5. *Let $P_\xi = (1 : \xi : 0)$, with $\xi \notin \mathbb{F}_{q^{\gcd(k,t)}}$. Then there is a unique branch centered at P_ξ. Thus, the multiplicity of two putative components of $\mathcal{C}$ (and therefore $\mathcal{A}$) in P_ξ is 0.*

5.2 Affine singularities of $\mathcal{C}$

Case $\min \deg_q(G(X)) = 2t$ Note that an affine point $P = (\overline{x}, \overline{y}) \in \mathcal{C}$ is singular if and only if

$$\begin{vmatrix} \overline{x}^{q^t} & G(\overline{x}) \\ \lambda^{q^t} & G(\lambda) \end{vmatrix} = \begin{vmatrix} \overline{y}^{q^t} & G(\overline{y}) \\ \lambda^{q^t} & G(\lambda) \end{vmatrix} = 0.$$

One can see immediately that

$$\begin{vmatrix} (X+\overline{x})^{q^t} & F(X+\overline{x}) & G(X+\overline{x}) \\ (Y+\overline{y})^{q^t} & F(Y+\overline{y}) & G(Y+\overline{y}) \\ \lambda^{q^t} & F(\lambda) & G(\lambda) \end{vmatrix} = H_{q^t} + H_{q^t+1} + \ldots,$$

where

$$H_{q^t} = \begin{vmatrix} F(\overline{y}) & G(\overline{y}) \\ F(\lambda) & G(\lambda) \end{vmatrix} X^{q^t} - \begin{vmatrix} F(\overline{x}) & G(\overline{x}) \\ F(\lambda) & G(\lambda) \end{vmatrix} Y^{q^t}; \tag{19}$$

$$H_{q^t+1} = G(\lambda)(X^{q^t}Y - XY^{q^t}). \tag{20}$$

As a direct consequence of Lemma 1 we have the following.

Proposition 5. *Let $\mathcal{C} : F_{\underline{f}}(X,Y,\lambda) = 0$ and $P = (\overline{x},\overline{y}) \in \mathcal{C}$ such that*

$$\begin{vmatrix} \overline{x}^{q^t} & G(\overline{x}) \\ \lambda^{q^t} & G(\lambda) \end{vmatrix} = \begin{vmatrix} \overline{y}^{q^t} & G(\overline{y}) \\ \lambda^{q^t} & G(\lambda) \end{vmatrix} = 0.$$

Then

- $I_{P,max}(\mathcal{C}) \leq \frac{(q^t+1)^2}{4}$ *if* $\begin{vmatrix} F(\overline{y}) & G(\overline{y}) \\ F(\lambda) & G(\lambda) \end{vmatrix} = \begin{vmatrix} F(\overline{x}) & G(\overline{x}) \\ F(\lambda) & G(\lambda) \end{vmatrix} = 0$;
- $I_{P,max}(\mathcal{C}) \leq q^t$ *otherwise.*

Remark 3. Observe that

$$\begin{vmatrix} \overline{x}^{q^t} & G(\overline{x}) \\ \lambda^{q^t} & G(\lambda) \end{vmatrix} = [\overline{x}G(\lambda)^{1/q^t} - \lambda(\overline{x}^{q^t} + \cdots + C^{1/q^t}\overline{x}^{q^{k-t}})]^{q^t} = g(\overline{x})^{q^t}$$

where g is a separable polynomial of degree q^{k-t}. Thus the number of affine singular points of $\mathcal{C}$ is at most $q^{2(k-t)}$. Moreover, if $m_P(\mathcal{C}) = q^t+1$, that is

$$\begin{vmatrix} F(\overline{y}) & G(\overline{y}) \\ F(\lambda) & G(\lambda) \end{vmatrix} = \begin{vmatrix} F(\overline{x}) & G(\overline{x}) \\ F(\lambda) & G(\lambda) \end{vmatrix} = 0, \tag{21}$$

then we obtain

$$\frac{\overline{x}^{q^t}}{\lambda^{q^t}} = \frac{G(\overline{x})}{G(\lambda)} = \frac{F(\overline{x})}{F(\lambda)},$$

i.e.

$$F(\overline{x})\lambda^{q^t} - F(\lambda)\overline{x}^{q^t} = \delta\lambda^{q^t}\overline{x}^{q^{2t}} - (\lambda + \delta\lambda^{q^{2t}})\overline{x}^{q^t} + \lambda^{q^t}\overline{x} = 0. \tag{22}$$

By combining Equation (22) with $g(\overline{x}) = \overline{x}G(\lambda)^{1/q^t} - \lambda(\overline{x}^{q^t} + \cdots + C^{1/q^t}\overline{x}^{q^{k-t}}) = 0$, we get that $\overline{x}$ (or $\overline{y}$ equivalently) must be a root of a polynomial

$$h(X)^{q^t} := [\alpha_0 X + \cdots + \alpha_{k-2t}X^{q^{k-2t}} + \tilde{\alpha}X^{q^t}]^{q^t} \tag{23}$$

for suitable $\alpha_0,\ldots,\alpha_{k-2t},\tilde{\alpha} \in \mathbb{F}_{q^n}$, where $\deg(h) \leq q^{\max\{k-2t,t\}}$. Thus

$$\#\{P \in \mathcal{C} : m_P(\mathcal{C}) = q^t+1\} \leq q^{\min\{\max\{2k-4t,2t\},4t\}}.$$

Case $\min \deg_q(G(X)) = t/2$ The procedure is exactly the same as for $\min \deg_q(G(X)) = 2t$. A point $P = (\overline{x}, \overline{y}) \in \mathcal{C}$ is singular if and only if

$$\begin{vmatrix} \overline{x}^{q^t} & G(\overline{x}) \\ \lambda^{q^t} & G(\lambda) \end{vmatrix} = \begin{vmatrix} \overline{y}^{q^t} & G(\overline{y}) \\ \lambda^{q^t} & G(\lambda) \end{vmatrix} = 0.$$

By direct computations one can observe that

$$\begin{vmatrix} (X+\overline{x})^{q^t} & F(X+\overline{x}) & G(X+\overline{x}) \\ (Y+\overline{y})^{q^t} & F(Y+\overline{y}) & G(Y+\overline{y}) \\ \lambda^{q^t} & F(\lambda) & G(\lambda) \end{vmatrix} = H_{q^{t/2}} + H_{q^{t/2}+1} + \dots,$$

where

$$H_{q^{t/2}} = \begin{vmatrix} \overline{y}^{q^t} & F(\overline{y}) \\ \lambda^{q^t} & F(\lambda) \end{vmatrix} X^{q^{t/2}} - \begin{vmatrix} \overline{x}^{q^t} & F(\overline{x}) \\ \lambda^{q^t} & F(\lambda) \end{vmatrix} Y^{q^{t/2}}; \tag{24}$$

$$H_{q^{t/2}+1} = \lambda^{q^t}(X^{q^{t/2}}Y - XY^{q^{t/2}}). \tag{25}$$

Thus $I_{P,max}(\mathcal{C}) \leq q^{t/2}$ or $I_{P,max}(\mathcal{C}) \leq \frac{(q^{t/2}+1)^2}{4}$ by Lemma 1.

6 Main result

Thanks to the results stated in Subsections 5.1 and 5.2, we have been able to find upper bounds on $\sum_{P \in Sing(\mathcal{A})} I_{P,max}(\mathcal{A})$ and prove the existence of a suitable component of $\mathcal{A}$ via Lemma 2.

Theorem 5. *Fix q, t, k integers, with $q > 2$, $t > 0$ and $k > 2t$. Let $n \geq k + t + 1$ and $F, G, \lambda, \mathcal{C}, \mathcal{A}$ as in Equations 5,6,7,8,9. If $(t, q) \notin \{(1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (4,3)\}$, then $\mathcal{A}$ has an absolutely irreducible component defined over $\mathbb{F}_{q^n}$ and not contained in $F_{(x,x^q,x^{q^2})}(X, Y, \lambda) = 0$.*

Now we are in position to prove our main result.

Theorem 6. *Let $\mathcal{C} = \langle X^{q^t}, X + \delta X^{q^{2t}}, G(X) \rangle \subseteq \mathcal{L}_{n,q}$ be an exceptional 3-dimensional $\mathbb{F}_{q^n}$-linear MRD code, where t is the minimum integer such that $X^{q^t} \in \mathcal{C}$, and suppose that*

$$(t, q) \notin \{(1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (4,3)\}.$$

Then $\deg_q(G(X)) < 2t$.

Proof. By Theorem 3 and Proposition 2 it follows that $\min \deg_q(G(X)) = 2t, t/2$. Suppose by way of contradiction that $\deg_q(G(X)) > 2t$; then $G(X)$ contains at least two terms (see Equations 6). Let consider $\mathcal{W}, \lambda, \mathcal{A}$ as in Equations 4, 7, 9. Theorems 5 ensure the existence of an absolutely irreducible component defined over $\mathbb{F}_{q^n}$ of $\mathcal{A}$ not contained in the curve defined by $F_{(x,x^q,x^{q^2})}(X, Y, \lambda) = 0$. By Lemma 4, $\mathcal{W}$ has an absolutely irreducible $\mathbb{F}_{q^n}$-rational component not contained in $\mathcal{V}$. Therefore Theorem 2 guarantees the existence of $\mathbb{F}_{q^n}$-rational points in $\mathcal{W} \setminus \mathcal{V}$ for n large enough, and a contradiction arises from Theorem 3 and Proposition 1.

References

1. Y. Aubry, G. McGuire and F. Rodier. A few more functions that are not APN infinitely often, finite fields theory and applications. *Contemporary Math.*, 518:23–31, 2010.
2. D. Bartoli, G. Marino, A. Neri, and L. Vicino. Exceptional scattered sequences, 2022.
3. D. Bartoli, C. Zanella and F. Zullo. A new family of maximum scattered linear sets in PG$(1, q^6)$. *Ars Math. Contemporanea*, 19(1):125-145, 2020
4. D. Bartoli and Y. Zhou. Exceptional scattered polynomials. *J. Algebra*, 509:507–534, 2018.
5. D. Bartoli and Y. Zhou Asymptotics of Moore exponent sets. *J. Comb. Theory Ser. A*, 175:105281, 2020.
6. D. Bartoli, G. Zini and F. Zullo. Linear Maximum Rank Distance Codes of Exceptional Type *IEEE Tran. Inf. Theory*, 69(6):3627–3636, 2023.
7. A. Blokhuis and M. Lavrauw. Scattered spaces with respect to a spread in PG(n, q). *Geometriae Dedicata*, 81(1):231–243, 2000.
8. A. Cafure and G. Matera. Improved explicit estimates on the number of solutions of equations over a finite field. *Finite Fields Appl.*, 12(2):155–185, 2006.
9. B. Csajbók, G. Marino, O. Polverino and C. Zanella. A new family of MRD-codes. *Linear Algebra Appl.*, 548:203–220, 2018.
10. B. Csajbók, G. Marino, O. Polverino and F. Zullo. Maximum scattered linear sets and MRD codes. *J. Algebraic Comb.*, 46(3-4):517–531, 2017.
11. B. Csajbók, G. Marino, O. Polverino and F. Zullo. Generalising the scattered property of subspaces. *Combinatorica*, 41(2):237–262, 2021.
12. B. Csajbók, G. Marino and F. Zullo, New maximum scattered linear sets of the projective line. *Finite Fields Their Appl.*, 54:133–150, 2018.
13. P. Delsarte. Bilinear forms over a finite field, with applications to coding theory. *J. Comb. Theory Ser. A*, 25(3):226–241, 1978.
14. E.M. Gabidulin. Theory of codes with maximum rank distance. *Problemy Peredachi Informatsii*, 21(1):3–16, 1985.
15. S. Lang and A. Weil. Number of points of varieties in finite fields. *Amer. J. Math.*, 76:819–827, 1954.
16. G. Longobardi and C. Zanella, Linear sets and MRD-codes arising from a class of scattered linearized polynomials. *J. Algebr. Combinatorics*, 53:639–661, 2021.
17. G. Lunardon, R. Trombetti and Y. Zhou, Generalized twisted Gabidulin codes. *J. Comb. Theory Ser. A*, 159:79–106, 2018.
18. W. Fulton, *Algebraic curves*, Advanced Book Classics, Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA, 1989.
19. F. Hernando and G. McGuire. Proof of a conjecture on the sequence of exceptional numbers classifying cyclic codes and APN functions. *J. Algebra*, 343:78–92, 2011.
20. J. W. P. Hirschfeld, G. Korchmáros and F. Torres. *Algebraic Curves over a Finite Field*. Princeton Series in Applied Mathematics, Princeton. 2008.
21. D. Jedlicka. APN monomials over $GF(2^n)$ for infinitely many n. *Finite Fields Appl.*, 13:1006–1028, 2007.
22. R. Koetter and F. Kschischang. Coding for errors and erasure in random network coding. *IEEE Trans. Inform. Theory*, 54(8):3579-3591, 2008.
23. G.L. Mullen and D. Panario. *Handbook of Finite Fields*. Chapman and Hall/CRC. 2013.
24. O. Polverino, P. Santonastaso, J. Sheekey and F. Zullo. Divisible Linear Rank Metric Codes. in IEEE Trans. Inform. Theory, 69(7):4528-4536, 2023.
25. A. Neri, P. Santonastaso and F. Zullo. Extending two families of maximum rank distance codes *Finite Fields Their Appl.*, 81:102045, 2022.
26. K.-U. Schmidt and Y. Zhou. Planar functions over fields of characteristic two. *J. Algebraic Comb.*, 40:503–526, 2014.

27. J. Sheekey. A new family of linear maximum rank distance codes. *Adv. Math. Commun.*, 10(3):475-488, 2016.
28. C. Zanella, A condition for scattered linearized polynomials involving Dickson matrices. *J. Geometry*, 110(3):1–9, 2019.

New scattered sequences of order $m \geq 3$

Alessandro Giannoni[1] and Giuseppe Marino[2]

[1] Department of Mathematics and applications "Renato Caccioppoli", University Federico II of Napoli, Napoli, Italy
[2] Department of Mathematics and applications "Renato Caccioppoli", University Federico II of Napoli, Napoli, Italy

Keywords: Scattered linear sets, Scattered sequences, Evasive subspaces

1 Introduction

Scattered sequences and exceptional scattered sequences can be seen as the geometrical counterparts of exceptional MRD codes. Rank-distance(RD) codes were introduced already in the late 70's by Delsarte [11] and then rediscovered by Gabidulin a few years later [13]. Due to their applications in network coding [24] and cryptography [14, 16], they attracted lots of attention in the last decade. RD codes are sets of matrices over a finite field $\mathbb{F}_q$ endowed with the so-called rank distance: the distance between two elements is defined as the rank of their difference. Among them, of particular interest is the family of rank-metric codes whose parameters are optimal, i.e. they have the maximum possible cardinality for a given minimum rank. Such codes are called *maximum rank distance (MRD) codes* and constructing new families is an important and active research task. From a different perspective, rank-metric codes can also be seen as sets of (restrictions of) $\mathbb{F}_q$-linear homomorphisms from $(\mathbb{F}_{q^n})^m$ to $\mathbb{F}_{q^n}$ equipped with the rank distance; see [3, Sections 2.2 and 2.3]. In the case of univariate linearized polynomials such a connection was already exploited in [23] by Sheekey, where the notion of scattered polynomials was introduced; see also [6]. Let $\mathcal{L}_{n,q}[X]$ be the set of q-linearized polynomials. For a polynomial $f \in \mathcal{L}_{n,q}[X]$ and a nonnegative integer $t \leq n-1$ we say that f is scattered of index t if for every $x, y \in \mathbb{F}_{q^n}^*$

$$\frac{f(x)}{x^{q^t}} = \frac{f(y)}{y^{q^t}} \iff \frac{y}{x} \in \mathbb{F}_q,$$

or equivalently

$$\dim_{\mathbb{F}_q}(\ker(f(x) - \alpha x^{q^t})) \leq 1, \text{ for every } \alpha \in \mathbb{F}_{q^n}.$$

In a more geometrical setting, a scattered polynomial is connected with a scattered subspace of the projective line; see [8]. From a coding theory point of view, f is scattered of index t if and only if $\mathcal{C}_{f,t} = \langle x^{q^t}, f(x)\rangle_{\mathbb{F}_{q^n}}$ is an MRD code with $\dim_{\mathbb{F}_{q^n}}(\mathcal{C}_{f,t}) = 2$. The polynomial f is said to be *exceptional scattered* of index t if it is scattered of index t as a polynomial in $\mathcal{L}_{\ell n,q}[X]$, for infinitely many ℓ; see [6]. The classification of exceptional scattered polynomials is still not complete, although it gained the attention of several researchers [1, 2, 4, 6, 12].

So far, many families of scattered polynomials have been constructed; see [5, 8–10, 17–23, 25, 26]. Among them, only two families are exceptional:

(Ps) $f(x) = x^{q^s}$ of index 0, with $\gcd(s, n) = 1$ (polynomials of so-called pseudoregulus type);
(LP) $f(x) = x + \delta x^{q^{2s}}$ of index s, with $\gcd(s, n) = 1$ and $\mathrm{N}_{q^n/q}(\delta) \neq 1$ (so-called LP polynomials).

The generalization of the notion of exceptional scattered polynomials – together with their connection with $\mathbb{F}_{q^n}$-linear MRD codes of $\mathbb{F}_{q^n}$-dimension 2 – yielded the introduction of the concept of $\mathbb{F}_{q^n}$-linear MRD codes of *exceptional type*; see [7]. An $\mathbb{F}_{q^n}$-linear MRD code $\mathcal{C} \subseteq \mathcal{L}_{n,q}[X]$ is an *exceptional MRD code* if the rank metric code

$$\mathcal{C}_\ell = \langle \mathcal{C} \rangle_{\mathbb{F}_{q^{\ell n}}} \subseteq \mathcal{L}_{\ell n,q}[X]$$

is an MRD code for infinitely many ℓ.

Only two families of exceptional $\mathbb{F}_{q^n}$-linear MRD codes are known:

(G) $\mathcal{G}_{k,s} = \langle x, x^{q^s}, \ldots, x^{q^{s(k-1)}} \rangle_{\mathbb{F}_{q^n}}$, with $\gcd(s,n) = 1$; see [11, 13, 15];
(T) $\mathcal{H}_{k,s}(\delta) = \langle x^{q^s}, \ldots, x^{q^{s(k-1)}}, x + \delta x^{q^{sk}} \rangle_{\mathbb{F}_{q^n}}$, with $\gcd(s,n) = 1$ and $\mathrm{N}_{q^n/q}(\delta) \neq (-1)^{nk}$; see [20,23].

The first family is known as *generalized Gabidulin codes* and the second one as *generalized twisted Gabidulin codes*, whereas in [6] it has been shown that the only exceptional $\mathbb{F}_{q^n}$-linear MRD codes spanned by monomials are the codes (G), in connection with so-called *Moore exponent sets*. Non-existence results on exceptional MRD codes were provided in [7, Main Theorem].

A generalization of MRD codes of exceptional type is connected with the notions of h-scattered sequences and exceptional h-scattered sequences, introduced in [3] as sequences of multivariate linearized polynomials $f_1, \ldots, f_s \in \mathcal{L}_{n,q}[X_1, \ldots, X_m]$, such that there exists $\mathcal{I} = (i_1, \ldots, i_m) \in \mathbb{N}^m$ so that the space

$$\{(x_1^{q^{i_1}}, \ldots, x_m^{q^{i_m}}, f_1(x_1, \ldots, x_m), \ldots, f_s(x_1, \ldots, x_m)) : x_1, \ldots, x_m \in \mathbb{F}_{q^n}\}$$

is h-scattered.

Using this new terminology, exceptional $\mathbb{F}_{q^n}$-linear MRD codes correspond to exceptional scattered sequences of order 1. In [3] exceptional scattered sequences of order 2 were investigated for the first time. Clearly, when considering sequences of order larger than one, one must check that these examples are really "new", i.e. they cannot be obtained as direct sum of two scattered sequences of smaller order. This led to the notion of indecomposability; see [3].

The aim of this talk is to present the first infinite family of exceptional scattered and indecomposable sequences of any order greater than two. In the last part of the talk it's also considered the equivalence issue it is worth mentioning that our family is quite large since it contains many non-equivalent sequences.

2 Definitions and Main Results

Let $q = p^h$, where p is a prime and $h > 0$ an integer, and denote by $\mathbb{F}_q$ the finite field with q elements.

We start with the definition of scattered sequences.

Definition 1. *[3, Definition 3.1] Consider $\mathcal{F} = (f_1, \ldots, f_s)$, with $f_1, \ldots, f_s \in \mathcal{L}_{n,q}[\underline{X}]$, which is the set of q-linearized polynomials in $X_1, \ldots, X_m$.*

We define

$$U_{\mathcal{F}} := \{(f_1(x_1, \ldots, x_m), \ldots, f_s(x_1, \ldots, x_m)) : x_1, \ldots, x_m \in \mathbb{F}_{q^n}\}.$$

Let $\mathcal{I} := (i_1, i_2, \ldots, i_m) \in (\mathbb{Z}/n\mathbb{Z})^m$, we define the ***$\mathcal{I}$-space*** *$U_{\mathcal{I},\mathcal{F}} := U_{\mathcal{F}'}$, where*

$$\mathcal{F}' = (X_1^{q^{i_1}}, \ldots, X_m^{q^{i_m}}, f_1, \ldots, f_s).$$

The s-tuple $\mathcal{F} := (f_1, \ldots, f_s)$ is said to be an ***$(\mathcal{I}; h)_{q^n}$-scattered sequence of order*** *m if the $\mathcal{I}$-space $U_{\mathcal{I},\mathcal{F}}$ is maximum h-scattered in $V(m+s, q^n)$. An $(\mathcal{I}; h)_{q^n}$-scattered sequence $\mathcal{F} := (f_1, \ldots, f_s)$ of order m is said to be* ***exceptional*** *if it is h-scattered over infinitely many extensions $\mathbb{F}_{q^{n\ell}}$ of $\mathbb{F}_{q^n}$.*

The main issue, when considering scattered sequences of order larger than one is given by its indecomposability.

Definition 2. *An nm-dimensional $\mathbb{F}_q$-subspace $U_{\mathcal{H}}$ of $V(k, q^n)$ is said to be* ***decomposable*** *if it can be written as*

$$U_{\mathcal{H}} = U_{\mathcal{F}} \oplus U_{\mathcal{G}}$$

for some nonempty $\mathcal{F}, \mathcal{G}$. When this happens we say that $\mathcal{F}$ and $\mathcal{G}$ are ***factors*** *of $\mathcal{H}$. Furthermore, U is then said to be* ***indecomposable*** *if it is not decomposable.*

Let $\mathcal{I} := (i_1, \ldots, i_m)$, $\mathcal{J} := (j_1, \ldots, j_{m'})$, let $\mathcal{F} = (f_1, \ldots, f_s)$ and $\mathcal{G} = (g_1, \ldots, g_{s'})$ be $(\mathcal{I}; h)_{q^n}$ and $(\mathcal{J}; h)_{q^n}$-scattered sequences of orders m and m', respectively. The direct sum $\mathcal{H} := \mathcal{F} \oplus \mathcal{G}$ is the $(s+s')$-tuple $(f_1, \ldots, f_s, g_1, \ldots, g_{s'})$. Since

$$U_{\mathcal{I}\oplus\mathcal{J},\mathcal{H}} = U_{\mathcal{I},\mathcal{F}} \oplus U_{\mathcal{J},\mathcal{G}},$$

$\mathcal{H}$ is an $(\mathcal{I} \oplus \mathcal{J}; h)_{q^n}$-scattered sequence of order $m + m'$.

To prove the indecomposability we need the concept of evasive subspace.

Definition 3. *Let h, r, k, n be positive integers, such that $h < k$ and $h \leq r$. An $\mathbb{F}_q$-subspace $U \subseteq V(k, q^n)$ is said to be (h, r)-evasive if for every h-dimensional $\mathbb{F}_{q^n}$-subspace $H \subseteq V(k, q^n)$, it holds $\dim_{\mathbb{F}_q}(U \cap H) \leq r$. When $h = r$, an (h, h)-evasive subspace is called h-scattered. Furthermore, when $h = 1$, a 1-scattered subspace is simply called scattered.*

Definition 4. *Let n, m be positive integers, with $m \geq 3$ and q a prime power. Consider the finite field $\mathbb{F}_{q^n}$. For each choice of $\alpha_1, \ldots, \alpha_m \in \mathbb{F}_{q^n}^*$ and $I, J \in \mathbb{N}$, $I < J < n$ we define the set:*

$$U_{\boldsymbol{A}}^{I,J} := \{(x_1, \ldots, x_m, f_1(\underline{x}), f_2(\underline{x}), \ldots, f_{m-1}(\underline{x}), f_m(\underline{x})) : x_1, \ldots, x_m \in \mathbb{F}_{q^n}\},$$

where $\boldsymbol{A} := (\alpha_1, \ldots, \alpha_m)$, $f_m(\underline{x}) := {x_m}^{q^I} + \alpha_1 {x_1}^{q^J}$ and $f_i(\underline{x}) := {x_i}^{q^I} + \alpha_{i+1} {x_{i+1}}^{q^J}$ with $i = 1, \ldots, m-1$.

From now on, we will denote $J - I$ as K and $(q^{h\ell} - 1)/(q^h - 1)$ as $C_{h,\ell}$.

Theorem 1. *Assume that* $\gcd(I, J) = 1$ *and that*

$$K_{\boldsymbol{A}}^{I,J} := \frac{\alpha_3 \cdot \alpha_4^{q^K} \cdot \alpha_5^{q^{2K}} \ldots \alpha_m^{q^{(m-3)K}} \cdot \alpha_1^{q^{(m-2)K}}}{\alpha_2^{C_{K,m-1}}}$$

is not a $C_{K,m}$-power in $\mathbb{F}_{q^n}$. Then the set $U_{\boldsymbol{A}}^{I,J}$ is scattered.

Lemma 1. *Let $n \in \mathbb{N}$, let $A \in \mathbb{N}$ such that* $\gcd(q, A) = 1$*, then there exist infinitely many $h \in \mathbb{N}$ such that* $\gcd(A, C_{n,h}) = 1$.

Corollary 1. *Assume that* $\gcd(I,J)=1$ *and that* $K_{\boldsymbol{A}}^{I,J}$ *is not an* $C_{K,m}$*-power in* $\mathbb{F}_{q^n}$*. Then the set* $U_{\boldsymbol{A}}^{I,J}$ *is exceptional scattered.*

Proof. From the previous lemma, there exist infinitely many $h\in\mathbb{N}$ such that

$$\gcd(C_{K,m},C_{n,h})=1.$$

Let us consider a fixed h satisfying the above property. By Bézout's identity, there exist integers c_1 and c_2 such that $c_1C_{K,m}+c_2C_{n,h}=1$. Suppose by the way of contradiction that there exists $\xi\in\mathbb{F}_{q^{hn}}\setminus\mathbb{F}_{q^n}$ such that $K_{\mathbf{A}}^{I,J}=\xi^{C_{K,m}}$. So $\xi^{C_{K,m}}\in\mathbb{F}_{q^n}$, and so $1=(\xi^{C_{K,m}})^{q^n-1}=(\xi^{q^n-1})^{C_{K,m}}$.

Raising both sides to the power of c_1, we obtain

$$1=(\xi^{q^n-1})^{c_1C_{K,m}}=(\xi^{q^n-1})^{-c_2C_{n,h}}(\xi^{q^n-1})=\xi^{q^n-1},$$

a contradiction to $\xi\notin\mathbb{F}_{q^n}$.

Therefore, there are infinitely many extensions of $\mathbb{F}_q$ where $K_{\mathbf{A}}^{I,J}$ is not an $C_{K,m}$-power, and by Theorem 1 the claim follows.

To prove the indecomposability we used the following lemma

Lemma 2. *Let* $\mathcal{F}:=(f_1,\dots,f_s)$ *be an exceptional* $(\mathcal{I};h)_{q^n}$*-scattered sequence of order* m*. If* $U_{\mathcal{I},\mathcal{F}}$ *is* $(r,rn/(h+1)-1)$*-evasive for any* $r\in[h+1,\lfloor(m+s)/2\rfloor]$ *with* $(h+1)\mid rn$ *then* $\mathcal{F}$ *is indecomposable.*

So, to satisfy the assumptions of the lemma, we proved the following two theorems

Theorem 2. *If* $n\geq 2(mJ+J+1)$ *then* $U_{\mathbf{A}}^{I,J}$ *is* $(r,\frac{rn}{2}-1)_q$*-evasive for any odd* $r\in[2,\dots,m]$.

Theorem 3. *If* $n\geq 2(mJ+J+1)$ *and* $\frac{\Pi_{\delta+2}}{\Pi_2}$ *is not a* $(q^{mK}-1)$*-power in* $\mathbb{F}_{q^n}$ *for any* $\delta=1,\dots,m-1$*, where* $\Pi_i=\alpha_i^{q^{(m-1)K}}\alpha_{i-1}^{q^{(m-2)K}}\cdots\alpha_{i+2}^{q^K}\alpha_{i+1}$*, where the indices of* α_i *are modulo* m *in the range* $[1,\dots,m]$*, then* $U_{\mathbf{A}}^{I,J}$ *is* $(r,\frac{rn}{2}-1)_q$*-evasive for any even* $r\in[2,\dots,m]$.

So we have

Corollary 2. *If* $n\geq 2(mJ+J+1)$*, and* $\frac{\Pi_{\delta+2}}{\Pi_2}$ *is not a* $(q^{mK}-1)$*-power in* $\mathbb{F}_{q^n}$ *for any* $\delta=1,\dots,m-1$*, then* $U_{\mathbf{A}}^{I,J}$ *is indecomposable.*

Proof. It follows from [3, Lemma 3.4].

Theorem 4. *Assume that* $\gcd(I,J)=1$*,* $K_{\boldsymbol{A}}^{I,J}$ *is not a* $C_{K,m}$*-power and* $\frac{\Pi_{\delta+2}}{\Pi_2}$ *is not a* $(q^{Km}-1)$*-power in* $\mathbb{F}_{q^n}$ *for any* $\delta=1,\dots,m-1$*. Then* $U_{\mathbf{A}}^{I,J}$ *is scattered and indecomposable in infinitely many extensions of* $\mathbb{F}_{q^n}$*.*

Proof. From Proposition (1) we have the existence of a sequence of positive integers $(h_k)_k$ such that $\gcd(q^{Km}-1,C_{n,h_k})=1$. This implies $\gcd(C_{K.m},C_{n,h_k})=1$, so from the calculations on Corollary (1), we have that $U_{\mathbf{A}}^{I,J}$ is scattered in $\mathbb{F}_{q^{nh_k}}$ for every k. With similar calculations, we obtain that $\frac{\Pi_{\delta+2}}{\Pi_2}$ is not a $(q^{Km}-1)$-power in $\mathbb{F}_{q^nh_k}$ for every k and δ.

Moreover, there exists an h_{k_0} such that

$$nh_{k_0}\geq 2(mJ+J+1)$$

and, by Corollary 2, $U_{\mathbf{A}}^{I,J}$ is indecomposable in every extension $\mathbb{F}_{q^{nh_k}}$ with $h_k\geq h_{k_0}$.

A necessary condition to avoid making the hypotheses of Theorem (4) empty is $m|n$. Given that, we have found a lower bound on the number of m-tuples $(\alpha_1, \ldots, \alpha_m)$ that make $U_{\mathbf{A}}^{I,J}$ scattered and indecomposable

$$Q_{n,m}^{I,J} := (q^n-1)^{m-1}\left((q^n-1) - \frac{q^n-1}{\gcd(q^n-1, C_{K,m})} - \sum_{j=1}^{\lceil \frac{m-1}{2} \rceil}(q^n-1)\frac{q^{\gcd(mn',j)}-1}{q^{m\gcd(n',K)}-1}\right),$$

where $n = mn'$.

We have also studied the $\Gamma L_q(2m, q^n)$-equivalence between sets of the type $U_{\mathbf{A}}^{I,J}$, obtaining these two results.

Theorem 5. *Let I, J, I_0, J_0 be nonnegative integers, such that $J + J_0 < n$, $I < J$, and $I_0 < J_0$. The two sets $U_{\mathbf{A}}^{I,J}$ and $U_{\overline{\mathbf{A}}}^{I_0,J_0}$ are not $\Gamma L(2m, q^n)$-equivalent if $(I, J) \neq (I_0, J_0)$.*

Theorem 6. *Let (I, J) be such that $J < n/2$, $\gcd(I, J) = 1$. Two sets $U_{\mathbf{A}}^{I,J}$ and $U_{\overline{\mathbf{A}}}^{I,J}$ are $\Gamma L(2m, q^n)$-equivalent if and only if $\exists\ \sigma \in Aut(\mathbb{F}_{q^n})$ such that one among these m elements is a $q^{mK}-1$ power:*

$$C_1 := \left(\frac{\overline{\alpha_2}}{\alpha_2^\sigma}\right)\left(\frac{\overline{\alpha_3}}{\alpha_3^\sigma}\right)^{q^K}\cdots\left(\frac{\overline{\alpha_m}}{\alpha_m^\sigma}\right)^{q^{(m-2)K}}\left(\frac{\overline{\alpha_1}}{\alpha_1^\sigma}\right)^{q^{(m-1)K}}$$

$$C_\delta := \left(\frac{\overline{\alpha_{\delta+1}}}{\alpha_2^\sigma}\right)\cdots\left(\frac{\overline{\alpha_m}}{\alpha_{m-\delta+1}^\sigma}\right)^{q^{(m-\delta-1)K}}\left(\frac{\overline{\alpha_1}}{\alpha_{m-\delta+2}^\sigma}\right)^{q^{(m-\delta)K}}\cdots\left(\frac{\overline{\alpha_{\delta-1}}}{\alpha_m^\sigma}\right)^{q^{(m-2)K}}\left(\frac{\overline{\alpha_\delta}}{\alpha_1^\sigma}\right)^{q^{(m-1)K}}$$

$$C_m := \left(\frac{\overline{\alpha_1}}{\alpha_2^\sigma}\right)\cdots\left(\frac{\overline{\alpha_{m-1}}}{\alpha_m^\sigma}\right)^{q^{(m-2)K}}\left(\frac{\overline{\alpha_m}}{\alpha_1^\sigma}\right)^{q^{(m-1)K}},$$

with $\delta = 2, \ldots, m-1$.

These two results allowed us to determine a lower bound on the number of $\Gamma L_q(2m, q^n)$-inequivalent scattered sets contained in our family. Fix $(\overline{\alpha_1}, \ldots, \overline{\alpha_m}) \in \mathbb{F}_{q^n}^m$ and consider all the $(\alpha_1, \ldots, \alpha_m) \in \mathbb{F}_{q^n}^m$ such that the corresponding sets are equivalent. For given $\alpha_2, \ldots, \alpha_m$, the function

$$\alpha_1 \longmapsto C_1 = \left(\frac{\overline{\alpha_2}}{\alpha_2^\sigma}\right)\left(\frac{\overline{\alpha_3}}{\alpha_3^\sigma}\right)^{q^K}\cdots\left(\frac{\overline{\alpha_m}}{\alpha_m^\sigma}\right)^{q^{(m-2)K}}\left(\frac{\overline{\alpha_1}}{\alpha_1^\sigma}\right)^{q^{(m-1)K}}$$

is a permutation of $\mathbb{F}_{q^n}$. Since $\alpha_2, \ldots, \alpha_m$ can vary in $(q^n-1)^{m-1}$ ways, C_1 is a $(q^{mk}-1)$-power for $(q^n-1)^m/\gcd(q^{mK}-1, q^n-1)$ m-uples $(\alpha_1, \ldots, \alpha_m)$.

An equivalence with $\sigma \neq id$ corresponds to an equivalence with $(\alpha_1^\sigma, \ldots, \alpha_1^\sigma)$. Via the condition on C_1, there are at most $nh(q^n-1)^m/(\gcd(q^{mK}-1, q^n-1))$ sets $U_{\mathbf{A}}^{I,J}$ equivalent to $U_{\overline{\mathbf{A}}}^{I,J}$.

Arguing analogously for $C_2, \ldots, C_m$, we obtain that

$$\frac{mnh(q^n-1)^m}{\gcd(q^{mK}-1, q^n-1)}$$

is an upper bound for the number of sets $U_{\mathbf{A}}^{I,J}$ equivalent to a fixed $U_{\overline{\mathbf{A}}}^{I,J}$.

Using the lower bound on the number of distinct instances of $(\alpha_1, \ldots, \alpha_m)$ giving rise to a scattered and indecomposable set $U_{\mathbf{A}}^{I,J}$, remarking that $n = mn'$, we can obtain a lower bound on the number of inequivalent scattered sequences

$$(q^n-1)^{m-1}\left((q^n-1) - \frac{q^n-1}{\gcd(q^n-1, C_{K,m})} - \sum_{j=1}^{\lceil \frac{m-1}{2} \rceil} (q^n-1)\frac{q^{\gcd(mn',j)}-1}{q^{m\gcd(n',K)}-1}\right) \cdot \frac{q^{m\gcd(n',K)}-1}{mnh(q^n-1)^m}.$$

References

1. D. Bartoli. Hasse-Weil type theorems and relevant classes of polynomial functions. *London Mathematical Society Lecture Note Series, Proceedings of 28th British Combinatorial Conference, Cambridge University Press*, pages 43–102, 2021.
2. D. Bartoli, M. Giulietti, and G. Zini. Towards the classification of exceptional scattered polynomials. *arXiv preprint arXiv:2206.13795*, 2022.
3. D. Bartoli, G. Marino, A. Neri, and L. Vicino. Exceptional scattered sequences. *arXiv preprint arXiv:2211.11477*, 2022.
4. D. Bartoli and M. Montanucci. On the classification of exceptional scattered polynomials. *J. Combin. Theory Ser. A*, 179:105386, 28, 2021.
5. D. Bartoli, C. Zanella, and F. Zullo. A new family of maximum scattered linear sets in PG$(1, q^6)$. *Ars Math. Contemp.*, 19(1):125–145, 2020.
6. D. Bartoli and Y. Zhou. Exceptional scattered polynomials. *J. Algebra*, 509:507–534, 2018.
7. D. Bartoli, G. Zini, and F. Zullo. Linear maximum rank distance codes of exceptional type. *IEEE Transactions on Information Theory*, pages vol. 69, no. 6, pp. 3627–3636, doi: 10.1109/TIT.2023.3243682, 2023.
8. A. Blokhuis and M. Lavrauw. Scattered spaces with respect to a spread in PG(n, q). *Geom. Dedicata*, 81(1):231–243, 2000.
9. B. Csajbók, G. Marino, O. Polverino, and C. Zanella. A new family of MRD-codes. *Linear Algebra Appl.*, 548:203–220, 2018.
10. B. Csajbók, G. Marino, and F. Zullo. New maximum scattered linear sets of the projective line. *Finite Fields Appl.*, 54:133–150, 2018.
11. P. Delsarte. Bilinear forms over a finite field, with applications to coding theory. *J. Combin. Theory Ser. A*, 25(3):226–241, 1978.
12. A. Ferraguti and G. Micheli. Exceptional scatteredness in prime degree. *J. Algebra*, 565:691–701, 2021.
13. E. M. Gabidulin. Theory of codes with maximum rank distance. *Problemy Peredachi Informatsii*, 21(1):3–16, 1985.
14. E. M. Gabidulin, A. Paramonov, and O. Tretjakov. Ideals over a non-commutative ring and their application in cryptology. In *Workshop on the Theory and Application of of Cryptographic Techniques*, pages 482–489. Springer, 1991.
15. A. Kshevetskiy and E. Gabidulin. The new construction of rank codes. In *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.*, pages 2105–2108. IEEE, 2005.
16. P. Loidreau. A new rank metric codes based encryption scheme. In *Post-quantum cryptography*, volume 10346 of *Lecture Notes in Comput. Sci.*, pages 3–17. Springer, Cham, 2017.
17. G. Longobardi, G. Marino, R. Trombetti, and Y. Zhou. A large family of maximum scattered linear sets of PG$(1, q^n)$ and their associated MRD codes. *Combinatorica*, 43:681–716, 2023.
18. G. Longobardi and C. Zanella. Linear sets and MRD-codes arising from a class of scattered linearized polynomials. *J. Algebraic Combin.*, pages 1–23, 2021.
19. G. Lunardon and O. Polverino. Blocking sets of size $q^t+q^{t-1}+1$. *J. Combin. Theory Ser. A*, 90(1):148–158, 2000.

20. G. Lunardon, R. Trombetti, and Y. Zhou. Generalized twisted Gabidulin codes. *J. Combin. Theory Ser. A*, 159:79–106, 2018.
21. G. Marino, M. Montanucci, and F. Zullo. MRD-codes arising from the trinomial $x^q+x^{q^3}+cx^{q^5} \in \mathbb{F}_{q^6}[x]$. *Linear Algebra Appl.*, 591:99–114, 2020.
22. A. Neri, P. Santonastaso, and F. Zullo. Extending two families of maximum rank distance codes. *Finite Fields Appl.*, 81:102045, 2022.
23. J. Sheekey. A new family of linear maximum rank distance codes. *Adv. Math. Commun.*, 10(3):475, 2016.
24. D. Silva, F. R. Kschischang, and R. Kötter. A rank-metric approach to error control in random network coding. *IEEE Trans. Inform. Theory*, 54(9):3951–3967, 2008.
25. C. Zanella. A condition for scattered linearized polynomials involving Dickson matrices. *J. Geom.*, 110(3):1–9, 2019.
26. C. Zanella and F. Zullo. Vertex properties of maximum scattered linear sets of $\mathrm{PG}(1, q^n)$. *Discrete Math.*, 343(5):111800, 2020.

Exceptional scattered polynomials in odd degree

Massimo Giulietti[1] and Giovanni Zini[2]

[1] University of Perugia massimo.giulietti@unipg.it
[2] University of Modena and Reggio Emilia giovanni.zini@unimore.it

Abstract. Scattered polynomials over finite fields attracted an increasing attention in the last years. One of the reasons is their deep connection with Maximum Rank Distance (MRD) codes. Known classification results for exceptional scattered polynomials, i.e. polynomials which are scattered over infinite field extensions, are limited to the cases where their index ℓ is small, or a prime number larger than the q-degree k of the polynomial, or an integer smaller than the k in the case where k is a prime. In this paper we completely classify exceptional scattered polynomials when the maximum of ℓ and k is odd.

MSC: 11T06.
Keywords: linearized polynomials, scattered polynomials, MRD codes, transitive finite linear groups.

1 Introduction

For q a prime power and n a positive integer, let $\mathbb{F}_{q^n}$ be the finite field with q^n elements. A q-linearized polynomial $f(x) = \sum_{i=0}^{n-1} a_i x^{q^i} \in \mathbb{F}_{q^n}[x]$ is said to be *scattered* of index $\ell \in \{0, \dots, n-1\}$ over $\mathbb{F}_{q^n}$ if, for any $y, z \in \mathbb{F}_{q^n}^*$,

$$\frac{f(y)}{y^{q^\ell}} = \frac{f(z)}{z^{q^\ell}} \Longrightarrow \frac{y}{z} \in \mathbb{F}_q; \tag{1}$$

see [4, 24]. The q-degree of a linearized polynomial $f(x) = \sum_{i=0}^{n-1} a_i x^{q^i}$ is defined as $\max\{i : a_i \neq 0\}$. For a scattered polynomial $f(x)$ of index ℓ and q-degree k we let $d := \max\{k, \ell\}$.

Scattered polynomials $f(x) \in \mathbb{F}_{q^n}[x]$ are connected with *scattered $\mathbb{F}_q$-subspaces* with respect to a Desarguesian spread. Recall that an *n-spread* of the r-dimensional vector space V over $\mathbb{F}_{q^n}$ is a set of n-dimensional $\mathbb{F}_q$-subspaces of V covering V and pairwise intersecting trivially. An $\mathbb{F}_q$-subspace U of V is *scattered* w.r.t. a spread S if U meets every element of S in an $\mathbb{F}_q$-space of dimension at most 1. A Deasarguesian spread of V arises by applying a field reduction to the vectors of V; see [18]. For a scattered polynomials $f(x) \in \mathbb{F}_{q^n}[x]$ of index ℓ, the set

$$U_f = \{(x^{q^\ell}, f(x)) \colon x \in \mathbb{F}_{q^n}\}$$

is a scattered $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^n} \times \mathbb{F}_{q^n}$ w.r.t. a Desarguesian spread; see [4, 24].

Scattered $\mathbb{F}_q$-subspaces have applications in different areas of mathematics, such as translation hyperovals [10], translation caps in affine spaces [2], two-intersection sets [6], blocking sets [1], translation spreads of the Cayley generalized hexagon [21], finite semifields [17], and graph theory [7]. Of particular interest is the connection between scattered subspaces and linear codes, namely MRD codes [23, 26, 24]:indeed, the polynomial $f(x)$ is scattered of index ℓ if and only if the set $\{ax^{q^\ell} + bf(x)\colon a, b \in \mathbb{F}_{q^n}\} \subseteq \mathrm{End}_{\mathbb{F}_q}(\mathbb{F}_{q^n}) \cong \mathbb{F}_q^{n\times n}$ is an $\mathbb{F}_q$-linear MRD code of $n \times n$ matrices, with $\mathbb{F}_q$-dimension $2n$ and minimum rank distance $n-1$. The relevance of MRD codes in communication theory relies on their applications to random linear network coding [25] and cryptography [9].

A scattered polynomial (of a certain index ℓ) over $\mathbb{F}_{q^n}$ is said to be *exceptional* if it is scattered (with respect to the same index ℓ) over infinitely many extensions $\mathbb{F}_{q^{nm}}$ of $\mathbb{F}_{q^n}$.

While several families of scattered polynomials have been constructed in recent years, only two families of exceptional polynomials are known so far and can be described as follows:

- polynomials of so-called pseudoregulus type, $f(x) = x^{q^s}$ of index 0, with $\gcd(s, n) = 1$, see [19];
- polynomials of so-called LP type (named after Lunardon and Polverino who introduced them in [20]), $f(x) = x + \delta x^{q^{2s}}$ of index s, with $\gcd(s, n) = 1$ and $\mathrm{N}_{q^n/q}(\delta) \neq 1$, see [20].

Classification of exceptional scattered polynomials is a natural problem, and it has been achieved only for index 0, 1, and 2 (see [3, 4]), and for prime values of d (see [8]).

In the investigation of exceptional scattered polynomials of index ℓ, we can assume that a q-linearized polynomial $f(x) \in \mathbb{F}_{q^n}[x]$ is ℓ-normalized, in the following sense (see [5, Remark 4.1]):

(i) the q-degree k of $f(x)$ is smaller than n;
(ii) $f(x)$ is monic;
(iii) the coefficient of x^{q^ℓ} in $f(x)$ is zero;
(iv) if $\ell > 0$, then the coefficient of x in $f(x)$ is nonzero, i.e. $f(x)$ is separable.

In this paper we completely classify scattered polynomials with odd $d = \max\{k, \ell\}$. A partial classification is also obtained when d is even. As in [8], our standpoint will be that of the theory of Galois extensions of function fields. Our main result is the following.

Main Theorem. *Let $f(x) \in \mathbb{F}_{q^n}[x]$ be an ℓ-normalized $\mathbb{F}_q$-linearized polynomial of q-degree k, and let $d := \max\{k, \ell\}$. If d is odd and $f(x)$ is exceptional scattered, then $f(x)$ is a monomial of pseudoregulus type.*

2 A result from Hering

In this section and the following one we deal with linear groups over finite fields, because Galois groups of linearized polynomials are linear.

In this section we recall a know result by Hering on linear groups which will be crucial in our argument. Although the proof of such a result is basically contained in [12, 13], we decided to sketch it to provide a precise reference.

Throughout this section, $f(x) \in \mathbb{F}_{q^n}[x]$ is an ℓ-normalized exceptional scattered polynomial of q-degree k, and $d = \max\{k, \ell\}$. Also, $\Gamma\mathrm{L}_q(a, q^b)$ denotes the subgroup of $\Gamma\mathrm{L}(a, q^b)$ defined as $\Gamma\mathrm{L}_q(a, q^b) = \mathrm{GL}(a, q^b) \rtimes \mathrm{Aut}(\mathbb{F}_{q^b} : \mathbb{F}_q)$.

As usual, $\mathrm{Sp}(e, q)$, with e even, denotes the symplectic group, i.e. the subgroup of $\mathrm{GL}(e, q)$ preserving a given non-degenerate alternating bilinear form. In terms of matrices, $\mathrm{Sp}(e, q)$ is made by the matrices $A \in \mathrm{GL}(e, q)$ such that $AHA^\top = H$, where H is a given invertible skew-symmetric matrix. Also, $G_2(q)$ denotes the Chevalley group made by the automorphisms of the split octonion algebra over $\mathbb{F}_q$, and is a subgroup of $\mathrm{Sp}(6, q)$. For $q > 2$, $G_2(q)$ is simple; for $q = 2$, $G_2(2)$ contains the simple group $G_2(2)' \cong \mathrm{PSU}(3, 3)$ with index 2.

Theorem 1. *Let p be a prime, $q = p^a$, $d \in \mathbb{N}$, $V = \mathbb{F}_q^d$, and G be a subgroup of $\mathrm{GL}(d, q)$ acting transitively on $V \setminus \{\mathbf{0}\}$. Then one of the following holds:*

1. $\mathrm{SL}(e, q^{d/e}) \lhd G \leq \Gamma\mathrm{L}_q(e, q^{d/e})$ *for some $e \mid d$;*
2. $\mathrm{Sp}(e, q^{d/e}) \lhd G \leq \Gamma\mathrm{L}_q(e, q^{d/e})$ *for some even $e \mid d$ with $e \geq 4$;*
3. $G_2(2^{d/6})' \lhd G \leq \Gamma\mathrm{L}_q(6, 2^{d/6})$, *where $p = 2$ and $6 \mid d$;*
4. $q^d \in \{5^2, 7^2, 11^2, 23^2, 29^2, 59^2, 2^4, 3^4, 3^6\}$.

Proof. The proof essentially goes back to Hering's papers [12] and [13], which build on his previous work [11]. Up to our knowledge, in the literature Hering's results have been summarized as in Theorem 1 only for $q = p$ a prime (for instance see [15, Theorem 69.7]), but in fact the case q a prime power can be easily dealt with.

To see this, use the same notations of [12, Section 5]. Let L be a subset of $\mathrm{Hom}(V, V)$ maximal with respect to the following conditions: L is normalized by G, L contains the identity, and L is a field with respect to the addition and multiplication in $\mathrm{Hom}(V, V)$. Then V is an L-vector space with scalar multiplication $\alpha v := \alpha(v)$ for any $\alpha \in L$ and $v \in V$. By [12, Lemma 5.2], up to excluding a specific case (namely $n = 2$, $p = 3$, and $|L| = 9$), L is uniquely defined.

Clearly, G normalizes the set $\mathcal{F}_q := \{\tau_\lambda \colon \lambda \in \mathbb{F}_q\}$, where τ_λ is defined by $v \mapsto \lambda v$ for any $v \in V$, and $\mathcal{F}_q$ is a field isomorphic to $\mathbb{F}_q$, with the operations of $\mathrm{Hom}(V, V)$. Therefore L contains $\mathcal{F}_q$ and has size $q^{d/e}$ for some divisor e of d. With the notation of [12, Section 5], we have $m = a \cdot \frac{d}{e}$ and $n^* = e$.

Then, as pointed out at the beginning of [12, Section 5], $G \leq \Gamma\mathrm{L}_q(e, q^{d/e})$, and the arguments of Hering's papers yield the claim.

3 Main result

This section is devoted to the proof of Main Theorem. To this aim, we first need to discuss the embeddings of $\Gamma\mathrm{L}_q(n, q^m)$ in $\mathrm{GL}(nm, q)$.

3.1 Embedding $\Gamma\mathrm{L}_q(n, q^m)$ in $\mathrm{GL}(nm, q)$

For a positive integer m, consider the field extension $\mathbb{F}_{q^m} : \mathbb{F}_q$. Let V be an n-dimensional vector space over $\mathbb{F}_{q^m}$. Then clearly V is an nm-dimensional vector space over $\mathbb{F}_q$ and any $\mathbb{F}_{q^m}$-linear automorphism of V is also an $\mathbb{F}_q$-linear automorphism. This provides the so-called *natural embedding* of the group of $\mathbb{F}_{q^m}$-linear automorphisms of V into the group of $\mathbb{F}_q$-linear automorphisms of V.

If a basis $\mathcal{A}$ of V over $\mathbb{F}_{q^m}$ and a basis $\mathcal{C}$ of V over $\mathbb{F}_q$ are fixed, then clearly the natural embedding induces an embedding $\eta_{\mathcal{A},\mathcal{C}}$ of $\mathrm{GL}(n, q^m)$ in $\mathrm{GL}(nm, q)$, which is again callled a natural embedding. Explicitly, for $T \in \mathrm{GL}(n, q^m)$, the matrix $\eta_{\mathcal{A},\mathcal{C}}(T)$ acts on a vector x in $\mathbb{F}_q^{nm}$ as follows: first, consider the vector $v \in V$ such that $x = (v)_{\mathcal{C}}$ (i.e. the vector of coordinates of v over the basis $\mathcal{C}$); then let $y = (v)_{\mathcal{A}} \in \mathbb{F}_{q^m}^n$ and compute $z = Ty$; let $w \in V$ be the vector such that $z = (w)_{\mathcal{A}}$; finally, the image of x by $\eta_{\mathcal{A},\mathcal{C}}(T)$ is $(w)_{\mathcal{C}} \in \mathbb{F}_q^{nm}$, that is

$$\eta_{\mathcal{A},\mathcal{C}}(T) \cdot (v)_{\mathcal{C}} = (w)_{\mathcal{C}}. \tag{2}$$

As far as the image of $\mathrm{GL}(n, q^m)$ in $\mathrm{GL}(nm, q)$ by a natural embedding is concerned, it is straightforward to check that different choices of bases produce conjugate subgroups of $\mathrm{GL}(nm, q)$.

If both n and m are odd, then any embedding of $\mathrm{GL}(n, q^m)$ in $\mathrm{GL}(nm, q)$ is actually natural (or, equivalently, all subgroups of $\mathrm{GL}(nm, q)$ which are isomorphic to $\mathrm{GL}(n, q^m)$ are conjugate). As we could not find a reference for this fact, we deduce it from a result by Kantor.

Proposition 2. *Any embedding of* $\mathrm{GL}(n, q^m)$, $nm > 1$, n, m *odd, in* $\mathrm{GL}(nm, q)$ *is natural.*

Proof. Let η' be any embedding of $\mathrm{GL}(n, q^m)$ in $\mathrm{GL}(nm, q)$. The group $G = \eta'(\mathrm{GL}(n, q^m))$ contains a Singer cycle of $\mathrm{GL}(nm, q)$ (of order $q^{nm} - 1$). By [16], $G \trianglerighteq \mathrm{GL}(nm/s, q^s)$, embedded naturally for some divisor s of nm. Clearly $s \geq m$, otherwise $|G| < |\mathrm{GL}(nm/s, q^s)|$.

- Suppose $s = nm$. Then $\mathrm{GL}(1, q^{nm})$ is a normal subgroup of G. Hence G is contained in the normalizer (in $\mathrm{GL}(nm, q)$) of $\mathrm{GL}(1, q^{nm})$, which by [14, Sect.II.7] is equal to $\mathrm{GL}(1, q^{nm}) \rtimes \mathrm{Aut}(\mathbb{F}_q^{nm} : \mathbb{F}_q)$. But $|G| > |\mathrm{GL}(1, q^{nm}) \rtimes \mathrm{Aut}(\mathbb{F}_q^{nm} : \mathbb{F}_q)|$, a contradiction.
- Suppose that $m < s < nm$. Since, by assumption, both n and m are odd, $nm \neq 2s$. Following the proof of [16, page 232], $G < \Gamma L(nm/s, q^s)$. Thus $|\mathrm{GL}(n, q^m)| = |G| < |\Gamma L(nm/s, q^s)|$, a contradiction.

Therefore $s = m$ and the embedding η' is natural.

For the rest of the paper it is convenient to write explicitly a (natural) embedding of $\mathrm{GL}(n, q^m)$ into $\mathrm{GL}(nm, q)$.

Let γ be a primitive element of $\mathbb{F}_{q^m}$, so that $\mathcal{B} = \{1, \gamma, \ldots, \gamma^{m-1}\}$ is an $\mathbb{F}_q$-basis of $\mathbb{F}_{q^m}$. Consider also an $\mathbb{F}_{q^m}$-basis $\mathcal{A} = \{\alpha_1, \ldots, \alpha_n\}$ of $\mathbb{F}_{q^{nm}}$. Clearly, the set

$$\mathcal{C} = \{\alpha_1, \alpha_1\gamma \ldots, \alpha_1\gamma^{m-1}, \ldots, \alpha_n, \alpha_n\gamma, \ldots, \alpha_n\gamma^{m-1}\} \tag{3}$$

is an $\mathbb{F}_q$-basis of $\mathbb{F}_{q^{nm}}$. We are going to describe $\eta_{\mathcal{A},\mathcal{C}}$ explicitly, as defined in (2).

Consider the map $\varphi : \mathbb{F}_{q^m} \to \mathbb{F}_q^{m\times m}$ which maps $a \in \mathbb{F}_{q^m}$ to the matrix whose $(i+1)$-th column consists of the components of $\gamma^i \cdot a$ over the basis $\mathcal{B}$, for $i = 0, \ldots, m-1$. In more explicit terms, if C is the $m \times m$ matrix over $\mathbb{F}_q$ representing the $\mathbb{F}_q$-linear map of $\mathbb{F}_{q^m}$ $x \mapsto \gamma \cdot x$ with respect to $\mathcal{B}$, then

$$\varphi(a) = \left((a)_{\mathcal{B}} \middle| C \cdot (a)_{\mathcal{B}} \middle| \cdots \middle| C^{m-1} \cdot (a)_{\mathcal{B}} \right). \tag{4}$$

It is well-known that C is the companion matrix of the minimal polynomial of γ over $\mathbb{F}_q$. Now, if $T = (a_{i,j})_{i,j=1,\ldots,n} \in \mathrm{GL}(n, q^m)$ then it is straightforward to check that $\eta_{\mathcal{A},\mathcal{C}}(T) = (\varphi(a_{i,j}))_{i,j=1,\ldots,n} \in GL(nm, q)$.

Let nm be odd and consider an embedding ζ of $\Gamma\mathrm{L}_q(n, q^m) = \mathrm{GL}(n, q^m) \rtimes \mathrm{Aut}(\mathbb{F}_{q^m} : \mathbb{F}_q)$ in $\mathrm{GL}(nm, q)$. By Proposition 2, $\zeta_{|\mathrm{GL}(n,q^m)}$ is a natural embedding. Therefore, there exists an inner automorphism β of $GL(nm, q)$ such that $(\beta \circ \zeta)(T) = (\varphi(a_{i,j}))_{i,j=1,\ldots,n}$ for $T = (a_{i,j})_{i,j=1,\ldots,n} \in \mathrm{GL}(n, q^m)$ with φ as in (4).

Let $M = (M_{i,j})_{i,j=1,\ldots,n} \in \mathrm{GL}(nm, q)$ be the image $(\beta\circ\zeta)(\phi)$ of the Frobenius map $\phi : a \mapsto a^q$ in $\mathrm{Aut}(\mathbb{F}_{q^m} : \mathbb{F}_q)$; here, $M_{i,j}$ is an $m \times m$ matrix over $\mathbb{F}_q$ for any $i, j = 1, \ldots, n$. For any $a \in \mathbb{F}_{q^m}$ we have

$$diag(a, 1, \ldots, 1) \circ \phi = \phi \circ diag(a^{q^{m-1}}, 1, \ldots, 1),$$

and hence

$$(\beta \circ \zeta)(diag(a, 1, \ldots, 1) \circ \phi) = (\beta \circ \zeta)(\phi \circ diag(a^{q^{m-1}}, 1, \ldots, 1)).$$

Therefore

$$diag(\varphi(a), \varphi(1), \ldots, \varphi(1)) \cdot M = M \cdot diag(\varphi(a^{q^{m-1}}), \varphi(1), \ldots, \varphi(1)),$$

that is

$$\begin{pmatrix} \varphi(a)\cdot M_{1,1} & \varphi(a)\cdot M_{1,2} & \cdots & \varphi(a)\cdot M_{1,n} \\ \varphi(1)\cdot M_{2,1} & \varphi(1)\cdot M_{2,2} & \cdots & \varphi(1)\cdot M_{2,n} \\ \vdots & \vdots & & \vdots \\ \varphi(1)\cdot M_{n,1} & \varphi(1)\cdot M_{n,2} & \cdots & \varphi(1)\cdot M_{n,n} \end{pmatrix} = \begin{pmatrix} M_{1,1}\cdot\varphi(a^{q^{m-1}}) & M_{1,2}\cdot\varphi(1) & \cdots & M_{1,n}\cdot\varphi(1) \\ M_{2,1}\cdot\varphi(a^{q^{m-1}}) & M_{2,2}\cdot\varphi(1) & \cdots & M_{2,n}\cdot\varphi(1) \\ \vdots & \vdots & & \vdots \\ M_{n,1}\cdot\varphi(a^{q^{m-1}}) & M_{n,2}\cdot\varphi(1) & \cdots & M_{n,n}\cdot\varphi(1) \end{pmatrix}.$$

In particular, for each $i \neq 1$ and each $a \in \mathbb{F}_{q^m}^*$, $M_{1,i}\cdot\varphi(1) = \varphi(a)\cdot M_{1,i}$. Note that $\det(\varphi(a) - \varphi(b)) \neq 0$ for each $a \neq b \in \mathbb{F}_{q^m}^*$. Therefore $M_{1,i} = \overline{O}$ (the zero matrix) for $i = 2, \ldots, n$. The same argument applies to each $M_{i,j}$ with $i \neq j$ and thus

$$M = diag(\overline{M}, \ldots, \overline{M}), \tag{5}$$

where $\overline{M}$ is provided by the (unique) embedding of $\Gamma\mathrm{L}_q(1, q^m)$ into $\mathrm{GL}(m, q)$ associated with the fixed $\mathbb{F}_q$-basis $\mathcal{B}$ of $\mathbb{F}_{q^m}$; see [14, Sect.II.7].

Observe that $\Gamma\mathrm{L}_q(n, q^m)$ contains a unique subgroup isomorphic to $\mathrm{SL}(n, q^m)$; this can be shown by noting that, for any pair $(n, q^m) \neq (2, 2), (2, 3)$, the subgroup $\mathrm{SL}(n, q^m)$ is the last term $\Gamma\mathrm{L}_q(n, q^m)^{(\infty)}$ of the commutator series of $\Gamma\mathrm{L}_q(n, q^m)$.

3.2 Proof of Main Theorem

Throughout this section, the notation on Galois extensions of global function fields is as in [8, Section 2]; see also [22]. For an ℓ-normalized $\mathbb{F}_q$-linearized polynomial $f(x) \in \mathbb{F}_{q^n}[x]$, let s be a transcendental over $\mathbb{F}_{q^n}$ and S be the splitting field of $f(x) - sx^{q^\ell} \in \mathbb{F}_{q^n}(s)[x]$ over $\mathbb{F}_{q^n}(s)$. For any positive integer r, denote by S_r the compositum function field $S \cdot \mathbb{F}_{q^{nr}}$, by k_r the field of constants of S_r, by G_r^{arith} and G_r^{geom} the arithmetic and the geometric Galois group of $S_r : \mathbb{F}_{q^{nr}}(s)$, and by φ_r the isomorphism $G_r^{\mathrm{arith}}/G_r^{\mathrm{geom}} \to \mathrm{Gal}(k_r : \mathbb{F}_{q^{nr}})$. By [8, Lemma 2.2] there exists a constant $C > 0$ depending on $S : \mathbb{F}_{q^n}(s)$ such that for any r satisfying $q^{nr} > C$ the following property holds: every $\gamma \in G_r^{\mathrm{arith}}$ such that $\varphi_r(\gamma)$ is the Frobenius automorphism for the extension $k_r : \mathbb{F}_{q^{nr}}$ is also a Frobenius at a rational unramified place of $\mathbb{F}_{q^{nr}}(s)$ different from the pole of s. For the rest of the paper, r is assumed to satisfy $q^{nr} > C$. Under this assumption, for scattered polynomials it holds $G_r^{\mathrm{arith}} \neq G_r^{\mathrm{geom}}$; see [8, Corollary 2.8].

Remark 1. Let η be any embedding of $\Gamma\mathrm{L}_q(e, q^{d/e}) = \mathrm{GL}(e, q^{d/e}) \rtimes \mathrm{Aut}(\mathbb{F}_{q^{d/e}} : \mathbb{F}_q) = \mathrm{GL}(e, q^{d/e}) \rtimes \langle \phi \rangle$ in $\mathrm{GL}(d, q)$. By Section 3.1, we can assume up to conjugation that $\eta_{|\mathrm{GL}(e,q^{d/e})} = \eta_{\mathcal{A},\mathcal{C}}$ for some $\mathbb{F}_{q^m}$-basis $\mathcal{A}$ of $\mathbb{F}_{q^{nm}}$ and $\mathcal{C}$ as in (3), such that $\eta(\phi) = M$ is as in (5).

Proposition 3. *Suppose that $d > 2$ and that there exists $e \mid d$ such that $e > 2$ and $\mathrm{SL}(e, q^{d/e}) \lhd G_r^{\mathrm{geom}} \lhd G_r^{\mathrm{arith}} \leq \Gamma L_q(e, q^{d/e})$. Then for any $\gamma \in G_r^{\mathrm{arith}}$ there exists $\alpha \in G_r^{\mathrm{geom}}$ with $rank(\eta(\alpha\gamma) - \mathbb{I}_d) < d - 1$.*

Proof. By Remark 1, it is enough to prove the claim for the case when the embedding η satisfies $\eta_{|\mathrm{GL}(e,q^{d/e})} = \eta_{\mathcal{A},\mathcal{C}}$ and $\eta(\phi) = M$, since the existence of $\alpha \in G_r^{\mathrm{geom}}$ with $rank(\eta(\alpha\gamma) - \mathbb{I}_d) < d - 1$ is invariant under conjugation in $\mathrm{GL}(d, q)$.

Write $\gamma = \beta\phi^j \in G_r^{\mathrm{arith}}$ with $\beta \in \mathrm{GL}(e, q^{d/e})$ and $j \in \{1, \ldots, d/e\}$, and write $B = \eta(\beta)$, so that $\eta(\gamma) = B \cdot M^j \in \mathrm{GL}(d, q)$. We aim to determine $A = \eta(\alpha) \in \eta(\mathrm{SL}(e, q^{d/e}))$ such that $rank(A \cdot B \cdot M^j - \mathbb{I}_d) < d - 1$. Since $A \cdot B \cdot M^j = M^j \cdot \overline{A} \cdot \overline{B}$ for some $\overline{A} \in \eta(\mathrm{SL}(e, q^{d/e}))$, $\overline{B} \in \eta(\mathrm{GL}(e, q^{d/e}))$, it is enough to find $\overline{A} \in \eta(\mathrm{SL}(e, q^{d/e}))$ such that $rank(M^j \cdot \overline{A} \cdot \overline{B} - \mathbb{I}_d) < d - 1$.

Let $(x_1^{(1)}, \ldots, x_{d/e}^{(1)}, \ldots, x_1^{(e)}, \ldots, x_{d/e}^{(e)})$ and $(\overline{x}_1^{(1)}, \ldots, \overline{x}_{d/e}^{(1)}, \ldots, \overline{x}_1^{(e)}, \ldots, \overline{x}_{d/e}^{(e)})$ be respectively the first and the $(d/e + 1)$-th column of $M^{d/e-j} = (M^j)^{-1}$ and let $y^{(i)} \in \mathbb{F}_{q^{d/e}}$, $i = 1, \ldots, e$, and $\overline{y}^{(i)} \in \mathbb{F}_{q^{d/e}}$, $i = 1, \ldots, e$, be such that $y_{\mathcal{B}}^{(i)} = (C^{i-1})^{-1}(x_1^{(i)}, \ldots, x_{d/e}^{(i)})$ and $\overline{y}_{\mathcal{B}}^{(i)} = (C^{i-1})^{-1}(\overline{x}_1^{(i)}, \ldots, \overline{x}_{d/e}^{(i)})$; here, C is the matrix defined in the proof of Proposition 2.

Consider a matrix $D \in \mathrm{GL}(e, q^{d/e})$ whose first two columns are

$$(y^{(1)}, y^{(2)}, \ldots, y^{(e)}) = (y^{(1)}, 0, \ldots, 0) \quad \text{and} \quad (\overline{y}^{(1)}, \overline{y}^{(2)}, \ldots, \overline{y}^{(e)}) = (0, \overline{y}^{(2)}, 0, \ldots, 0)$$

and such that $\det(D) = \det(\eta^{-1}(\overline{B}))$; such a matrix D exists, because $e > 2$.

Now, consider the matrix $\overline{A} := \eta(D) \cdot \overline{B}^{-1} \in \eta(\mathrm{SL}(e, q^{d/e}))$. It is readily seen that the first and the $(d/e + 1)$-th column of $M^j \cdot \overline{A} \cdot \overline{B}$ are

$$(1, 0, \ldots, 0) \quad \text{and} \quad (0, 0, \ldots, 0, \underbrace{1}_{(d/e+1)}, 0, \ldots, 0),$$

respectively. This shows that $rank(M^j \cdot \overline{A} \cdot \overline{B} - \mathbb{I}_d) \leq d-2$ and the claim follows.

Theorem 4. *Let $f(x) \in \mathbb{F}_{q^n}[x]$ be an ℓ-normalized $\mathbb{F}_q$-linearized polynomial of q-degree $k < n$, and let $d := \max\{k, \ell\}$. Then $f(x)$ is not exceptional scattered unless one of the following holds:*

- *$f(x)$ is of pseudoregulus type; or*
- *$G_r^{\mathrm{arith}} \not\leq \Gamma\mathrm{L}_q(1, q^d)$ and $\mathrm{SL}(e, q^{d/e}) \not\leq G_r^{\mathrm{geom}}$ for any divisor $e > 2$ of d.*

Proof. The case $q^d = 9^3$ follows by [8]. Suppose that $f(x)$ is exceptional scattered and not of pseudoregulus type.

- Suppose that $\mathrm{SL}(e, q^{d/e}) \leq G_r^{\mathrm{geom}}$ for some divisor $e > 2$ of d. By Theorem 1,
$$\mathrm{SL}(e, q^{d/e}) \lhd G_r^{\mathrm{geom}} \lhd G_r^{\mathrm{arith}} \leq \Gamma L_q(e, q^{d/e}).$$
Then by Proposition 3 there exist $\gamma \in G_r^{\mathrm{arith}}$ and $\alpha \in G_r^{\mathrm{geom}}$ such that $rank(\eta(\alpha\gamma) - \mathbb{I}_d) < d-1$. By [8, Theorem 2.7], $f(x)$ is not exceptional scattered, a contradiction.
- Suppose that
$$G_r^{\mathrm{geom}} \lhd G_r^{\mathrm{arith}} \leq \Gamma L_q(1, q^d).$$
Now we argue as in [8, Section 4]. Since $f(x) - sx^{q^\ell} \in \mathbb{F}_{q^n}(s)[x]$ has exactly q^d-1 non-zero roots, the transitivity of G_r^{geom} on such roots implies $(q^d-1) \mid |G_r^{\mathrm{geom}}|$. Thus, $|G_r^{\mathrm{geom}}| = i(q^d-1)$ and $|G_r^{\mathrm{arith}}| = j(q^d-1)$ with $1 \leq i \mid j \leq d$. As in [8, Proof of Theorem 1.4], one gets that $|q^\ell - q^k|$ divides $i(q^d-1)$.
Suppose $\ell < k = d$. Then
$$\frac{q^k - q^\ell}{q^{\gcd(k,\ell)} - 1} \,\Big|\, i < r. \tag{6}$$
By considering separately $\ell \mid k$ or $\ell \nmid k$, one gets $k > q^{k/2} - 1$, and hence, by direct computations, a contradiction to (6).
Then $k < \ell = d$ and
$$\frac{q^\ell - q^k}{q^{\gcd(k,\ell)} - 1} \,\Big|\, i < \ell.$$
If $k > 0$, then $\ell > q^{\ell/2} - 1$ with $\ell > 1$, whence a contradiction as above. Then $k = 0$, $(x^{q^\ell}, f(x)) = (x^{q^\ell}, x)$, and thus $f(x)$ is of pseudoregulus type, a contradiction.

It is readily seen that Main Theorem follows from Theorems 1 and 4, which provide a partial classification also for the d even case.

Theorem 5. *Let $f(x) \in \mathbb{F}_{q^n}[x]$ be an ℓ-normalized $\mathbb{F}_q$-linearized polynomial of q-degree $k < n$, and let $d := \max\{k, \ell\}$. Then $f(x)$ is not exceptional scattered unless one of the following holds:*

1. $\mathrm{SL}(2, q^{d/2}) \lhd G_r^{\mathrm{geom}} \lhd G_r^{\mathrm{arith}} \leq \Gamma\mathrm{L}_q(2, q^{d/2})$;
2. $Sp(e, q^{d/e}) \lhd G_r^{\mathrm{geom}} \lhd G_r^{\mathrm{arith}} \leq \Gamma\mathrm{L}_q(e, q^{d/e})$, *for some even $e \mid d$ with $e \geq 4$;*
3. $G_2(2^{d/6})' \lhd G_r^{\mathrm{geom}} \lhd G_r^{\mathrm{arith}} \leq \Gamma\mathrm{L}_q(6, 2^{d/6})$, *where $p = 2$ and $6 \mid d$;*
4. $q^d \in \{5^2, 7^2, 11^2, 23^2, 29^2, 59^2, 2^4, 3^4, 3^6\}$.

References

1. Ball, S., Blokhuis, A., and Lavrauw, M. Linear $(q+1)$-fold blocking sets in $\mathrm{PG}(2,q^4)$. *Finite Fields Appl. 6*, 4 (2000), 294–301.
2. Bartoli, D., Giulietti, M., Marino, G., and Polverino, O. Maximum scattered linear sets and complete caps in Galois spaces. *Combinatorica 38*, 2 (2018), 255–278.
3. Bartoli, D., and Montanucci, M. On the classification of exceptional scattered polynomials. *J. Combin. Theory Ser. A 179* (2021), 105386, 28.
4. Bartoli, D., and Zhou, Y. Exceptional scattered polynomials. *J. Algebra 509* (2018), 507–534.
5. Bartoli, D., Zini, G., and Zullo, F. Investigating the exceptionality of scattered polynomials. *Finite Fields Appl. 77* (2022), 101956.
6. Blokhuis, A., and Lavrauw, M. Scattered spaces with respect to a spread in $\mathrm{PG}(n,q)$. *Geom. Dedicata 81*, 1-3 (2000), 231–243.
7. Calderbank, R., and Kantor, W. M. The geometry of two-weight codes. *Bull. London Math. Soc. 18*, 2 (1986), 97–122.
8. Ferraguti, A., and Micheli, G. Exceptional scatteredness in prime degree. *J. Algebra 565* (2021), 691–701.
9. Gabidulin, E. M., Paramonov, A., and Tretjakov, O. Ideals over a non-commutative ring and their application in cryptology. In *Workshop on the Theory and Application of of Cryptographic Techniques* (1991), Springer, pp. 482–489.
10. Glynn, D. G., and Steinke, G. F. Laguerre planes of even order and translation ovals. *Geom. Dedicata 51*, 2 (1994), 105–112.
11. Hering, C. Zweifach transitive Permutationsgruppen, in denen 2 die maximale Anzahl von Fixpunkten von Involutionen ist. *Math. Z. 104* (1968), 150–174.
12. Hering, C. Transitive linear groups and linear groups which contain irreducible subgroups of prime order. *Geometriae Dedicata 2* (1974), 425–460.
13. Hering, C. Transitive linear groups and linear groups which contain irreducible subgroups of prime order. II. *J. Algebra 93*, 1 (1985), 151–164.
14. Huppert, B. *Endliche Gruppen. I.* Die Grundlehren der Mathematischen Wissenschaften, Band 134. Springer-Verlag, Berlin-New York, 1967.
15. Johnson, N. L., Jha, V., and Biliotti, M. *Handbook of finite translation planes*, vol. 289 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, 2007.
16. Kantor, W. M. Linear groups containing a Singer cycle. *J. Algebra 62*, 1 (1980), 232–234.
17. Lavrauw, M., and Polverino, O. Finite semifields. In *Current research topics in Galois geometry*. New York, NY: Nova Science Publishers/Novinka, 2014, pp. 131–159.
18. Lavrauw, M., and Van de Voorde, G. Field reduction and linear sets in finite geometry. In *Topics in finite fields*, vol. 632 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 2015, pp. 271–293.
19. Lunardon, G., Marino, G., Polverino, O., and Trombetti, R. Maximum scattered linear sets of pseudoregulus type and the Segre variety $\mathcal{S}_{n,n}$. *J. Algebraic Combin. 39*, 4 (2014), 807–831.
20. Lunardon, G., and Polverino, O. Blocking sets of size $q^t+q^{t-1}+1$. *J. Combin. Theory Ser. A 90*, 1 (2000), 148–158.
21. Marino, G., and Polverino, O. On translation spreads of $H(q)$. *J. Algebraic Combin. 42*, 3 (2015), 725–744.

22. Micheli, G. Constructions of locally recoverable codes which are optimal. *IEEE Trans. Inform. Theory 66*, 1 (2020), 167–175.
23. Polverino, O., and Zullo, F. Connections between scattered linear sets and MRD-codes. *Bull. Inst. Combin. Appl. 89* (2020), 46–74.
24. Sheekey, J. A new family of linear maximum rank distance codes. *Advances in Mathematics of Communications 10*, 3 (2016), 475–488.
25. Silva, D., Kschischang, F. R., and Kötter, R. A rank-metric approach to error control in random network coding. *IEEE Trans. Inform. Theory 54*, 9 (2008), 3951–3967.
26. Zini, G., and Zullo, F. Scattered subspaces and related codes. *Des. Codes Cryptogr. 89*, 8 (2021), 1853–1873.

PIR Codes, Unequal-Data-Demand Codes, and the Griesmer Bound

Henk D.L. Hollmann[1][0000−0003−4005−2369], Martin Puškin[2], and Ago-Erik Riet[2][0000−0002−8310−6809]

[1] Institute of Computer Science, University of Tartu, Tartu 50409, Estonia
[2] Institute of Mathematics and Statistics, University of Tartu, Tartu 50409, Estonia
{henk.hollmann,martin.puskin,ago-erik.riet}@ut.ee

Abstract. Unequal Error-Protecting (UEP) codes are error-correcting (EC) designed to protect some parts of the encoded data better than other parts. Here, we introduce a similar generalization of PIR codes that we call Unequal-Data-Demand (UDD) PIR codes. These codes are PIR-type codes designed for the scenario where some parts of the encoded data are in higher demand than other parts. We generalize various results for PIR codes to UDD codes. Our main contribution is a new approach to the Griesmer bound for linear EC codes involving an Integer Linear Programming (ILP) problem that generalizes to linear UEP codes and linear UDD PIR codes.

Keywords: PIR codes · UEP codes · Griesmer bound.

1 Introduction

A t-PIR code stores a data record in encoded form on a collection of servers in such a way that the data symbol in any position in the record can be recovered from the encoded data symbols stored by any of t disjoint groups of servers; such a group of servers is called a *recovery set* for that position. We refer to Section 5 for a more formal definition of PIR codes.

A Private Information Retrieval (PIR) scheme stores a database in encoded form on a multi-server distributed data storage system in such a way that a user can extract a bit of information from the database without leaking information about which particular bit the user was interested in. Originally, PIR codes were employed to reduce the amount of storage needed to implement such a system. Here, linear t-PIR codes can be used to implement a classical (linear) t-server PIR scheme [1] with less storage overhead than the original scheme, by using the PIR code to emulate the t servers [3]; see also [13] for a nice explanation of how PIR codes can achieve this.

Unequal-error-protecting or UEP codes are error-correcting codes that protect some parts of the encoded data better than other parts. A simple example of an UEP code can be obtained by the concatenation of two codes with different error-correcting capabilities. Interestingly, there exist UEP codes that are more efficient than any code obtained by concatenation of smaller codes, see, e.g.,

[2], [4, Chapter 1]. In analogy with UEP codes, we define *Unequal Data Demand* or UDD codes as PIR codes designed for cases where some parts of the data are more in demand, more popular, than other parts. Again, the basic question is whether we can do better than just using a concatenation of two PIR codes with different values for t? It turns out that this question again has an affirmative answer and in this paper we will give several examples to show this.

The Griesmer bound [8, Chapter 17, Theorem 24] is a famous and fundamental bound on the length of a linear error-correcting (EC) code with a given minimum distance (see Section 2 for definitions). This bound has been generalized to UEP codes in [4, Chapter 1], and to PIR codes in [10]. In this paper, we generalize the Griesmer bound to the case of UDD PIR codes. First we show that the Griesmer bound for UDD PIR codes can be obtained as a consequence of the corresponding bound for UEP codes; to this end, we first generalize a well-known bound for the minimum distance for PIR codes [5], [7], [9], [12], [14] to the case of UDD codes. We also provide an alternative, direct proof, using an Integer Linear Program (ILP) formulation. Interestingly, we show that the ILP can be used to provide a uniform proof for *all* the Griesmer bounds mentioned above.

The contents of this paper are the following. In Sections 2, 3, and 4, we briefly review error-correcting codes, PIR codes, and UEP codes, respectively. The new notion of UDD PIR codes is introduced and discussed in Section 5. In Section 6, we first derive the new distance bound for UDD PIR codes, which we then use to give a first proof of the Griesmer bound for UDD PIR codes. In Section 7 we derive an ILP bound for UDD PIR codes and use it to generalize a bound for PIR codes from [10]. We determine a lower bound to the optimum of the ILP in Section 8, which then provides a second proof of the Griesmer bound for this type of codes. In Section 9 we show that our ILP bound provides a uniform proof of the Griesmer bound for all codes mentioned earlier. A few open problems are discussed in Section 10. Finally, in Section 11 we present some conclusions. This work is based on [11], where many other results for PIR codes are generalized to UDD codes.

Throughout this paper, $\mathbb{F}_q$ denotes the finite field of order q, where q is a power of a prime p, and we write $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$. We let $\mathbb{F}_q^n$ denote the n-dimensional vector space over $\mathbb{F}_q$. For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{F}_q^n$, we let $\langle \boldsymbol{a}, \boldsymbol{b} \rangle := a_1 b_1 + \cdots a_n b_n$ denote the *inner product* of $\boldsymbol{a}$ and $\boldsymbol{b}$, see, e.g., [6]. For $\boldsymbol{h} \in \mathbb{F}_q^n$, we let $\boldsymbol{h}^\perp := \{\boldsymbol{x} \in \mathbb{F}_q^k \mid \langle \boldsymbol{x}, \boldsymbol{h} \rangle = 0\}$ denote the hyperplane with normal vector $\boldsymbol{h}$. We use the set $[n] := \{1, 2, \ldots, n\}$ to index the symbols in vectors (or codewords) $\boldsymbol{c} \in \mathbb{F}_q^n$, that is, the symbol in $\boldsymbol{c}$ with index $i \in [n]$ is c_i. We use $\mathbb{Z}_+$ to denote the set of nonnegative integers.

2 Error-correcting codes

The *Hamming weight* $w(\boldsymbol{c})$ of a vector $\boldsymbol{c} \in \mathbb{F}_q^n$ is the number of positions i in $[n]$ for which $c_i \neq 0$, and the *Hamming distance* $d(\boldsymbol{x}, \boldsymbol{y})$ between $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_q^n$ is defined as $d(\boldsymbol{x}, \boldsymbol{y}) = w(\boldsymbol{x} - \boldsymbol{y})$, that is, the number of positions in which $\boldsymbol{x}$ and $\boldsymbol{y}$

differ. Note that the Hamming distance is a *metric*, see, e.g., [8]. The minimum (Hamming) distance $d(C)$ of a set $C \subseteq \mathbb{F}_q^n$ is the minimum Hamming distance $d(\boldsymbol{x}, \boldsymbol{y})$ between two *distinct* vectors $\boldsymbol{x}, \boldsymbol{y} \in C$.

A $(n, M, d)_q$ code C is a set of M vectors from $\mathbb{F}_q^n$ with minimum distance $d(C) = d$. The elements of C are referred to as *codewords*. We say that C is *linear* if C is a linear subspace of $\mathbb{F}_q^n$; if $\dim(C) = k$ then we say that C is a $[n, k]_q$ code, or a $[n, k, d]_q$ code if $d(C) = d$. A *generator matrix* $\boldsymbol{G}$ for a linear code C is a $k \times n$ matrix with entries from $\mathbb{F}_q$ with the property that the rows of $\boldsymbol{G}$ consist of codewords that together form a *basis* for (the $\mathbb{F}_q$-linear subspace) C. An *encoder* for C is a one-to-one map $\epsilon : \mathbb{F}_q^k \to C$. Given a generator matrix $\boldsymbol{G}$ for C, the map $\epsilon : \mathbb{F}_q^k \to \mathbb{F}_q^n$ defined by $\epsilon(\boldsymbol{a}) = \boldsymbol{a}^\top \boldsymbol{G}$ is a *linear* encoder for C, referred to as *the encoder* G for C; note that any linear encoder is of this form.

3 PIR codes

We first provide a formal definition of a PIR code.

Definition 1. *Given a (one-to-one) encoder map $\epsilon : \mathbb{F}_q^k \longrightarrow \mathbb{F}_q^n$, a set of positions $I = \{i_1, \ldots, i_s\} \subseteq [n]$ is called a* recovery set for the j-th data symbol *if the restriction $\mathbf{c}_I = (c_{i_1}, \ldots, c_{i_s})$ of a codeword $\mathbf{c} = \epsilon(\boldsymbol{a})$ uniquely determines the j-th data symbol a_j. The encoder map ϵ is a t-PIR code if there exists for every $j = 1, \ldots, k$ a collection of t disjoint recovery sets for the j-th data symbol.*

We say that a $k \times n$ matrix $\boldsymbol{G}$ with entries from $\mathbb{F}_q$ is a (linear) t-PIR code if the corresponding encoder $\epsilon : \boldsymbol{a}^\top \to \boldsymbol{a}^\top \boldsymbol{G}$ is t-PIR. In that case we say that $\boldsymbol{G}$ generates *a t-PIR code, or that $\boldsymbol{G}$ is* t-PIR.

Here it is important to realize that the t-PIR property is a property of the *encoder* of the code. Note that if the span of the columns from a $k \times n$ matrix $\boldsymbol{G}$ with indices in a set of positions I contains the j-th unit vector $\boldsymbol{e}_j$ ($j \in [k]$), then I is a recovery set for the j-th data symbol in the PIR code generated by $\boldsymbol{G}$. In [9, Theorem 1] it was shown that every recovery set I arises in this way.

Example 1. Let $q = 2$, and let C be the binary linear code with (linear) encoder $\epsilon : \boldsymbol{a}^\top \to \boldsymbol{a}^\top \boldsymbol{G}$, where

$$G = \begin{pmatrix} 1\,0\,1\,1 \\ 0\,1\,1\,0 \end{pmatrix}. \tag{1}$$

Then the first data symbol has recovery sets $\{1\}, \{2, 3\}, \{4\}$ and the second data symbol has recovery sets $\{2\}, \{1, 3\}$, so this encoder and this matrix $\boldsymbol{G}$ are both 2-PIR. Indeed, note that if $\boldsymbol{c}^\top = \boldsymbol{a}^\top \boldsymbol{G} = (a_1, a_2, a_1 + a_2, a_1)$, then for example a_1 can be recovered as $c_2 + c_3$, in accordance with the fact that $\boldsymbol{e}_1$ is the sum of the second and third column of $\boldsymbol{G}$. Since every recovery set for the second data symbol except $\{2\}$ has size at least 2, that data symbol does not have 3 disjoint recovery sets and so $\boldsymbol{G}$ is not a 3-PIR code. □

4 Unequal-Error-Protection codes

An error-correcting code is designed to protect data against the occasional occurrence of errors: by sending the data in encoded form, the original data can still be recovered from the received codeword as long as not too many errors have occurred. Unequal-Error-Protecting codes play a similar role, but are designed to protect certain data symbols better than others, see, e.g., [2], [4, Chapter 1].

For an encoder map $\epsilon : \mathbb{F}_q^k \to \mathbb{F}_q^n$, define the *separation vector* $\boldsymbol{s}(\epsilon) \in \mathbb{Z}_+^k$ by defining for every $j \in [k]$

$$s_j(\epsilon) = \min\{d(\epsilon(\boldsymbol{a}), \epsilon(\boldsymbol{a}')) \mid \boldsymbol{a}, \boldsymbol{a}' \in \mathbb{F}_q^k, a_j \neq a'_j\}. \tag{2}$$

So $s_i(\epsilon)$ is just the minimal distance between two codewords in distinct subcodes $C_{i,\beta} = \{\boldsymbol{c} = \epsilon(\boldsymbol{a}) \mid a_i = \beta\}$ ($\beta \in \mathbb{F}_q$) of the code $C = \epsilon(\mathbb{F}_q^k)$. It is not difficult to see that by decoding to the nearest codeword, we can decode the i-th data symbol correctly if at most $\lfloor (s_i(\epsilon) - 1)/2 \rfloor$ errors have occurred. For more details, see, e.g., [4, Chapter 1, Section II]. We note that

$$d(C) = \min_{j \in [k]} s_i(\epsilon).$$

We will write $\boldsymbol{s}(\boldsymbol{G})$ to denote the separation vector of a linear code encoded with generator matrix $\boldsymbol{G}$.

A "trivial" construction of an UEP code is to use an $(n, q^{k_1}, d_1)_q$ code C_1 to protect part of the data, and a $(n_2, q^{k_2}, d_2)_q$ code C_2 to protect another part of the data. Then the *concatenation* of C_1 and C_2, the code with codewords $(\boldsymbol{c}_1, \boldsymbol{c}_2)$ with $\boldsymbol{c}_i \in C_i$ $(i = 1, 2)$ has a separation vector $\boldsymbol{s}(\epsilon)$ for which

$$s_i(\epsilon) \geq \begin{cases} d_1, \text{ if } i \text{ is among the first } n_1 \text{ positions;} \\ d_2, \text{ if } i \text{ is among the last } n_2 \text{ positions.} \end{cases}$$

However, often one can do better than this trivial construction.

Example 2. Suppose we want to protect two data bits against errors or erasures, and we want to realize a separation vector $\boldsymbol{s}(\epsilon) = (3, 2)$. For the trivial construction, we would need two repetition codes, one of length 3 and one of length 2, with an encoder $\epsilon(ab) = aaabb$ $(a, b \in \mathbb{F}_2)$. So the minimum length of a "trivial" construction would be 5. Now consider the linear PIR code of length 4 generated by the matrix $\boldsymbol{G}$ as in (1). Here again $\boldsymbol{s}(\boldsymbol{G}) = (3, 2)$, but now with a code of length only 4. □

It turns out that a given k-dimensional linear code C has an *optimal* generator matrix $\boldsymbol{G}^*$, in the sense that the separation vector $\boldsymbol{s}(\boldsymbol{G}^*)$ of the encoder determined by $\boldsymbol{G}^*$ is componentwise larger than or equal to the separation vector $\boldsymbol{s}(\boldsymbol{G})$ of any other generator matrix $\boldsymbol{G}$ of the code, that is, $s_j(\boldsymbol{G}^*) \geq s_j(\boldsymbol{G})$ for every $j \in [k]$, see [2], [4, Chapter 1, Section II]. This allows us to speak of $\boldsymbol{s}(\boldsymbol{G}^*)$ as *the separation vector of the code* C. Such a matrix $\boldsymbol{G}^*$ can be obtained by a greedy construction, where the first row of $\boldsymbol{G}^*$ is a vector from C with minimum weight and each further row is a vector from C of minimum weight outside the span of the rows already chosen. For further details, see [2].

5 Unequal-Data-Demand codes

t-PIR codes are designed so that up to t users can obtain each a particular data symbol from data that is stored in encoded form on a number of servers, where every server can be read off at most once. Unequal-Data-Demand (UDD) codes enable a similar scenario, but now for the situation where some parts of the data are in higher demand than other parts. We first present a formal definition.

Definition 2. *Let $T = (t_1, \ldots, t_k)$ where $t_1, \ldots, t_k$ are integers with $t_1 \geq t_2 \geq \cdots \geq t_k \geq 0$. An UDD T-PIR code of length n is an encoder $\epsilon : \mathbb{F}_q^k \longrightarrow \mathbb{F}_q^n$ where the j-data symbol has at least t_j mutually disjoint recovery sets ($j = 1, \ldots, k$).*

We say that a $k \times n$ matrix $\boldsymbol{G}$ with entries from $\mathbb{F}_q$ is a (linear) T-PIR code if the corresponding encoder $\epsilon : \boldsymbol{a}^\top \to \boldsymbol{a}^\top \boldsymbol{G}$ is T-PIR. In that case we say that $\boldsymbol{G}$ generates *a T-PIR code.*

As for UEP-codes, we can use the "trivial" construction by concatenation to obtain examples of UDD PIR-codes, but often we can do better.

Example 3. The properties of the matrix $\boldsymbol{G}$ in (1) as stated in Example 1 show that $\boldsymbol{G}$ generates a $(3, 2)$-PIR code of length 4. To achieve this by concatenation would require a length-3 repetition code and a length-2 repetition code, for a total length equal to 5. □

6 The Griesmer bound for UEP PIR codes

It is well-known that for a t-PIR code, the minimum distance d of the associated code satisfies $d \geq t$, see, e.g., [5], [7], [9], [12], [14]. We will need the following generalization.

Theorem 1. *Let C be an $(n, q^k, d)_q$-code with encoder $\epsilon : \mathbb{F}_q^k \to \mathbb{F}_q^n$, and let ϵ have separation vector $\boldsymbol{s}(\epsilon)$. If ϵ is an UDD T-PIR code, where $T = (t_1, \ldots, t_k)$ and $t_1 \geq t_2 \geq \cdots \geq t_k$, then $s_j(\epsilon) \geq t_j$ for all $j \in [k]$.*

Proof. Let $j \in [k]$. Since ϵ is T-PIR, there are t_j mutually disjoint recovery sets $I_1, \ldots, I_{t_j}$ for the j-th data symbol. Now let $\boldsymbol{a}, \boldsymbol{a}' \in \mathbb{F}_q^n$ with $a_j \neq a'_j$, and let $\mathbf{c} = \epsilon(\boldsymbol{a})$ and $\mathbf{c}' = \epsilon(\boldsymbol{a}')$ be the corresponding codewords. For every $i \in [t_j]$, since I_i determines the j-th data symbol and since $a_j \neq a'_j$, we must have $\mathbf{c}_{I_i} \neq \mathbf{c}'_{I_i}$, that is, $\mathbf{c}$ and $\mathbf{c}'$ differ in a position in I_i. We conclude that $d(\mathbf{c}, \mathbf{c}') \geq t_j$. Now the claim follows from the definition of $s_j(\epsilon)$ in (2).

We will use this result to prove the following generalization of the Griesmer bound.

Theorem 2 (Griesmer for UDD PIR codes). *Suppose that the $k \times n$ matrix $\boldsymbol{G}$ over $\mathbb{F}_q$ generates a linear UDD T-PIR code, where $T = (t_1, \ldots, t_k)$ with $t_1 \geq t_2 \geq \cdots \geq t_k \geq 0$. Then*

$$n \geq \sum_{j=1}^{k} \left\lceil \frac{t_j}{q^{j-1}} \right\rceil . \tag{3}$$

Proof. Suppose that $\boldsymbol{s}(\boldsymbol{G}) = (s_1, \ldots, s_k)$ is the separation vector of the UEP code generated by $\boldsymbol{G}$. Then by the Griesmer bound for linear UEP codes [4, Chapter I, Part III, Corollary 14] we have that $n \geq \sum_{j=1}^{k} \lceil s_j / q^{j-1} \rceil$. By Theorem 1, we have $s_j \geq t_j$, hence (3) follows immediately.

It would be nice to have an argument that would prove all these Griesmer-type bounds *simultaneously*, in a *uniform* way. In the next sections we will provide such an approach.

7 An ILP problem related to PIR codes

Fix a prime power q. There is a one-to-one correspondence between the collection of hyperplanes in $\mathbb{F}_q^k$ and the collection $\mathcal{P}_k$ of vectors $\boldsymbol{h} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}$ of the form $\boldsymbol{h} = (0, \ldots, 0, 1, \ldots)$, so with the first nonzero entry equal to 1, where a vector $\boldsymbol{h} \in \mathcal{P}_k$ corresponds to the hyperplane $\boldsymbol{h}^\perp := \{\boldsymbol{a} \in \mathbb{F}_q^k \mid \langle \boldsymbol{h}, \boldsymbol{a} \rangle = 0\}$. For later use, note that $|\mathcal{P}_k| = (q^k - 1)/(q - 1)$. (Note also that the vectors in $\mathcal{P}_k$ are in one-to-one correspondence with the points in the $(k-1)$-dimensional projective geometry $\mathrm{PG}(k-1, q)$, see, e.g., [8, Appendix B].) For $\mathbf{h} \in \mathcal{P}_k$, define

$$\nu(\boldsymbol{h}) = \min\{j \in \{1, \ldots, k\} \mid h_j \neq 0\};$$

as a consequence, $h_{\nu(\boldsymbol{h})} = 1$. We now have the following. (Here and below, for less cumbersome notation, we will write $\sum_{\mathrm{Condition}(\boldsymbol{i})} n_{\boldsymbol{i}}$ to denote the sum of all numbers $n_{\boldsymbol{i}}$ with $\boldsymbol{i} \neq \mathbf{0}$ for which $\boldsymbol{i}$ satisfies $\mathrm{Condition}(\boldsymbol{i})$.)

Theorem 3. (Cf. [10, Lemma 6]) *Let $\boldsymbol{G}$ be a $k \times n$ matrix over $\mathbb{F}_q$ that generates an UDD T-PIR code, where $T = (t_1, \ldots, t_k)$ and $t_1 \geq t_2 \geq \cdots \geq t_k$. Suppose that $\boldsymbol{G}$ has $n_{\boldsymbol{i}}$ columns equal to $\boldsymbol{i}$, for $\boldsymbol{i} \in \mathbb{F}_q^k$. Then for all $\mathbf{h} \in \mathcal{P}_k$, we have*

$$\sum_{\langle \boldsymbol{i}, \boldsymbol{h} \rangle \neq 0} n_{\boldsymbol{i}} \geq t_{\nu(\boldsymbol{h})} \tag{4}$$

Proof. To see this, note that if the j-th unit vector $\boldsymbol{e}_j$ is not contained in the hyperplane $\mathbf{h}^\perp$ ($j \in [k], \mathbf{h} \in \mathcal{P}_k$), that is, when $h_j \neq 0$, then every set of columns of $\boldsymbol{G}$ whose span contains $\boldsymbol{e}_j$ must contain a column outside $\mathbf{h}^\perp$; so by our assumptions on $\boldsymbol{G}$, there are at least t_j columns of $\boldsymbol{G}$ outside $\mathbf{h}^\perp$. Taking $j = \nu(\mathbf{h})$ gives (4).

So for $T = (t_1, \ldots, t_k) \in \mathbb{Z}^k$ with $t_1 \geq \cdots \geq t_k \geq 0$, define $\mu(T)$ to be the solution of the following Integer Linear Programming (ILP) problem:

$$\mathrm{ILP}(T) : \begin{cases} n_{\boldsymbol{i}} \in \mathbb{Z}, n_{\boldsymbol{i}} \geq 0 & (\boldsymbol{i} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}) \\ \displaystyle\sum_{\{\boldsymbol{i} \mid \langle \boldsymbol{i}, \mathbf{h} \rangle \neq 0\}} n_{\boldsymbol{i}} \geq t_{\nu(\boldsymbol{h})} & (\boldsymbol{h} \in \mathcal{P}_k) \\ \text{minimize } n = \displaystyle\sum_{\boldsymbol{i} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}} n_{\boldsymbol{i}} \end{cases} \tag{5}$$

Then, according to Theorem 3, if the $k \times n$ matrix $\boldsymbol{G}$ generates a $(t_1, \ldots, t_k)$-PIR code with $t_1 \geq \cdots \geq t_k$, then $n \geq n - n_{\mathbf{0}} \geq \mu(T)$, where for an optimal solution, we should of course take $n_{\mathbf{0}} = 0$.

Example 4. Let $q = 2$ and $k = 2$, and let $T = (t_1, t_2) \in \mathbb{Z}^3$ with $t_1 \geq t_2 \geq 0$. Associating the numbers $1, 2, 3$ with the vectors $(1, 0)$, $(0, 1)$, and $(1, 1)$, the ILP(T) is the problem to minimize $n = n_1 + n_2 + n_3$, where $n_i \geq 0$ is integer $(i = 1, 2, 3)$, under the conditions

$$n_1 + n_3 \geq t_1 \tag{6}$$
$$n_2 + n_3 \geq t_2 \tag{7}$$
$$n_1 + n_2 \geq t_1, \tag{8}$$

where the inequalities correspond to the hyperplanes $(1, 0)^\perp$, $(0, 1)^\perp$, and $(1, 1)^\perp$, respectively. It is not difficult to see that the minimum value for n under these conditions equals $t_1 + \lceil t_2/2 \rceil$. □

It is easy to give a lower bound for $\mu(T)$.

Proposition 1. *We have that* $\mu(T) \geq \sum_{j=1}^{k} t_j / q^{j-1}$.

Proof. If $\boldsymbol{i} \neq \mathbf{0}$, then $|\boldsymbol{i}^\perp| = q^{k-1}$, hence $|\mathbb{F}_q^k \setminus \boldsymbol{i}^\perp| = q^k - q^{k-1} = q^{k-1}(q-1)$. So there are q^{k-1} vectors $\boldsymbol{h} \in \mathcal{P}_k$ such that $\langle \boldsymbol{i}, \boldsymbol{h} \rangle \neq 0$. As a consequence, using (5) we have

$$\sum_{j=1}^{k} t_j q^{k-j} = \sum_{\boldsymbol{h} \in \mathcal{P}_k} t_{\nu(\boldsymbol{h})} \leq \sum_{\boldsymbol{h} \in \mathcal{P}_k} \sum_{\langle \boldsymbol{i}, \boldsymbol{h} \rangle \neq 0} n_{\boldsymbol{i}} = \sum_{\boldsymbol{i} \neq \mathbf{0}} \sum_{\boldsymbol{h} \in \mathcal{P}_k, \langle \boldsymbol{i}, \boldsymbol{h} \rangle \neq 0} n_{\boldsymbol{i}} = \sum_{\boldsymbol{i} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}} q^{k-1} n_{\boldsymbol{i}},$$

so $n \geq \sum_{\boldsymbol{i} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}} n_{\boldsymbol{i}} \geq \sum_{j=1}^{k} t_j / q^{j-1}$.

In the next section, we will provide a better bound for $\mu(T)$.

8 A sharper lower bound for the ILP problem

Our aim is to prove the following theorem.

Theorem 4. *Let $\mu(T)$ be the optimal solution of the ILP problem (5), where $\boldsymbol{G}$ and T are as in Theorem 3. Then*

$$\mu(T) \geq \sum_{j=1}^{k} \left\lceil \frac{t_j}{q^{j-1}} \right\rceil. \tag{9}$$

Proof. To prove this, we will use induction on the dimension k. First note that the theorem obviously holds for $k = 1$. Assume that the theorem holds for dimension $k - 1$, and suppose that the $n_{\boldsymbol{i}}$ with $\boldsymbol{i} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}$ satisfy the ILP constraints.

First consider the q^{k-1} inequalities in (5) that involve t_1. Let $\delta \in \mathbb{Z}_+$ and $\boldsymbol{b} \in \mathcal{P}_k$ with $\nu(\boldsymbol{b}) = 1$ be such that for all $\boldsymbol{u} \in \mathcal{P}_k$ with $\nu(\boldsymbol{u}) = 1$, we have

$$\sum_{\langle \boldsymbol{i},\boldsymbol{u}\rangle \neq 0} n_{\boldsymbol{i}} \geq \sum_{\langle \boldsymbol{i},\boldsymbol{b}\rangle \neq 0} n_{\boldsymbol{i}} = t_1 + \delta. \tag{10}$$

Then for every $\boldsymbol{u} \in \mathcal{P}_k$ with $\nu(\boldsymbol{u}) = 1$, we have that $\sum_{\langle \boldsymbol{i},\boldsymbol{u}\rangle \neq 0} n_{\boldsymbol{i}} \geq \sum_{\langle \boldsymbol{i},\boldsymbol{b}\rangle \neq 0} n_{\boldsymbol{i}}$, and hence

$$\sum_{\substack{\langle \boldsymbol{i},\boldsymbol{u}\rangle \neq 0 \\ \langle \boldsymbol{i},\boldsymbol{b}\rangle = 0}} n_{\boldsymbol{i}} \geq \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{b}\rangle \neq 0 \\ \langle \boldsymbol{i},\boldsymbol{u}\rangle = 0}} n_{\boldsymbol{i}}. \tag{11}$$

Fix $\boldsymbol{h}' \in \mathcal{P}_{k-1}$, and set $\boldsymbol{h} = (0, \boldsymbol{h}')$. Note that $\boldsymbol{h} \in \mathcal{P}_k$ and $\nu(\boldsymbol{h}) = 1 + \nu(\boldsymbol{h}') > 1$. For every $\lambda \in \mathbb{F}_q$, define $\boldsymbol{u}_\lambda := \boldsymbol{b} + \lambda \boldsymbol{h}$; note that $\boldsymbol{u}_\lambda \in \mathcal{P}_k$ and $\nu(\boldsymbol{u}_\lambda) = 1$, so that (11) holds for $\boldsymbol{u} = \boldsymbol{u}_\lambda$. Note also that $\langle \boldsymbol{i}, \boldsymbol{u}_\lambda\rangle = \langle \boldsymbol{i}, \boldsymbol{b}\rangle + \lambda \langle \boldsymbol{i}, \boldsymbol{h}\rangle$. Now

$$t_{\nu(\boldsymbol{h})} \leq \sum_{\langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0} n_{\boldsymbol{i}} = \sum_{\lambda \in \mathbb{F}_q} \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0 \\ \langle \boldsymbol{i},\boldsymbol{b}\rangle = -\lambda\langle \boldsymbol{i},\boldsymbol{h}\rangle}} n_{\boldsymbol{i}} = \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0 \\ \langle \boldsymbol{i},\boldsymbol{b}\rangle = 0}} n_{\boldsymbol{i}} + \sum_{\lambda \in \mathbb{F}_q^*} \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0 \\ \langle \boldsymbol{i},\boldsymbol{u}_\lambda\rangle = 0}} n_{\boldsymbol{i}}. \tag{12}$$

Now note that if $\lambda \neq 0$ and $\langle \boldsymbol{i}, \boldsymbol{u}_\lambda\rangle = 0$, then $\langle \boldsymbol{i}, \boldsymbol{h}\rangle = 0$ holds if and only if $\langle \boldsymbol{i}, \boldsymbol{b}\rangle = 0$; note also that if $\lambda \neq 0$ and $\langle \boldsymbol{i}, \boldsymbol{b}\rangle = 0$, then $\langle \boldsymbol{i}, \boldsymbol{u}_\lambda\rangle = 0$ if and only if $\langle \boldsymbol{i}, \boldsymbol{h}\rangle = 0$. So noting that $\nu(\boldsymbol{u}_\lambda) = 1$ and using (11), we find that

$$\sum_{\substack{\langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0 \\ \langle \boldsymbol{i},\boldsymbol{u}_\lambda\rangle = 0}} n_{\boldsymbol{i}} = \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{b}\rangle \neq 0 \\ \langle \boldsymbol{i},\boldsymbol{u}_\lambda\rangle = 0}} n_{\boldsymbol{i}} \leq \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{b}\rangle = 0 \\ \langle \boldsymbol{i},\boldsymbol{u}_\lambda\rangle \neq 0}} n_{\boldsymbol{i}} = \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{b}\rangle = 0 \\ \langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0}} n_{\boldsymbol{i}} \tag{13}$$

So by combining (12) and (13) and recalling that $\nu(\boldsymbol{h}) = 1 + \nu(\boldsymbol{h}')$, we conclude that

$$t_{1+\nu(\boldsymbol{h}')} = t_{\nu(\boldsymbol{h})} \leq \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0 \\ \langle \boldsymbol{i},\boldsymbol{b}\rangle = 0}} n_{\boldsymbol{i}} + \sum_{\lambda \in \mathbb{F}_q^*} \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{b}\rangle = 0 \\ \langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0}} n_{\boldsymbol{i}} = q \sum_{\substack{\langle \boldsymbol{i},\boldsymbol{b}\rangle = 0 \\ \langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0}} n_{\boldsymbol{i}}, \tag{14}$$

hence

$$\sum_{\substack{\langle \boldsymbol{i},\boldsymbol{b}\rangle = 0 \\ \langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0}} n_{\boldsymbol{i}} \geq \lceil t_{1+\nu(\boldsymbol{h}')}/q \rceil \tag{15}$$

holds for all $\boldsymbol{h}' \in \mathcal{P}_{k-1}$.

Now we set up a 1-1 correspondence between vectors $\boldsymbol{i}' \in \mathbb{F}_q^{k-1}$ and vectors $\boldsymbol{i} \in \mathbb{F}_q^k$ for which $\langle \boldsymbol{i}, \boldsymbol{b}\rangle = 0$. Writing $\boldsymbol{b} = (1, \boldsymbol{b}')$, we let $\boldsymbol{i}'$ correspond with $\boldsymbol{i} = (i_1, \boldsymbol{i}')$, with $i_1 := -\langle \boldsymbol{i}', \boldsymbol{b}'\rangle$. Note that then

$$\langle \boldsymbol{i}, \boldsymbol{h}\rangle = \langle \boldsymbol{i}', \boldsymbol{h}'\rangle. \tag{16}$$

Furthermore, if $\boldsymbol{i} \in \mathbb{F}_q^k \backslash \{\boldsymbol{0}\}$ with $\langle \boldsymbol{i}, \boldsymbol{b}\rangle = 0$ corresponds with $\boldsymbol{i}'$ in $\mathbb{F}_q^{k-1} \backslash \{\boldsymbol{0}\}$, then we define $n'_{\boldsymbol{i}'} := n_{\boldsymbol{i}}$. Finally, we set $T' := (t'_1, \ldots, t'_{k-1})$, where $t'_j := \lceil t_{j+1}/q \rceil$. Then from (15) and (16), we find that

$$\sum_{\langle \boldsymbol{i}',\boldsymbol{h}'\rangle \neq 0} n'_{\boldsymbol{i}'} \geq t'_{\nu(\boldsymbol{h}')} \tag{17}$$

holds for all $\boldsymbol{h}' \in \mathcal{P}_{k-1}$. Hence by induction, we have that

$$\sum_{\langle \boldsymbol{i},\boldsymbol{b}\rangle=0} n_{\boldsymbol{i}} = \sum_{\boldsymbol{i}' \in \mathbb{F}_q^{k-1}\setminus\{\boldsymbol{0}\}} n'_{\boldsymbol{i}'} = n' \geq \sum_{j=1}^{k-1} \lceil \lceil t_{j+1}/q \rceil / q^{j-1} \rceil = \sum_{j=2}^{k} \lceil t_j/q^{j-1} \rceil. \quad (18)$$

By combining (10) and (18), we now find that $n \geq \sum_{j=1}^{k} \left\lceil \frac{t_j}{q^{j-1}} \right\rceil$. So the claim for dimension k follows.

9 The Griesmer bound for linear codes and for linear UEP-codes from the ILP problem

The Griesmer bound for linear codes can also be proved by our ILP argument. To see that, assume that $\boldsymbol{G} = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n]$ is a $k \times n$ matrix over $\mathbb{F}_q$ that generates a k-dimensional q-ary linear code of length n with minimum distance d. Suppose that $\boldsymbol{G}$ has $n_{\boldsymbol{i}}$ columns equal to $\boldsymbol{i}$ ($\boldsymbol{i} \in \mathbb{F}_q^k$). Let $\boldsymbol{h}^{\perp}$ be a hyperplane, where $\boldsymbol{h} \in \mathbb{F}_q^k \setminus \{\boldsymbol{0}\}$. Consider $\boldsymbol{c}^{\top} = \boldsymbol{h}^{\top}\boldsymbol{G}$. Then $c_j = 0$ if and only if $\boldsymbol{h}^{\top}\boldsymbol{g}_j = 0$, so $w(\boldsymbol{c}) = \sum_{\langle \boldsymbol{h},\boldsymbol{i}\rangle \neq 0} n_{\boldsymbol{i}}$. We conclude that

$$\sum_{\langle \boldsymbol{h},\boldsymbol{i}\rangle \neq 0} n_{\boldsymbol{i}} \geq d$$

holds for every $\boldsymbol{h} \in \mathbb{F}_q^k \setminus \{\boldsymbol{0}\}$. So a linear code with generator matrix $\boldsymbol{G}$ has minimum distance at least d if and only if every hyperplane contains at most $n - d$ columns of $\boldsymbol{G}$, or equivalently, if there are at least d columns of $\boldsymbol{G}$ outside every hyperplane. This establishes the ILP (5) for the case where $t_i = d$ for all $i \in [k]$, hence shows that the Griesmer bound holds for linear codes.

The Griesmer bound for UEP codes (see [4, page 23]) can also be obtained from the ILP (5). Indeed, suppose that the linear UEP code is generated by a $k \times n$ matrix $\boldsymbol{G}$ over $\mathbb{F}_q$. Then the separation vector $(s_1, \ldots, s_k)$ of the code is given by

$$s_j = s_j(\boldsymbol{G}) = \min\{w(\boldsymbol{h}^{\top}\boldsymbol{G}) \mid h_j \neq 0\},$$

($j \in [k]$) [4]. Suppose that the rows of $\boldsymbol{G}$ are ordered in such a way that $s_1 \geq \cdots \geq s_k$. Suppose furthermore that $\boldsymbol{G}$ has $n_{\boldsymbol{i}}$ columns equal to $\boldsymbol{i}$ ($\boldsymbol{i} \in \mathbb{F}_q^k$). Then if $\boldsymbol{h} \in \mathbb{F}_q^k$ and $h_j = 1$ and $h_1 = \cdots = h_{j-1} = 0$ (so if $\boldsymbol{h} \in \mathcal{P}_k$ and $\nu(\boldsymbol{h}) = j$), then

$$\sum_{\langle \boldsymbol{i},\boldsymbol{h}\rangle \neq 0} n_{\boldsymbol{i}} = |\{l \mid \boldsymbol{h}^{\top}\boldsymbol{g}_l \neq 0\}| = w(\boldsymbol{h}^{\top}\boldsymbol{G}) \geq s_j = s_{\nu(\boldsymbol{h})},$$

so again we obtain the ILP (5); hence the Griesmer bound for UEP-codes also follows from the ILP bound. Conversely, it is not difficult to see that if $(n_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathbb{F}_q \setminus \{\boldsymbol{0}\}}$ is a feasible solution to the ILP (5), then with $n = \sum n_{\boldsymbol{i}}$, the $k \times n$ matrix G that has $n_{\boldsymbol{i}}$ columns equal to $\boldsymbol{i}$ for all $\boldsymbol{i} \in \mathbb{F}_q^n \setminus \{\boldsymbol{0}\}$ satisfies $s_j(\boldsymbol{G}) \geq t_j$ for all j. So the problem of finding a linear UEP code with the smallest length n for which $\boldsymbol{s} \geq (t_1, \ldots, t_k)$ is in fact equivalent to the ILP problem (5).

10 Open problems

1. It would be interesting to find a constructive proof of the Griesmer bound for UDD PIR codes, along the lines of the usual proofs for Griesmer-type bounds for linear codes. So the question is, given a linear q-ary $(t_1, \ldots, t_k)$-PIR code of length n with $t_1 \geq \cdots \geq t_k$, can we construct a $(\lceil t_2/q \rceil, \ldots, \lceil t_k/q \rceil)$-PIR code of length $n - t_1$?

2. Earlier we mentioned that linear UEP codes have an optimal generator matrix. This matrix can easily be constructed and thus the separation vector of the code can be determined relatively easily. Do PIR codes, and, more generally, UDD PIR codes, also have an optimal generator matrix? Does the optimal generator matrix for the code, considered as a UEP code, always provide the optimal encoder for the code as an UDD PIR code?

11 Conclusions

Unequal Error Protection (UEP) error-correcting codes were designed for the scenario where some parts of the encoded data need more protection than other parts. The correction properties of an encoder for an UEP code are captured by a generalization of the minimum distance called the *separation vector*. In this paper, we investigate Unequal Data Demand (UDD) PIR codes, generalizing the notion of a t-PIR code to include scenarios where some parts of the encoded data are in higher demand than other parts. First we have proved a generalized distance bound for UDD PIR encoders in terms of the separation vector of the associated UEP code. This bound has been used to derive a Griesmer-type bound for linear UDD PIR codes from the corresponding Griesmer bound for linear UEP codes. For an alternative proof of this Griesmer-type bound, we have derived an Integer Linear Programming (ILP) bound for the minimum length of a linear UDD PIR code, and we have determined a lower bound for the optimal solution of this ILP. In addition, we show that this ILP bound can be used to give a uniform proof for the Griesmer bound for linear codes, for linear UEP codes, and for linear UDD PIR codes.

Acknowledgments. This research was supported by the Estonian Research Council grants PRG49 and PSG114, and by the European Regional Development Fund via CoE project EXCITE.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. B. Chor, E. Kushilevitz and O. Goldreich, M. Sudan. Private information retrieval. in: Proc. 36-th IEEE Symp. on Foundations of Computer Science (FOCS), 41–50, 1995

2. L.A. Dunning, W. Robbins. Optimal encodings of linear block codes for unequal error protection. *Inform. Control*, 37:150–177, 1978
3. A. Fazeli, A. Vardy, E. Yaakobi. Codes for distributed PIR with low storage overhead. In: Proc. IEEE Symp. Information Theory (ISIT), 2852–2856, Hong Kong (2015)
4. W.J. van Gils. Design of error-control coding schemes for three problems of noisy information transmission, storage and processing. PhD thesis, Eindhoven University of Technology, 1988. `https://doi.org/10.6100/IR274904`
5. H.D.L. Hollmann, U. Luhaäär. Optimal possibly nonlinear 3-PIR codes of small size. in: Arithmetic of Finite Fields, Proceedings WAIFI 2022, LNCS 13638, Chapter 9
6. J.H. van Lint. Introduction to Coding Theory (3ed). Springer, 1999
7. H.-Y. Lin, E. Rosnes. Lengthening and extending binary private information retrieval codes. In: Proc. International Zurich Seminar on Information and Communication (IZS), 113 –117, ETH Zurich. February 21–23 (2018)
8. F.J. MacWilliams, N.J.A. Sloane. The Theory of Error-Correcting Codes. North Holland, 1977
9. H. Lipmaa, V. Skachek. Linear batch codes. In: Proc. 4th International Castle Meeting on Coding Theory and Applications (ICMCTA), 245–253, Palmela, Portugal (2014)
10. S. Kurz, E. Yaakobi. PIR codes with short block length. *Des. Codes, Cryptogr.*, 89:559–587, June 2021
11. M. Puškin, On Unequal Data Demand Private Information Retrieval Codes. Bachelor Thesis, University of Tartu, 2022
12. V. Skachek. Batch and PIR codes and their connections to locally repairable codes. In: Greferath, M., Pavčević, M. O., Silberstein, N., Ángeles Vázquez-Castro, M. (eds) Network Coding and Subspace Designs, 427–442. Springer (2018)
13. A. Vardy. Private Information Retrieval: Coding instead of Replication. Talk at the Institate Henri Poincaré, March 25 (2016). Available at `https://www.youtube.com/watch?v=WU2-6Da8IyE&t=934s`
14. J. Zumbrägel, V. Skachek. Talk: On bounds for batch codes. Algebraic Combinatorics and Applications (ALCOMA), March 15–20 (2015)

Understanding the new distinguisher of alternant codes at degree 2

Axel Lemoine[1], Rocco Mora[2], and Jean-Pierre Tillich[1]

[1] Inria Paris, France
{axel.lemoine,jean-pierre.tillich}@inria.fr
[2] CISPA, Germany
rocco.mora@cispa.de

Abstract. Distinguishing Goppa codes or alternant codes from generic linear codes [FGO$^+$11] has been shown to be a first step before being able to attack the McEliece type cryptosystem based on those codes [BMT23]. Whereas the distinguisher of [FGO$^+$11] is only able to distinguish Goppa codes or alternant codes of rate very close to 1, in [CMT23a] a much more powerful (and more general) distinguisher is proposed. It is based on computing the Hilbert series $\{\mathrm{HF}(d),\ d \in \mathbb{N}\}$ of a Pfaffian modeling. The distinguisher of [FGO$^+$11] can be interpreted as computing HF(1). Computing HF(2) still gives a polynomial time distinguisher for alternant or Goppa codes and is apparently able to distinguish Goppa or alternant codes in a much broader regime of rates as the one of [FGO$^+$11]. However, the scope of this distinguisher was unclear. We give here a formula for HF(2) corresponding to generic alternant codes when the field size q satisfies $q \geqslant r$ where r is the degree of the alternant code. We also show that this expression for HF(2) is a lower bound in general on it. The HF(2) corresponding to random linear codes is known and this yields a precise description of the new regime of rates that can be distinguished by this new method which shows that the new distinguisher improves significantly upon the one given in [FGO$^+$11].

1 Introduction

The McEliece cryptosystem [McE78] is the oldest code-based scheme and it is based on binary Goppa codes, a subfamily of alternant codes. It is believed to be quantum-resistant and its IND-CCA secure variation [ABC$^+$22] is currently a fourth round finalist of the NIST post-quantum competition. For a long time, it was believed that structural attacks aiming at recovering the underlying Goppa structure from an arbitrary generator matrix of the code were much more expensive than message recovery attacks. The latter ignore completely the algebraic structure and aim just at decoding a generic linear code.

In [FGO$^+$11] another approach was tried. Instead of trying to recover directly the algebraic structure from a generator matrix of a Goppa code, a potentially easier problem is solved first, namely that of *distinguishing* a Goppa code from a generic linear code just by the knowledge of a generator matrix of the code. This

is a promise problem where either we are given a generator matrix of a Goppa code or one of a random linear code and one must decide in which case we are. It turned out that there is a way to solve this problem in polynomial time for Goppa codes and more generally for the slightly more general family of alternant codes, at least for very high rate codes [FGO$^+$11]. It took a while to transform this distinguisher into an algorithm recovering the algebraic structure of the Goppa or the alternant code, but this has recently been achieved in [BMT23, CMT23b] (but binary Goppa codes could not be handled by these papers).

Interestingly enough, [CMT23b] also puts forward a new algebraic object, namely the matrix code of quadratic relations. The point is that this matrix code can be associated to any linear code. However, the matrix code associated to Goppa or alternant codes contains matrices of unusually low rank, namely rank 3 in characteristic $\neq 2$ and rank 2 in characteristic 2, which are consequences of structured quadratic relations. Finding such low rank matrices can in principle be achieved by solving the corresponding MinRank problem. Moreover, in characteristic 2, the matrix code is a subspace of skew-symmetric matrices and the MinRank problem can be modeled with a system where the Pfaffians of principal submatrices of order 4 are equated to 0. The polynomials corresponding to these equations define what we call the Pfaffian ideal. The existence of low-rank matrices has been exploited to mount a distinguisher attack and its complexity has been partially analyzed [CMT23b] as we recall below.

This work focuses on characteristic 2 and aims to advance the knowledge of a fundamental object associated with the above-mentioned Pfaffian ideal (and with polynomial ideals in general): its Hilbert function (or series). This Hilbert series $\{\mathrm{HF}(d), d \in \mathbb{N}\}$ turns out to be a very good way to distinguish alternant or Goppa codes from generic linear codes. Whereas $\mathrm{HF}(d)$ never vanishes in the first case, it turns out to be equal to 0 for a large enough degree in the second case. This gives a new distinguisher for Goppa or alternant codes. The Hilbert function associated to a generic linear code can be easily derived by making some assumptions that have been verified experimentally [CMT23b, Conjecture 1] and the smallest degree for which the Hilbert series vanishes can be computed. Interestingly, when the co-dimension $n-k$ of the code is of the form $n-k = \mathcal{O}\left(n^{\alpha}\right)$ when $\alpha < 1$ and n is the codelength, the degree d at which this happens is low enough so that the actual computation of the Hilbert series can be done with a complexity which is smaller than the aforementioned message recovery attacks. Potentially, this also paves the way to key attacks on the McEliece cryptosystem based on such codes of very large rate which are *less complex* than message recovery attacks.

Unfortunately, whereas the Hilbert series $\{\mathrm{HF}_{\mathrm{R}}(d), d \in \mathbb{N}\}$ of a generic linear code is well understood in [CMT23b], the Hilbert series $\{\mathrm{HF}_{\mathrm{A}}(d), d \in \mathbb{N}\}$ that corresponds to an alternant code is much less understood. This is a pity, since this would allow to understand precisely the scope of the distinguisher based on the computation of the Hilbert series. The only case, which was understood right now is the Hilbert series at degree 1, $\mathrm{HF}(1)$. It turns out that knowing $\mathrm{HF}(1)$ is equivalent to knowing the dimension of the square of the dual code and the

distinguisher of alternant or Goppa codes based on the fact that their HF(1) differs is actually equivalent to the distinguisher of [FGO+11].

The aim of this work is to understand the value of $\mathrm{HF}_{\mathrm{A}}(2)$. We will provide here a formula for it together with a proof using a natural conjecture that has been verified experimentally. We also prove that this formula is actually a rigorous lower bound on $\mathrm{HF}_{\mathrm{A}}(2)$ in general. It turns out that the distinguisher based on $\mathrm{HF}_{\mathrm{A}}(2) \neq \mathrm{HF}_{\mathrm{R}}(2)$ works for a much broader set of of parameters than the distinguisher $\mathrm{HF}_{\mathrm{A}}(1) \neq \mathrm{HF}_{\mathrm{R}}(1)$ (which is equivalent to the one of [FGO+11]). This could open the way to key attacks in the regime of parameters for which $\mathrm{HF}_{\mathrm{A}}(2) \neq \mathrm{HF}_{\mathrm{R}}(2)$, much in the same way that [FGO+11, MT23] were a first step before the attacks of [BMT23, CMT23b]. On top of that, knowing the Hilbert series precisely is crucial when it comes to solve the Pfaffian system and our work can be viewed as a significant step in this direction.
Note: The proofs of all results announced in this extended abstract can be found in the full version of this paper which can be found on eprint.

2 Preliminaries

2.1 Reed-Solomon and alternant codes

We work in characteristic 2 throughout the paper. We denote by $\mathbb{F}_q$ the finite field of size q which is therefore assumed here to be a power of 2.

Definition 1 (Generalized Reed-Solomon code). *Let $n \leqslant q$ be an integer, $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a vector of pairwise-distinct elements of $\mathbb{F}_q$, and $\boldsymbol{y} \in (\mathbb{F}_q^\times)^n$. The generalized Reed-Solomon code of dimension r, support $\boldsymbol{x}$ and multiplier $\boldsymbol{y}$ is*

$$\mathbf{GRS}_r(\boldsymbol{x}, \boldsymbol{y}) \stackrel{def}{=} \{(y_1 f(x_1), \ldots, y_n f(x_n) \mid f \in \mathbb{F}_q[X]_{<r})\}.$$

A (generalized) Reed-Solomon code of degree r is an MDS code of dimension r. It is well-known that the dual of a GRS code is also a GRS code [MS77].

Definition 2 (Alternant code). *Let r, m be two integers, $\boldsymbol{x} \in \mathbb{F}_{q^m}^n$ be a support, and $\boldsymbol{y} \in (\mathbb{F}_{q^m}^\times)^n$ be a multiplier. The alternant code of support $\boldsymbol{x}$ and multiplier $\boldsymbol{y}$ is*

$$\mathscr{A}_r(\boldsymbol{x}, \boldsymbol{y}) \stackrel{def}{=} (\mathbf{GRS}_r(\boldsymbol{x}, \boldsymbol{y})^\perp)_{|\mathbb{F}_q}.$$

We know that [MS77] $\dim_{\mathbb{F}_q} \mathscr{A}_r(\boldsymbol{x}, \boldsymbol{y}) \geqslant n - rm$ and this bound is generally tight.

2.2 Quadratic relations over a basis of a code

We first recall the concept of *code of quadratic relations* with respect to a basis of a linear code.

Definition 3. *Let $\mathscr{C}$ be an $[n, k]$-linear code over $\mathbb{F}_q$, and let $\mathcal{V} \stackrel{def}{=} (v_1, \ldots, v_k)$ be a basis of $\mathscr{C}$. The code of quadratic relations of $\mathscr{C}$ with respect to $\mathcal{V}$ is defined as*

$$\mathscr{C}_{rel}(\mathcal{V}) \stackrel{def}{=} \left\{ \boldsymbol{c} = (c_{i,j})_{1 \leqslant i \leqslant j \leqslant n} \in \mathbb{F}_q^{\binom{k+1}{2}} \;\middle|\; \sum_{i \leqslant j} c_{i,j} v_i \star v_j = 0 \right\},$$

where $\boldsymbol{a} \star \boldsymbol{b} \stackrel{def}{=} (a_1 b_1, \ldots, a_n b_n)$ *for* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{F}_q^n$.

As usual, a codeword $\boldsymbol{c} \in \mathscr{C}_{rel}(\mathcal{V})$ may be seen as a quadratic form whose matrix $M_{\boldsymbol{c}} = (m_{i,j})_{1 \leqslant i,j \leqslant k}$ is defined by $m_{i,j} = c_{i,j}$ if $i < j$, $m_{i,j} = c_{j,i}$ if $i > j$ and $m_{i,i} = 2c_{i,i}$. In characteristic 2, the diagonal of these matrices is thus always zero. We define the code of matrices

$$\mathscr{C}_{mat}(\mathcal{V}) \stackrel{\text{def}}{=} \{M_{\boldsymbol{c}} \mid \boldsymbol{c} \in \mathscr{C}_{rel}(\mathcal{V})\}.$$

We recall that a lot of interesting features of the code of relations remain invariant under a change of basis.

Lemma 1 ([CMT23b], Proposition 4).

$$\dim \mathscr{C}_{mat}(\mathcal{V}) = \dim \mathscr{C}_{rel}(\mathcal{V}).$$

$\dim \mathscr{C}_{rel}(\mathcal{V})$ *and the rank distribution of* $\mathscr{C}_{mat}(\mathcal{V})$ *is invariant under a change of basis.*

As a consequence, we sometimes write $\mathscr{C}_{rel}$ or $\mathscr{C}_{mat}$ without specifying the basis when we refer to invariants.

3 Codes of relations of a generalized Reed-Solomon code

The key for understanding $\mathrm{HF}_{\mathrm{A}}(2)$ will be to treat the case $m = 1$ first, *i.e.* when the alternant code is actually a generalized Reed-Solomon code.

3.1 Fundamental relations in the canonical basis

Definition 4 (Canonical basis). $\mathcal{A} \stackrel{def}{=} (\boldsymbol{a}_0, \ldots, \boldsymbol{a}_{r-1})$, *where* $\boldsymbol{a}_i = \boldsymbol{x}^i \star \boldsymbol{y}$ *forms a basis of* $\mathbf{GRS}_r(\boldsymbol{x}, \boldsymbol{y})$, *which we call a canonical basis*[3].

Definition 5 (Fundamental relations). *Quadratic relations of the form*

$$\boldsymbol{a}_i \star \boldsymbol{a}_j = \boldsymbol{a}_k \star \boldsymbol{a}_l$$

when $i + j = k + l$ *will be called* ***fundamental relations****.*

It is well known that there are $\binom{r-1}{2}$ linearly independent fundamental relations in $\mathscr{C}_{rel}(\mathcal{A})$, and in some cases these relations generate the code of relations.

Proposition 1. *If* $2r - 1 \leqslant n$, *then the fundamental relations form a basis of* $\mathscr{C}_{rel}(\mathcal{A})$.

[3] Since different vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ may generate the same **GRS** code, we talk about *a* canonical basis rather than *the* canonical basis.

When the above proposition holds, not only $\mathscr{C}_{mat}(\mathcal{A})$ *does not* depend on the choice of $\boldsymbol{x}$ and $\boldsymbol{y}$, but *every* GRS code of dimension r has the same code of relations with respect to any canonical basis. We can even describe a simple algorithm that builds a basis of $\mathscr{C}_{mat}(\mathcal{A})$. The idea is to list the relations in the following order :

$$\boldsymbol{a}_0 \star \boldsymbol{a}_2 = \boldsymbol{a}_1 \star \boldsymbol{a}_1, \boldsymbol{a}_0 \star \boldsymbol{a}_3 = \boldsymbol{a}_1 \star \boldsymbol{a}_2, \boldsymbol{a}_0 \star \boldsymbol{a}_4 = \boldsymbol{a}_2 \star \boldsymbol{a}_2,\ \boldsymbol{a}_1 \star \boldsymbol{a}_3 = \boldsymbol{a}_2 \star \boldsymbol{a}_2, \ldots$$

Before describing the algorithm, let us introduce a few notations.

Definition 6 (Sparse notation). *Let $I = \{(i_1, j_1), \ldots, (i_t, j_t)\} \subseteq [\![1, r]\!]^2$ be a list of distinct tuples such that $i_s < j_s$ for all s. We define $[(i_1, j_1), \ldots, (i_t, j_t)]$ as the matrix $(m_{k,l})_{1 \leqslant k,l \leqslant r}$ such that $m_{k,l} = 1$ if $(k, l) \in I$ or $(l, k) \in I$ and 0 otherwise.*

Finally, if $M \in \mathbb{F}_q^{r \times r}$, $\check{M}$ denotes the *transpose* of M with respect to the anti-diagonal $(i + j = r + 1)$, *i.e* $\check{m}_{i,j} = m_{r-j+1,r-i+1}$. Using this notation, Algorithm 1 returns a basis of $\mathscr{C}_{mat}(\mathcal{A})$. One can check that the algorithm produces exactly

Algorithm 1 Generation of a basis $\mathcal{B} = (M_1, \ldots, M_N)$ of $\mathscr{C}_{mat}(\mathcal{A})$

```
B ← ∅
s ← 4
while s ⩽ r + 1 do
    (i, j) ← (1, s − 1)          ▷ (i, j) will run along the sub-anti-diagonal i + j = s.
    if s is even then
        while i < j do
            M ← [(i, j)]
            B ← B ∪ {M, M̌}
            (i, j) ← (i + 1, j − 1)
    else
        k ← (s−1)/2
        while i + 1 < j − 1 do
            M ← [(i, j), (k − 1, k)]
            B ← B ∪ {M, M̌}
            (i, j) ← (i + 1, j − 1)
    s ← s + 1
return B
```

$N = \binom{r-1}{2}$ matrices (note that when $s = r + 1$, the set $\{M, \check{M}\}$ contains only one matrix). Furthermore, for each matrix $M \in \mathcal{B}$ returned by the algorithm, its nonzero coefficients that are not right above/under the diagonal do not appear in any other matrix in $\mathcal{B}$, which clearly implies that the elements of $\mathcal{B}$ are linearly independent over $\mathbb{F}_q$. Applying Proposition 1, we get

Corollary 1. *When $2r - 1 \leqslant n$, Algorithm 1 returns a basis of $\mathscr{C}_{mat}(\mathcal{A})$. This basis is referred to as* ***the canonical basis***[4] *of $\mathscr{C}_{mat}$.*

[4] This basis does not depend on the choice $\boldsymbol{x}$ and $\boldsymbol{y}$, but depends on r.

3.2 Rank 2 matrices in $\mathscr{C}_{mat}$

Among the fundamental relations, the ones of the form

$$\boldsymbol{a}_i \star \boldsymbol{a}_j = \boldsymbol{a}_k^{\star 2}$$

(when $i + j = 2k$) give a matrix of rank 2 when $\mathbb{F}$ is of characteristic 2. This suggests that there are many matrices of rank 2 in $\mathscr{C}_{mat}$.

Implicit modeling of [CMT23b]. To find rank 2 matrices in $\mathscr{C}_{mat}$, we may adopt the inverse point of view, *i.e* finding matrices belonging to $\mathscr{C}_{mat}$ inside the variety of skew-symmetric matrices of rank $\leqslant 2$. Writing the generic skew-symmetric matrix

$$\boldsymbol{M} = \begin{pmatrix} 0 & X_{1,2} & \dots & X_{1,r} \\ X_{1,2} & 0 & \dots & X_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ X_{1,r} & X_{2,r} & \dots & 0 \end{pmatrix},$$

we know that the variety of rank $\leqslant 2$ skew-symmetric matricies in characteristic 2 may be described by the following equations [Wim12]:

$$X_{i,j}X_{k,l} + X_{i,k}X_{j,l} + X_{i,l}X_{j,k} = 0,\ 1 \leqslant i < j < k < l \leqslant r \tag{1}$$

We call the left-hand side of (1) a *Pfaffian* of $\boldsymbol{M}$, as it is *the* Pfaffian of some 4×4 principal submatrix of $\boldsymbol{M}$. More generally, if $\boldsymbol{N}$ is any skew-symmetric matrix whose coefficients lie in some polynomial ring, we denote by $\mathrm{Pf}(\boldsymbol{N}, 2)$ the set of all Pfaffians of size 4 of $\boldsymbol{N}$. Adding linear equations expressing the fact that $\boldsymbol{M}$ belongs to $\mathscr{C}_{mat}$, we obtain the first algebraic modeling of rank $\leqslant 2$ matrices in $\mathscr{C}_{mat}$:

Modeling 1 (Implicit modeling) *The implicit modeling of rank $\leqslant 2$ matrices in $\mathscr{C}_{mat}$ consists of the ideal I generated by the $\binom{r}{4}$ Pfaffians of the generic skew-symmetric $r \times r$ matrix $\boldsymbol{M}$ and parity-check equations expressing the fact that $\boldsymbol{M}$ belongs to $\mathscr{C}_{mat}$.*

Explicit modeling. Another strategy is to compute a basis $(\boldsymbol{M}_1, \dots, \boldsymbol{M}_N)$ of $\mathscr{C}_{mat}$ and solve the MinRank problem with matrix

$$\boldsymbol{M} \stackrel{\text{def}}{=} \sum_{i=1}^{N} X_i \boldsymbol{M}_i \tag{2}$$

$\boldsymbol{M}$ looks like this in the canonical basis when $r = 5$ or $r = 6$

$$\underset{(r=5)}{\boldsymbol{M}} = \begin{pmatrix} 0 & 0 & X_1 & X_2 & X_4 \\ 0 & 0 & X_2 & X_3 & X_5 \\ X_1 & X_2 & 0 & X_5 & X_6 \\ X_2 & X_3 & X_5 & 0 & 0 \\ X_4 & X_5 & X_6 & 0 & 0 \end{pmatrix}, \underset{(r=6)}{\boldsymbol{M}} = \begin{pmatrix} 0 & 0 & X_1 & X_2 & X_4 & X_6 \\ 0 & 0 & X_2 & X_3 & X_5 & X_8 \\ X_1 & X_2 & 0 & X_5 + X_6 & X_7 & X_9 \\ X_2 & X_3 & X_5 + X_6 & 0 & X_9 & X_{10} \\ X_4 & X_5 & X_7 & X_9 & 0 & 0 \\ X_6 & X_8 & X_9 & X_{10} & 0 & 0 \end{pmatrix}$$

The matrix $\boldsymbol{M}$ is the generic matrix in $\mathscr{C}_{mat}$. Since it is skew-symmetric, one may consider its Pfaffians of degree 2, *i.e* the Pfaffians of all 4×4 submatrices of $\boldsymbol{M}$, which leads to the following algebraic modeling.

Modeling 2 (Explicit Pfaffian modeling) *The explicit modeling consists of* $\binom{r}{4}$ *equations* $f = 0$ *for* $f \in \mathbf{Pf}(\boldsymbol{M}, 2)$. *More explicitely, writing* $\boldsymbol{M} = (\boldsymbol{m}_{i,j})_{1 \leqslant i,j \leqslant r}$, *the equations are*

$$\boldsymbol{m}_{i,j}\boldsymbol{m}_{k,l} + \boldsymbol{m}_{i,k}\boldsymbol{m}_{j,l} + \boldsymbol{m}_{i,l}\boldsymbol{m}_{j,k} = 0.$$

where each coefficient $\boldsymbol{m}_{i,j}$ *is a polynomial of degree 1.*

We are interested in computing the Hilbert function at degree 2 of the ideal generated by $\mathbf{Pf}(\boldsymbol{M}, 2)$. We recall the concept of Hilbert function.

Definition 7 (Hilbert function). *Let* I *be a homogeneous ideal of a polynomial ring* $\mathbb{F}[\boldsymbol{X}]$. *Writing* $\mathbb{F}[\boldsymbol{X}]_d$ *the (finite dimensional)* $\mathbb{F}$*-vector space spanned by monomials of degree* d *and* $I_d = I \cap \mathbb{F}[\boldsymbol{X}]_d$, *the Hilbert function of* I *is defined as*

$$\mathrm{HF}_{\mathbb{F}[\boldsymbol{X}]/I}(d) \stackrel{def}{=} \dim_{\mathbb{F}} \mathbb{F}[\boldsymbol{X}]_d/I_d, \ d \in \mathbb{N}.$$

Experimentally, we always find that the elements of $\mathbf{Pf}(\boldsymbol{M}, 2)$ are linearly independent, which leads us to state the following as a conjecture.

Conjecture 1. $\mathrm{HF}(2) = \binom{\binom{r-1}{2}+1}{2} - \binom{r}{4} = \dfrac{1}{12}(r-1)(r-2)(r^2-3r+6)$.

Note that it might be useful to consider the implicit Pfaffian modeling. However, since the Hilbert function strongly depends on how the equations are written, one must be careful when changing the modeling. In our case, we can safely do so thanks to the following theorem.

Theorem 1. *Let* I *(resp.* J*) be the ideal of the polynomial ring* $\boldsymbol{A}$ *(resp.* $\boldsymbol{B}$*) produced by the implicit (resp. explicit) modeling.* $\boldsymbol{A}/I$ *and* $\boldsymbol{B}/J$ *both have a structure of graded* $\mathbb{F}$*-algebra. There exists a map*

$$\Phi : \boldsymbol{A}/I \longrightarrow \boldsymbol{B}/J$$

that defines an isomorphism of graded $\mathbb{F}$*-algebras.*

In the following, we sometimes talk about *the Pfaffian modeling associated to a code* without specifying whether it is implicit or explicit, since we only deal with Hibert functions.

4 Hilbert function of a Pfaffian ideal associated with a generic alternant code

4.1 The block-diagonal code of relations

In the case of alternant codes, the crux for having rank 2 matrices in $\mathscr{C}_{mat}$ is to consider [CMT23b] the extension to $\mathbb{F}_{q^m}$ of the *dual* code. Let us then recall the following fact.

Proposition 2 ([BMT23], Proposition 14). *For any code $\mathscr{C} \subseteq \mathbb{F}_q^n \subseteq \mathbb{F}_{q^m}^n$, we denote by $\mathscr{C}_{\mathbb{F}_{q^m}}$ the $\mathbb{F}_{q^m}$-vector space spanned by $\mathscr{C}$. Let $\mathscr{C} = \mathscr{A}_r(\boldsymbol{x}, \boldsymbol{y})$ be an alternant code of extension degree m. Then*

$$(\mathscr{A}_r(\boldsymbol{x}, \boldsymbol{y})^\perp)_{\mathbb{F}_{q^m}} = \sum_{j=0}^{m-1} \mathbf{GRS}_r(\boldsymbol{x}^{q^j}, \boldsymbol{y}^{q^j})$$

*With the usual assumption that $\dim_{\mathbb{F}_q} \mathscr{A}_r(\boldsymbol{x}, \boldsymbol{y}) = n - rm$, the above sum becomes a direct sum and the sequence $\mathcal{A} = (\boldsymbol{a}_0, \ldots, \boldsymbol{a}_{r-1}, \boldsymbol{a}_0^q, \ldots, \boldsymbol{a}_{r-1}^q, \ldots, \boldsymbol{a}_{r-1}^{q^{m-1}})$ is a basis of $(\mathscr{A}_r(\boldsymbol{x}, \boldsymbol{y})^\perp)_{\mathbb{F}_{q^m}}$, called the **canonical basis**.*

When $r < q+1$, it follows from the analysis of [FGO+11] that $\mathscr{C}_{rel}(\mathcal{A})$ is spanned by relations like

$$\boldsymbol{a}_a^{q^l} \star \boldsymbol{a}_b^{q^l} = \boldsymbol{a}_c^{q^l} \star \boldsymbol{a}_d^{q^l}$$

for $0 \leqslant l < m$ and $0 \leqslant a, b, c, d < r$ such that $a + b = c + d$. This implies that any matrix $A \in \mathscr{C}_{mat}(\mathcal{A})$ has a block-diagonal structure, *i.e.*

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_0 & \ldots & \boldsymbol{0}_r \\ \vdots & \ddots & \vdots \\ \boldsymbol{0}_r & \ldots & \boldsymbol{A}_{m-1} \end{pmatrix}$$

where $\boldsymbol{A}_j \in \mathscr{C}_{mat}(\boldsymbol{a}_0^{q^i}, \ldots, a_{r-1}^{q^j})$ is the matrix associated with some element of the code of quadratic relations of $\mathbf{GRS}_r(\boldsymbol{x}, \boldsymbol{y})^{q^j} = \mathbf{GRS}_r(\boldsymbol{x}^{q^j}, \boldsymbol{y}^{q^j})$ with respect to its canonical basis.

4.2 The Hilbert function at degree 2

The authors of [CMT23b] noticed that the Hilbert function of the Pfaffian modeling at degree 1 can be used as a distinguisher that boils down to the one presented in [FGO+11]. We recall that a generic alternant code is *square-distinguishable* if it is 1-distinguishable in the sense of [CMT23b]. The Hilbert function at degree 2 can also be used as a distinguisher which seems to work on a larger range of parameters. Our goal here is to find a formula for HF(2) when the code is 2-distinguishable, assuming $r < q + 1$.

Let $(B_1, \ldots, B_N)$ be the matrices returned by Algorithm 1. As a reminder, $N = \binom{r-1}{2}$. One can find rank $\leqslant 2$ matrices in $\mathscr{C}_{mat}(\mathcal{A})$ by solving the linear MinRank problem associated with the matrix

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{M}_0 & \ldots & \boldsymbol{0}_r \\ \vdots & \ddots & \vdots \\ \boldsymbol{0}_r & \ldots & \boldsymbol{M}_{m-1} \end{pmatrix} \in \mathbb{F}_{q^m}[X_i \mid 1 \leqslant i \leqslant mN]^{rm \times rm} \tag{3}$$

where $\boldsymbol{M}_j = \sum_{i=1}^{N} X_{Nj+i} B_i \in \mathbb{F}_{q^m}[X_i \mid Nj + 1 \leqslant i \leqslant (N+1)j]^{r \times r}$ is essentially the matrix associated with a generalized Reed-Solomon code as described in the previous section. The main result of this work is the following theorem.

Theorem 2. *Assume Conjecture 1 is true. Let $\mathscr{A}_r(\boldsymbol{x}, \boldsymbol{y})$ be a* ***generic*** *square-distinguishable alternant code with $r < q + 1$. Then the Hilbert function HF_A at degree 2 of the Pfaffian modeling associated with $\mathscr{A}_r(\boldsymbol{x}, \boldsymbol{y})$ is given by*

$$\mathrm{HF}_A(2) = \frac{m}{12}(r-1)(r-2)(r^2-3r+6).$$

Theorem 2 requires Conjecture 1 to hold, and is also limited to the case $r < q + 1$. Indeed, equality cannot be claimed in general, because there might exist alternant codes for which additional relations occur. Analogously to previous results on the Hilbert function at degree 1 [MT23], the value provided by Theorem 2 for degree 2 still represents a lower bound.

Corollary 2. *The Hilbert function of the Pfaffian modeling associated with an alternant code of order r and extension degree m satisfies*

$$\mathrm{HF}_A(2) \geqslant \frac{m}{12}(r-1)(r-2)(r^2-3r+6).$$

Corollary 2 allows us to state when a generic alternant code is 2-distinguishable. We use here the following definition of d-distinguishability

Definition 8. *An alternant code is d-distinguishable if the associated Hilbert function HF_A satisfies*

$$\mathrm{HF}_A(d) > \mathrm{HF}_R(d)$$

where HF_R is the Hilbert series of a random linear code of the same length and dimension as the alternant code.

Corollary 3. *If $\mathrm{HF}_R(2) < \frac{m}{12}(r-1)(r-2)(r^2-3r+6)$, then an alternant code $\mathscr{A}_r$ is 2-distinguishable.*

In the case $r < q+1$, we experimentally found that the Hilbert function at degree 2 for a generic 2-distinguishable alternant code was always equal to $\frac{m}{12}(r-1)(r-2)(r^2-3r+6)$.

5 The new distinguisher regime

From [CMT23a], it can be readily deduced that $\mathrm{HF}_{\mathrm{R}}(2)$ corresponding to a random linear code of the same dimension $k = n - rm$ as a generic alternant code of length n, degree r and extension degree m is given by

$$\mathrm{HF}_{\mathrm{R}}(2) = \max\left\{0, \frac{1}{2}\left(k^2 - k(s^2-s+1) + \frac{s^4-s^2}{6}\right)\right\}, \tag{4}$$

where $s \stackrel{\text{def}}{=} rm$. Combined with Corollary 3, this implies

Proposition 3. *For a given degree r and extension degree m and assuming that the field size q satisfies $q \geqslant r$, a generic alternant code whose dimension satisfies $k > k_0$ where*

$$k_0 \stackrel{def}{=} \frac{s^2 - s + 1 - \sqrt{\frac{s^4}{3} + \frac{2H}{3} - 2s^3 + \frac{11}{3}s^2 - 2s + 1}}{2}$$

with $s \stackrel{def}{=} rm$, $H \stackrel{def}{=} m(r-1)(r-2)(r^2-3r+6)$ is 2-distinguishable.

A natural asymptotic choice of parameters is to let r go to infinity and assume that $m = \mathcal{O}(\log r)$. This is in general the range which is chosen for m, since in order to maximize the decoding capacity one chooses the smallest possible m such that $q^m \geqslant n$. In such a case, it is straightforward to check that

$$k_0 \underset{r\to\infty}{\sim} \frac{1 - \sqrt{\frac{1+\frac{2}{m^3}}{3}}}{2} m^2 r^2.$$

When m also goes to infinity with r, we have

$$k_0 \sim \frac{1 - \sqrt{\frac{1}{3}}}{2} m^2 r^2 \approx 0.21 m^2 r^2.$$

This is much better than the distinguisher of [FGO+11]. In the regime where $q \geqslant r$, it is able to distinguish a generic alternant code from a generic linear code when $n > \binom{mr+1}{2} - \frac{m(r-1)(r-2)}{2}$, that is when $k > \binom{mr+1}{2} - rm - \frac{m(r-1)(r-2)}{2}$. This corresponds to $k > k_1 \stackrel{\text{def}}{=} \binom{mr}{2} - \frac{m(r-1)(r-2)}{2}$ with $k_1 = \frac{1-\frac{1}{m}}{2} m^2 r^2 + o\left(m^2 r^2\right)$ as $r \to \infty$ and if m goes to infinity as well, $k_1 \sim \frac{m^2 r^2}{2}$.

References

[ABC+22] Martin Albrecht, Daniel J. Bernstein, Tung Chou, Carlos Cid, Jan Gilcher, Tanja Lange, Varun Maram, Ingo von Maurich, Rafael Mizoczki, Ruben Niederhagen, Edoardo Persichetti, Kenneth Paterson, Christiane Peters, Peter Schwabe, Nicolas Sendrier, Jakub Szefer, Cen Jung Tjhai, Martin Tomlinson, and Wang Wen. Classic McEliece (merger of Classic McEliece and NTS-KEM). `https://classic.mceliece.org`, November 2022. Fourth round finalist of the NIST post-quantum cryptography call.

[BMT23] Magali Bardet, Rocco Mora, and Jean-Pierre Tillich. Polynomial time key-recovery attack on high rate random alternant codes. *CoRR*, abs/2304.14757, 2023. To appear in the Transactions on Information Theory.

[CMT23a] Alain Couvreur, Rocco Mora, and Jean-Pierre Tillich. A new approach based on quadratic forms to attack the McEliece cryptosystem. In Jian Guo and Ron Steinfeld, editors, *Advances in Cryptology - ASIACRYPT 2023 - 29th International Conference on the Theory and Application of Cryptology and Information Security, Guangzhou, China, December 4-8, 2023, Proceedings, Part IV*, volume 14441 of *LNCS*, pages 3–38. Springer, 2023.

[CMT23b] Alain Couvreur, Rocco Mora, and Jean-Pierre Tillich. A new approach based on quadratic forms to attack the McEliece cryptosystem. *arXiv preprint arXiv:2306.10294*, 2023.

[FGO+11] Jean-Charles Faugère, Valérie Gauthier, Ayoub Otmani, Ludovic Perret, and Jean-Pierre Tillich. A distinguisher for high rate McEliece cryptosystems. In *Proc. IEEE Inf. Theory Workshop- ITW 2011*, pages 282–286, Paraty, Brasil, October 2011.

[McE78] Robert J. McEliece. *A Public-Key System Based on Algebraic Coding Theory*, pages 114–116. Jet Propulsion Lab, 1978. DSN Progress Report 44.

[MS77] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error Correcting Codes*. North Holland, 1977.

[MT23] Rocco Mora and Jean-Pierre Tillich. On the dimension and structure of the square of the dual of a Goppa code. *Des. Codes Cryptogr.*, 91(4):1351–1372, 2023.

[Wim12] Michael Wimmer. Algorithm923: Efficient numerical computation of the Pfaffian for dense and banded skew-symmetric matrices. *ACM Trans. Math. Software*, 38(4), aug 2012.

On equidistant single-orbit cyclic subspace codes

Mahak and Maheshanand Bhaintwal

Indian Institute of Technology Roorkee, Roorkee-247667, India

Abstract. A code is said to be equidistant if the distance between any two distinct codewords is the same. In this paper, we have studied equidistant single-orbit cyclic subspace codes. If a subspace U has dimension one or $n-1$, or U is a cyclic shift of a field, then the orbit code generated by U is equidistant and is termed a trivial equidistant orbit code. Using the concept of cyclic difference sets, we have proved that only the trivial equidistant single-orbit cyclic subspace codes exist. Further, examples have been provided to illustrate the existence of non-trivial single-orbit quasi-cyclic subspace codes.

Keywords: Subspace codes · Cyclic orbit codes · Equidistant codes

1 Introduction

Subspace codes are used in random network coding to correct errors and erasures. A well-known paper [11] by Kötter and Kschischang sparked the main interest in subspace codes. Since 2008, researchers have been actively engaged in working on subspace codes. Of special interest among subspace codes is the class of cyclic subspace codes, introduced by Etzion and Vardy [5]. The algebraic structure of cyclic subspace codes and their efficient encoding and decoding algorithms motivate the study of these codes.

A cyclic subspace code C is a collection of $\mathbb{F}_q$-subspaces in $\mathbb{F}_{q^n}$ that is closed under the multiplication by the elements of $\mathbb{F}_{q^n}^*$, i.e., $\alpha U \in C$ for all $U \in C$ and $\alpha \in \mathbb{F}_{q^n}^*$. In [15], it is shown that the set $\{\alpha U \mid \alpha \in \mathbb{F}_{q^n}^*\}$ can be seen as the orbit of a subspace U under the action of the group $\mathbb{F}_{q^n}^*$ on the set of subspaces of $\mathbb{F}_{q^n}$. Thus, a cyclic subspace code is the union of the orbits of the subspaces contained in it. The code $\mathrm{Orb}(U) = \{\alpha U \mid \alpha \in \mathbb{F}_{q^n}^*\}$ is called a single-orbit cyclic subspace code. Quasi-cyclic subspace codes are a natural generalization of cyclic subspace codes and are studied in [8]. A subspace code C is called a quasi-cyclic subspace code if $\alpha U \in C$ for all $U \in C$ and $\alpha \in G$, where G is a multiplicative subgroup of $\mathbb{F}_{q^n}^*$ [12].

A code with the property that the distance between any two distinct codewords is same is called an equidistant code. Equidistant subspace codes have been explored by many researchers [1, 2, 4, 9]. The applications of equidistant subspace codes to distributed storage are discussed in [13]. In this paper, we focus on equidistant single-orbit cyclic subspace codes. A code $\mathrm{Orb}(U)$ is always equidistant if either $\dim(U)$ is one or $n-1$, or U is a cyclic shift of a subfield. We call such codes trivial equidistant orbit codes. The concept of cyclic difference sets has

been used in this paper to prove that there are only trivial equidistant single-orbit cyclic subspace codes. We have given some examples to prove the existence of single-orbit quasi-cyclic subspace codes.

2 Preliminaries

Let $q = p^h$, p a prime and h a positive integer. Let $\mathbb{F}_q$ be a finite field of size q, and let $\mathbb{F}_q^* := \mathbb{F}_q \backslash \{0\}$. Let $\mathbb{F}_{q^n}$ denotes the extension field of degree n of $\mathbb{F}_q$, and let $\mathbb{F}_{q^n}^* := \mathbb{F}_{q^n} \backslash \{0\}$. For $\beta \in \mathbb{F}_{q^n}^*$, $|\beta|$ denotes the order of β in multiplicative group $(\mathbb{F}_{q^n}^*, \times)$. For any subset $\{x_1, x_2, \ldots, x_r\}$ of $\mathbb{F}_q^n$ the subspace of $\mathbb{F}_q^n$ spanned by this set is denoted by $\langle x_1, x_2, \ldots, x_r \rangle_{\mathbb{F}_q}$. For any element $\alpha \in \mathbb{F}_{q^n}$, we set $\overline{\alpha} = \alpha \mathbb{F}_q = \{\alpha\lambda \mid \lambda \in \mathbb{F}_q\}$.

The set of all $\mathbb{F}_q$-subspaces of $\mathbb{F}_{q^n}$ is denoted by $\mathcal{P}_q(n)$ and is called the projective space of order n over $\mathbb{F}_q$. For $0 \leq k \leq n$, the set of all k-dimensional $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^n$ is denoted by $\mathcal{G}_q(n,k)$ and is called a Grassmanian. Clearly,

$$\mathcal{P}_q(n) = \bigcup_{k=0}^{n} \mathcal{G}_q(n,k) \ .$$

The size of $\mathcal{G}_q(n,k)$ is given by the q-binomial coefficient $\begin{bmatrix} n \\ k \end{bmatrix}_q$, i.e.,

$$|\mathcal{G}_q(n,k)| = \begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q^n-1)(q^n-q)\ldots(q^n-q^{k-1})}{(q^k-1)(q^k-q)\ldots(q^k-q^{k-1})} \ .$$

The projective space $\mathcal{P}_q(n)$ is a metric space with respect to the metric d_s defined by

$$d_s(U,V) = \dim(U) + \dim(V) - 2\dim(U \cap V)$$

for any $U, V \in \mathcal{P}_q(n)$. A subspace code C is a subset of $\mathcal{P}_q(n)$ containing atleast two elements with the metric d_s. The minimum distance of a subspace code C, denoted by $d_s(C)$, is defined by

$$d_s(C) = \min\{d_s(U,V) \mid U, V \in C,\ U \neq V\} \ .$$

The subspace code C is said to be a constant dimension subspace code if every element in C is of the same dimension, i.e., $C \subseteq \mathcal{G}_q(n,k)$ for some positive integer $k \leq n$. For a constant dimension subspace code $C \subseteq \mathcal{G}_q(n,k)$, we have

$$d_s(C) = 2k - 2\ \max\{\dim(U \cap V) \mid U, V \in C,\ U \neq V\} \ .$$

A constant dimension subspace code C is said to be an *equidistant subspace code* if for all $U, V \in C$ with $U \neq V$ we have $d_s(U,V) = d_s(C)$. An equidistant subspace code $C \subseteq \mathcal{G}_q(n,k)$ is said to be c-intersecting if $d_s(C) = 2(k-c)$, i.e., $\dim(U \cap V) = c$ for all $U, V \in C$ with $U \neq V$.

It is well known that the extension field $\mathbb{F}_q^n$ is a vector space of dimension n over $\mathbb{F}_q$, and both $\mathbb{F}_q^n$ and $\mathbb{F}_{q^n}$ are isomorphic as a vector space over $\mathbb{F}_q$. Due to

the rich algebraic structure of $\mathbb{F}_{q^n}$ than $\mathbb{F}_q^n$, we consider the subspaces of $\mathbb{F}_q^n$ in $\mathbb{F}_{q^n}$ in the study of cyclic subspace codes.

For a subspace $U \subseteq \mathbb{F}_{q^n}$ and $\alpha \in \mathbb{F}_{q^n}^*$, the cyclic shift of U with respect to α is defined as $\alpha U = \{\alpha u \mid u \in U\}$. Clearly $\alpha U \subseteq \mathbb{F}_{q^n}$. It is easy to see that αU is a vector space over $\mathbb{F}_q$ and its dimension is same as the dimension of U over $\mathbb{F}_q$. In fact, we can define a group action $\mathbb{F}_{q^n}^* \times \mathcal{P}_q(n) \to \mathcal{P}_q(n)$ of $\mathbb{F}_{q^n}^*$ on $\mathcal{P}_q(n)$ [15] as

$$(\alpha, U) \to \alpha U \ .$$

For any $\mathbb{F}_q$-subspace $U \subseteq \mathbb{F}_q^n$, the orbit of U, denoted by $\mathrm{Orb}(U)$, is defined by

$$\mathrm{Orb}(U) = \{\alpha U \mid \alpha \in \mathbb{F}_{q^n}^*\} \ .$$

The *stabilizer* of U, denoted by $\mathrm{Stab}(U)$, is defined by $\mathrm{Stab}(U) = \{\alpha \in \mathbb{F}_{q^n}^* \mid \alpha U = U\}$. Clearly, $\mathrm{Stab}(U)$ is a subgroup of $\mathbb{F}_{q^n}^*$, and since $aU = U$ for all $a \in \mathbb{F}_q^*$, we have $\mathbb{F}_q^* \subseteq \mathrm{Stab}(U)$. By [8, Lemma 3.3], $\mathrm{Stab}(U) \cup \{0\}$ is a subfield of $\mathbb{F}_{q^n}$, and U is a vector space over $\mathrm{Stab}(U) \cup \{0\}$. Thus $\mathrm{Stab}(U) \cup \{0\} = \mathbb{F}_{q^t}$ for some t which is a divisor of $\gcd(\dim_{\mathbb{F}_q}(U), n)$. For $t = n$, i.e., $U = \mathbb{F}_{q^n}$, we have $\mathrm{Stab}(U) = \mathbb{F}_{q^n}^*$. Thus in this case, $\mathrm{Orb}(U)$ contains only one element. So, we always consider $t < n$. Using the orbit-stabilizer theorem, for any subspace U of $\mathbb{F}_q^n$, we have

$$|\mathrm{Orb}(U)| = \frac{q^n - 1}{|\mathrm{Stab}(U)|} = \frac{q^n - 1}{q^t - 1} \ .$$

A subspace code C with the property that for any $\alpha \in \mathbb{F}_{q^n}^*$ and $U \in C$, $\alpha U \in C$, is said to be a cyclic subspace code. Thus, $\mathrm{Orb}(U)$ is a cyclic subspace code of constant dimension, and we call $\mathrm{Orb}(U)$ a *single-orbit cyclic subspace code* or simply an *orbit code*. In general, a cyclic subspace code is the union of the orbits of the subspaces contained in it.

If $\mathrm{Stab}(U) = \mathbb{F}_q^*$, i.e., $|\mathrm{Orb}(U)| = \frac{q^n-1}{q-1}$, then $\mathrm{Orb}(U)$ is called a *full-length orbit code* and we say that U generates a full-length orbit. Otherwise, $\mathrm{Orb}(U)$ is a degenerate orbit.

Let U be an $\mathbb{F}_q$-subspace of dimension k in $\mathbb{F}_{q^n}$. Then $\mathrm{Orb}(U) \subseteq \mathcal{G}_q(n, k)$. By the definition of subspace distance, for any $\alpha, \beta \in \mathbb{F}_{q^n}^*$, we have

$$\begin{aligned} d_s(\alpha U, \beta U) &= \dim(\alpha U) + \dim(\beta U) - 2\dim(\alpha U \cap \beta U) \\ &= 2k - 2\dim(\alpha U \cap \beta U) \ . \end{aligned}$$

As, $\dim(\alpha U \cap \beta U) = \dim(U \cap \alpha^{-1}\beta U)$, so we get

$$d_s(\alpha U, \beta U) = 2k - 2\dim(U \cap \alpha^{-1}\beta U) \ .$$

Therefore,

$$\begin{aligned} d_s(\mathrm{Orb}(U)) &= \min\{d_s(\alpha U, \beta U) \mid \alpha, \beta \in \mathbb{F}_{q^n}^*, \alpha U \neq \beta U\} \\ &= 2k - 2\ \max\{\dim(U \cap \gamma U) \mid \gamma \in \mathbb{F}_{q^n}^*, \gamma U \neq U\} \ . \end{aligned}$$

Now, for any subspace U in $\mathbb{F}_{q^n}$, $\mathrm{Orb}(U)$ is an equidistant subspace code if there exists some non-negative integer c such that $\dim(U \cap \alpha U) = c$ for all $\alpha \in \mathbb{F}_{q^n}^*$. If $\mathrm{Orb}(U)$ is an equidistant code, then we call it *equidistant single-orbit cyclic subspace code*.

For a subspace U of dimension one or $n-1$ in $\mathbb{F}_{q^n}$, $\mathrm{Orb}(U)$ is trivially an equidistant subspace code. It is known that for a subspace U of dimension k in $\mathbb{F}_{q^n}$, $d_s(\mathrm{Orb}U) = 2k$ if and only if $U = \beta\mathbb{F}_{q^k}$, for some $\beta \in \mathbb{F}_{q^n}^*$[7]. For such a subspace U, $\mathrm{Stab}(U) = \mathbb{F}_{q^k}^*$ and $\dim(U \cap \gamma U) = 0$ for $\gamma \in \mathbb{F}_{q^n} \backslash \mathbb{F}_{q^k}$. Thus $\mathrm{Orb}(U)$ is an equidistant code. Hence, if $\dim(U) = 1$ or $n-1$, or if U is a cyclic shift of a subfield of $\mathbb{F}_{q^n}$, then $\mathrm{Orb}(U)$ is an equidistant subspace code, and we call it as trivial equidistant single-orbit cyclic subspace code. For a subspace U of dimension k in $\mathbb{F}_{q^n}$ with $\mathrm{Stab}(U) = \mathbb{F}_{q^k}^*$, we have $U = \gamma\mathbb{F}_{q^k}$ for some $\gamma \in \mathbb{F}_{q^n}^*$. Therefore, from now onward, we will assume that $\dim(U) > 1$ and $\mathrm{Stab}(U) = \mathbb{F}_{q^t}^*$ where $t < \dim(U)$.

Definition 1. *[14] Suppose $(G, +)$ is a finite group of order v in which the identity element is denoted by "0". Let k and λ be positive integers such that $2 \leq k < v$. A (v, k, λ)-difference set in $(G, +)$ is a subset $D \subseteq G$ that satisfies the following properties:*

1. *$|D| = k$,*
2. *the multiset $[x - y : x, y \in D, x \neq y]$ contains every element in $G\backslash\{0\}$ exactly λ times.*

Note that

$$\lambda(v-1) = k(k-1) \tag{1}$$

if a (v, k, λ)-difference set exists.

Let D be a (v, k, λ)-difference set in a group $(G, +)$. For any $g \in G$, define

$$D + g = \{x + g : x \in D\} .$$

Any set $D + g$ is called a translate of D.

Lemma 1. *[16] Let G be a group of order v and $D \subseteq G$ is a (v, k, λ)-difference set. Then for any $g, g' \in G$, $g \neq g'$, we have $|(D + g) \cap (D + g')| = \lambda$.*

Definition 2. *[10] Let $(G, +)$ be a group of order nm and let $(N, +)$ be a subgroup of G of order n. Then a k-subset D of G is called a relative difference set with parameters n, m, k, λ_1 and λ_2 (relative to N) or briefly an $(n, m, k, \lambda_1, \lambda_2)$-RDS, provided that the list of differences $\{d_1 - d_2 : d_1, d_2 \in D, d_1 \neq d_2\}$ contain each element of N (excepting 0) precisely λ_1 times and each element of $G\backslash N$ exactly λ_2 times.*

Lemma 2. *[10] Let D be an $(n, m, k, \lambda_1, \lambda_2)$-RDS in G. Then*

$$k(k-1) = n(m-1)\lambda_2 + (n-1)\lambda_1 .$$

Proposition 1. *[6] Let $\alpha \in \mathbb{F}_{q^n}^*$ be a primitive element over $\mathbb{F}_q$. Consider the orbit of the k-dimensional subspace $U = \{0, \alpha^{i_1}, \ldots, \alpha^{i_{q^k-1}}\}, i_j \in Z_{q^n-1}$ for all $j = 1, \ldots, q^k - 1$ under the action of the Singer subgroup generated by C_{u_α}. If the indices i_j constitute a $(v = q^n - 1, k = q^k - 1, \lambda)$ difference set, where $\lambda \leq q^d - 1$, the orbit code so formed has minimum subspace distance $2(k - d)$.*

3 Equidistant cyclic subspace codes

Consider an extension field $\mathbb{F}_{q^n}$. Let α be a primitive element of $\mathbb{F}_{q^n}$. Then $\mathbb{F}_{q^n}^* = \{\alpha^i \mid i = 0, 1, \ldots, q^n - 2\}$. Now consider the group $\mathbb{Z}_{q^n-1} = \{0, 1, \ldots, q^n - 2\}$ under the operation addition modulo $q^n - 1$. Let $G = \{\alpha^0 = 1, \alpha^{j_1}, \alpha^{j_2}, \ldots, \alpha^{j_m}\}$ be a subgroup of the multiplicative group $\mathbb{F}_{q^n}^*$. Then $I = \{t \mid \alpha^t \in G\}$ is a subgroup in $\mathbb{Z}_{q^n-1}$. Similarly, for a subgroup in $(\mathbb{Z}_{q^n-1}, \oplus_{q^n-1})$ there is a subgroup in $(\mathbb{F}_{q^n}^*, \times)$.

Lemma 3. *Let a, b and c be two positive integers such that $a < b \leq c$. If $m^c - 1$ divides $(m^a - 1)(m^b - 1)$, then we must have $c = b$.*

Proof. As $(m^c - 1) \mid (m^a - 1)(m^b - 1)$, there exists a positive integer k such that

$$(m^a - 1)(m^b - 1) = k(m^c - 1) \ .$$

From this we get $m^{a+b} - m^a - m^b + 1 = km^c - k$. This gives $m^a(m^b - 1 - m^{b-a} - km^{c-a}) = -(k + 1)$, and hence m^a divides $(k + 1)$. Let $k + 1 = sm^a$ for some positive integer s. From this we get $k = sm^a - 1$, and thus $(m^a - 1)(m^b - 1) = (sm^a - 1)(m^c - 1)$. Since $(m^a - 1) < (m^c - 1)$ and $(m^b - 1) \leq (m^c - 1)$, we get $s = 1$. From this follows that $c = b$.

Lemma 4. *Let G be a group of order v and $D \subseteq G$ with $|D| = k$. If for every $0 \neq g \in G$, $|D \cap (D + g)| = \lambda$ then D is a (v, k, λ)-difference set in G.*

Proof. Let $0 \neq g \in G$ be an arbitrary element. As $|D \cap (D + g)| = \lambda$, there exist λ number of distinct pairs (d, d'), $d, d' \in D$ such that $d = d' + g$, i.e., $d - d' = g$. As g is an arbitrary element of G, the multiset $[x - y : x, y \in D, x \neq y]$ contains every element in $G \backslash \{0\}$ exactly λ times. Hence the result.

Theorem 1. *Let α be a primitive element of $\mathbb{F}_{2^n}$ over $\mathbb{F}_2$. Let $U = \{0, \alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{i_{2^k-1}}\}$ be a subspace of dimension k in $\mathbb{F}_{2^n}$ such that U generates a full-length orbit. The subspace code $Orb(U)$ is an equidistant code if and only if the set of indices i_j, $1 \leq j \leq 2^k - 1$, is a difference set in $\mathbb{Z}_{2^n-1}$.*

Proof. Let $\mathrm{Orb}(U)$ be an equidistant code and let $d_s(\mathrm{Orb}(U)) = 2(k - r)$. As U generates a full-length orbit, $\dim(U \cap \beta U) = r$ for all $\beta \in \mathbb{F}_{2^n} \backslash \mathbb{F}_2$. Now consider the set $D = \{i_j \mid \alpha^{i_j} \in U\}$. Clearly $D \subseteq \mathbb{Z}_{2^n-1}$ and $|D| = 2^k - 1$. Let $j(\neq 0)$ be an arbitrary element in $\mathbb{Z}_{2^n-1}$. Then $\alpha^j \in \mathbb{F}_{2^n} \backslash \mathbb{F}_2$, and $\dim(U \cap \alpha^j U) = r$, i.e., $|\{0, \alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{i_{2^k-1}}\} \cap \{0, \alpha^{j+i_1}, \alpha^{j+i_2}, \ldots, \alpha^{j+i_{2^k-1}}\}| = 2^r$. From this we get $|D \cap (j + D)| = 2^r - 1$. As j is an arbitrary element in $\mathbb{Z}_{2^n-1} \backslash \{0\}$, by Lemma 4 we get the set of indices D to be a $(2^n - 1, 2^k - 1, 2^r - 1)$-difference set in $\mathbb{Z}_{2^n-1}$.

For the converse, let $D = \{i_j \mid \alpha^{i_j} \in U\}$ constitutes a $(2^n-1, 2^k-1, s)$-difference set in $\mathbb{Z}_{2^n-1}$. From equation (1), $s(2^n-2) = (2^k-1)(2^k-2)$. From this we get $s(2^{n-1}-1) = (2^k-1)(2^{k-1}-1)$. As $k < n$, from Lemma 3 we get $s = (2^{k-1}-1)$. This implies that the multiset $[x-y : x, y \in D, x \neq y]$ contains every element of $\mathbb{Z}_{2^n-1}\backslash\{0\}$ exactly $2^{k-1}-1$ times. Let $\alpha^m U \neq U$ be an arbitrary element in $\mathrm{Orb}(U)$. Then $m \in \mathbb{Z}_{2^n-1}\backslash\{0\}$. By Lemma 1, $|D \cap (m+D)| = 2^{k-1}-1$. Therefore, $|U \cap \alpha^m U| = 2^{k-1}$ and $\dim(U \cap \alpha^m U) = k-1$. Hence $\mathrm{Orb}(U)$ is an equidistant code.

Remark 1. The argument used in Theorem 1 cannot be applied for a subspace U in $\mathbb{F}_{q^n}$ with $q > 2$. Let α be a primitive element of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ and let $U = \{0, \alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{q^k-1}\}$ be a subspace in $\mathbb{F}_{q^n} (q > 2)$ over $\mathbb{F}_q$. Suppose that U generates a full-length orbit. Then $\alpha U = U$ for all $\alpha \in \mathbb{F}_q$. Let $D = \{i_j \mid \alpha^{i_j} \in U\}$. For $2 \in \mathbb{F}_q$, there exist a $j \in \mathbb{Z}_{q^n-1}\backslash\{0\}$ such that $2 = \alpha^j$. Now, $|D \cap (j+D)| = q^k - 1$. Thus, D is not a difference set in G.

Theorem 2. *Let α be a primitive element of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. Let $U = \{0, \alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{i_{q^k-1}}\}$ be a subspace in $\mathbb{F}_{q^n}$ of dimension k such that U generates a full-length orbit. If the subspace code $Orb(U)$ is an equidistant code then the indices i_j, $1 \leq j \leq q^k - 1$, form a relative difference set in $\mathbb{Z}_{q^n-1}$.*

Proof. Let $\mathrm{Orb}(U)$ be an equidistant subspace code, and let $d_s(\mathrm{Orb}(U)) = 2(k-r)$. Let $D = \{i_j \mid \alpha^{i_j} \in U\}$ and $N = \{j \mid \alpha^j \in \mathbb{F}_q^*\}$. Then N is a subgroup of $\mathbb{Z}_{q^n-1}$ and $|N| = q-1$. For any $i \in \mathbb{Z}_{q^n-1}\backslash N$, $\alpha^i \in \mathbb{F}_{q^n}\backslash\mathbb{F}_q$ and thus $\dim(U \cap \alpha^i U) = r$. From this, we get $|D \cap (i+D)| = q^r - 1$ for all $i \in \mathbb{Z}_{q^n-1}\backslash N$. Now for any $t \in N$, $\alpha^t \in \mathbb{F}_q$ and $\dim(U \cap \alpha^t U) = q^k$. Thus, for any $t \in N$, $|D \cap (t+D)| = q^k - 1$. Hence the set of indices D constitutes a $(q-1, \frac{q^n-1}{q-1}, q^k-1, q^k-1, q^r-1)$ relative difference set in $\mathbb{Z}_{q^n-1}$ (relative to N).

Remark 2. If we take $q = 2$ in Theorem 2, then subgroup $N = \{0\}$. In this case relative differnce set is a difference set.

Theorem 3. *There is only the trivial equidistant (full length) single-orbit cyclic subspace code in $\mathcal{P}_q(n)$ for $n \geq 3$.*

Proof. Let α be a primitive element of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ and let $U = \{0, \alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{i_{q^k-1}}\}$ be a subspace of dimension k in $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. Let $\mathrm{Orb}(U)$ be an equidistant subspace code with subspace distance $2(k-r)$. By Theorem 2 the set of indices $\{i_j \mid \alpha^{i_j} \in U\}$ constitutes a $(q-1, \frac{q^n-1}{q-1}, q^k-1, q^k-1, q^r-1)$-difference set in $\mathbb{Z}_{q^n-1}$. By Lemma 2, we get

$$(q^k-1)(q^k-2) = (q-1)\left(\frac{q^n-1}{q-1} - 1\right)(q^r-1) + (q-2)(q^k-1) \ .$$

On simplifying the above equation, we get $(q^k-1)(q^{k-1}-1) = (q^{n-1}-1)(q^r-1)$. Further this gives

$$q^{2k-1} - (q+1)q^{k-1} = q^{n+r-1} - q^{n-1} - q^r \ . \tag{2}$$

Case 1. Let $r > k-1$. On dividing both sides of equation (2) by q^{k-1}, we get

$$q^k - (q+1) = q^{n+r-k} - q^{n-k} - q^{r-k+1} \ .$$

As $n > k, r-k+1 > 0$, the right side of the above equation is a multiple of q but the left side is not. This is a contradiction.
Case 2. Let $r < k-1$. On dividing both sides of equation (2) by q^r, we get

$$q^{2k-r-1} - (q+1)q^{k-r-1} = q^{n-1} - q^{n-r-1} - 1 \ .$$

As $n > k > r+1$, the left side of the above equation is a multiple of q but the right side is not. This is a contradiction. So, we conclude that $r = k-1$. By putting the value of $r = k-1$ in (2), we get $k = n-1$. Therefore, $\dim(U) = n-1$ and $d_s(\mathrm{Orb(U)}) = 2$. Hence the result.

Remark 3. From the above proposition we conclude that for a subspace U in $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ which generates a full-length orbit, $\mathrm{Orb}(U)$ is an equidistant code if and only if $\dim(U) = n-1$.

Now, we consider the subspaces which do not generate a full-length orbit. The result of Theorem 2 holds for such codes. This is shown in the next theorem.

Theorem 4. *Let α be a primitive element of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. Let $U = \{0, \alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{i_{q^k-1}}\}$ be a subspace in $\mathbb{F}_{q^n}$ of dimension k such that U does not generate a full-length orbit. If the subspace code $Orb(U)$ is an equidistant code, then the indices i_j, $1 \le j \le q^k - 1$ form a relative difference set in $\mathbb{Z}_{q^n-1}$.*

Proof. Let $\mathrm{Orb}(U)$ be an equidistant subspace code with subspace distance $2(k-r)$. Let $\mathrm{Stab}(U) = \mathbb{F}^*_{q^t}$ for some $1 < t < k$ and t divides $\gcd(k,n)$. Let $N = \{i_j \mid \alpha^{i_j} \in \mathbb{F}^*_{q^t}\}$. Then N is a subgroup of $\mathbb{Z}_{q^n-1}$. Clearly, the cardinality of N is $q^t - 1$. Let $D = \{i_j \mid \alpha^{i_j} \in U\}$. For any $j \in N$, $U = \alpha^j U$. This gives $|D \cap (j + D)| = q^k - 1$. For any $m \in \mathbb{Z}_{q^n-1} \backslash N$, $\dim(U \cap \alpha^m U) = q^r$. So, we get $|D \cap (m + D)| = q^r - 1$. By Lemma 4, the set of indices D constitutes a $(q^t - 1, \frac{q^n-1}{q^t-1}, q^k - 1, q^k - 1, q^r - 1)$-relative difference set in $\mathbb{Z}_{q^n-1}$.

Theorem 5. *There is only trivial equidistant single-orbit cyclic subspace code in $\mathcal{P}_q(n)$ for $n \ge 3$.*

Proof. We have proved the result in Theorem 3 for subspaces which generate the full-length orbit. Now we prove the result for subspaces which do not generate a full-length orbit. Let α be a primitive element in $\mathbb{F}_{q^n}$. Let $U = \{0, \alpha^{i_1}, \alpha^{i_2}, \ldots, \alpha^{q^k-1}\}$ be a subspace in $\mathbb{F}_{q^n}$ and let $\mathrm{Stab}(U) = \mathbb{F}^*_{q^t}$. Let $D = \{i_j \mid \alpha^{i_j} \in U\}$. Let $\mathrm{Orb}(U)$ be an equidistant subspace code with subspace distance $2(k-r)$. By Theorem 4 the set of indices D constitutes a $(q^t - 1, \frac{q^n-1}{q^t-1}, q^k - 1, q^k - 1, q^r - 1)$-relative difference set in $\mathbb{Z}_{q^n-1}$. By Lemma 2,

$$(q^k-1)(q^k-2) = (q^t-1)\left(\frac{q^n-1}{q^t-1} - 1\right)(q^r-1) + (q^t-2)(q^k-1) \ .$$

As $k > t$, by simplifying the above equation we get

$$q^{2k-t} - q^k - q^{k-t} = q^{n+r-t} - q^{n-t} - q^r \ . \tag{3}$$

Case 1. Let $r > k-t$. On dividing both sides of the equation (2) by q^{k-t}, we get

$$q^k - q^t - 1 = q^{n+r-k} - q^{n-k} - q^{r-k+t} \ .$$

Clearly, the right side is a multiple of q but the left side is not. This is a contradiction.
Case 2. Let $r < k-t$. On dividing both sides of equation (2) by q^r, we get

$$q^{2k-t-r} - q^{k-r} - q^{k-t-r} = q^{n-t} - q^{n-t-r} - 1 \ .$$

As $n > k > t$ and $k > r$, the left side is a multiple of q but the right side is not. This is a contradiction.

So, we conclude that $r = k-t$. By putting the value of $r = k-t$ in equation (3) we get $k = n-t$. Thus, the dimension of subspace U is $n-t$ and the subspace distance of $\mathrm{Orb}(U)$ is $2t$. Hence the result.

4 Equidistant quasi-cyclic subspace codes

Definition 3. *[8] Fix an element $\beta \in \mathbb{F}_{q^n}^* \backslash \{1\}$. Let U be an $\mathbb{F}_q$-subspace in $\mathbb{F}_{q^n}$. The β-cyclic orbit code generated by U is defined as the set*

$$Orb_\beta(U) = \{\beta^i U \mid i = 0, 1, \ldots, |\beta| - 1\} \ .$$

If β is a primitive element of $\mathbb{F}_{q^n}$, i.e., $\mathbb{F}_{q^n}^* = \langle\beta\rangle$. We write $Orb_\beta(U)$ simply as $Orb(U)$.

A subspace code C is called a *quasi-cyclic* subspace code if $\alpha U \in C$ for all $U \in C$ and $\alpha \in G$, where G is a multiplicative subgroup of $\mathbb{F}_{q^n}^*$[12]. If $\beta \in \mathbb{F}_{q^n}^* \backslash \{1\}$ is not a primitive element of $\mathbb{F}_{q^n}$, we call $\mathrm{Orb}_\beta(U)$ single-orbit quasi-cyclic subspace code.

In the previous section we have proved that there exist only trivial equidistant single-orbit cyclic subspace codes. Now, we turn to the β-cyclic orbit code $\mathrm{Orb}_\beta(U)$, where β is not a primitive element of $\mathbb{F}_{q^n}$. The following examples show that there exist equidistant single-orbit quasi-cyclic subspace codes.

Example 1. Consider an irreducible monic polynomial $p(x) = x^9 + x^8 + x^5 + 4x^4 + x^3 + 2x^2 + 4x + 3$ of degree 9 over $\mathbb{F}_5$. Let α be a root of the polynomial $p(x)$. Then $\mathbb{F}_5(\alpha)$ is the extension field of degree 9 over $\mathbb{F}_5$. Consider the subspace $U = \delta_1\mathbb{F}_5 \oplus \delta_2\mathbb{F}_5 \oplus \delta_3\mathbb{F}_5 \oplus \delta_4\mathbb{F}_5$, where $\delta_1 = \alpha^6 + 2\alpha^2 + 1$, $\delta_2 = \alpha^5 + \alpha^4 + \alpha^2 + 2$, $\delta_3 = 2\alpha^4 + \alpha^3 + 2\alpha^2 + \alpha$ and $\delta_4 = \alpha^8 + \alpha^6 + \alpha^3$. The dimension of U over $\mathbb{F}_5$ is 4 and the subspace U generates a full-length orbit. Let $\omega = 2\alpha^8 + 3\alpha^6 + 3\alpha^5 + 2\alpha^4 + 2\alpha^3$. The order of ω in $\mathbb{F}_{5^9}^*$ is 76. Consider the orbit code $\mathrm{Orb}_\omega(U) = \{\omega^i U \mid 1 \leq i \leq |\omega|\}$. Using the Magma computational algebra system [3], we computed that

$\dim(U \cap zU) = 0$ for all $zU \in \mathrm{Orb}_\omega(U),\ zU \neq U$ and $|\mathrm{Orb}_\omega(U)| = 19$. Thus $\mathrm{Orb}_\omega(U)$ is an equidistant code with subspace distance 8. Now let $G = \mathbb{F}_{5^3}$ and let $\beta = 4\alpha^8 + \alpha^7 + \alpha^6 + \alpha^5 + 3\alpha^4 + 2\alpha^2 + 4\alpha + 3$ is a generator of group $G\backslash\{0\}$. Consider the subspace code $\mathrm{Orb}_\beta(U) = \{\beta^i U \mid 0 \leq i \leq |\beta| - 1\}$. Using the Magma computational algebra system [3], we computed that $\dim(U \cap cU) = 0$ for all $cU \in \mathrm{Orb}_\beta U,\ cU \neq U$ and $|\mathrm{Orb}_\beta(U)| = 31$. Hence $\mathrm{Orb}_\beta(U)$ is an equidistant code with subspace distance 8.

Example 2. Consider an irreducible monic polynomial $p(x) = x^{12} + x^6 + x^5 + x^4 + x^2 + 2$ of degree 12 over $\mathbb{F}_3$. Let α be a root of the polynomial $p(x)$. Then $\mathbb{F}_3(\alpha)$ is the extension field of degree 12 over $\mathbb{F}_3$. Consider the subspace $U = \langle z^{66430}, z^{199290}, z^{40880}, z^{81760}, z^{286540}, z^{374556} \rangle_{\mathbb{F}_3}$. The dimension of U over $\mathbb{F}_3$ is 6 and the subspace U generates a full-length orbit. The subfield $\mathbb{F}_{3^3}$ of $\mathbb{F}_{3^{12}}$ is contained in U. Let c be a primitive element of $\mathbb{F}_{3^3}$. Then the order of c is 26. Consider the orbit code $\mathrm{Orb}_c(U) = \{c^i U : i = 1, \ldots, |c|\}$. Clearly $\mathbb{F}_{3^3} \subseteq U \cap c^i U$. Using the magma computational algebra system[3], we computed that $U \cap c^i U = \mathbb{F}_{3^3}$ for all $c^i U \in \mathrm{Orb}_c(U)$ with $c^i U \neq U$. Thus $\mathrm{Orb}_c(U)$ is an equidistant subspace code of size 13 and minimum distance 6. Now, let $\gamma = z^{9490}$. The order of γ in $\mathbb{F}^*_{3^{12}}$ is 56. Consider the orbit code $\mathrm{Orb}_\gamma(U) = \{\gamma^i U : 1 \leq i \leq |\beta|\}$. By using the magma compuatational algebra system [3], we obtained that $\dim(U \cap \omega U) = 2$ for all $\omega U \in \mathrm{Orb}_\gamma(U)$ with $\omega U \neq U$. Thus $\mathrm{Orb}_\gamma(U)$ is an equidistant code of size 28 and minimum distance 8.

Note 1. While considering the equidistant quasi-cyclic subspace code $\mathrm{Orb}_\beta(U)$, we will take $\beta \notin \mathrm{Stab}(U)$. Otherwise, being a subfield of $\mathbb{F}_{q^n}$ stabilizer of U will contain the multiplicative subgroup of $\mathbb{F}^*_{q^n}$ generated by β, and $\mathrm{Orb}_\beta(U)$ will contain a single subspace U.

Proposition 2. *Let n be an even number and U be a subspace in $\mathbb{F}_{q^n}$. Let β be an element of degree 2 in $\mathbb{F}_{q^n}$ such that $\beta \notin Stab(U)$. Then $Orb_\beta(U)$ is an equidistant subspace code.*

Proof. Since $\beta \notin \mathrm{Stab}(U)$, we have $|\mathrm{Orb}_\beta(U)| \geq 2$. As β is an element of degree 2 in $\mathbb{F}_{q^n}$, $\mathbb{F}_q[\beta] = \{a + c\beta \mid a, c \in \mathbb{F}_q\}$. Clearly, $\{\beta^i \mid 0 \leq i \leq |\beta| - 1\} \subseteq \mathbb{F}_q[\beta]$. For any $\gamma \in \mathbb{F}^*_{q^n}$ and $s \in \mathbb{F}^*_q$, $\dim(U \cap \gamma U) = \dim(U \cap (\gamma + s)U)$. Therefore, $\dim(U \cap \beta U) = \dim(U \cap \delta U)$ for all $\delta \in \mathbb{F}_q[\beta]\backslash\mathbb{F}_q$. Thus, $\dim(U \cap \beta U) = \dim(U \cap \beta^i U)$ for all $i,\ 1 \leq i \leq |\beta| - 1$. Hence $\mathrm{Orb}_\beta(U)$ is an equidistant code.

Remark 4. Let $G = \mathbb{F}^*_{q^2}$ be a multiplicative subgroup of $\mathbb{F}^*_{q^n}$ and let γ be a generator of G, i.e., $G = \{\gamma^i \mid 0 \leq i \leq q^2 - 2\}$. Clearly, degree of γ over $\mathbb{F}_q$ is 2. If $\mathbb{F}^*_{q^2} \nsubseteq \mathrm{Stab}(U)$, by Proposition 2, $\mathrm{Orb}_\gamma(U)$ is an equidistant subspace code.

Acknowledgments The first author would like to thank Ministry of Education, India for providing financial support.

References

1. Bartoli, D., Pavese, F.: A note on equidistant subspace codes. Discret. Appl. Math. **198**, 291–296 (2016)
2. Basu, P.: Equidistant linear codes in projective spaces (2021), preprint at `https://arxiv.org/abs/2107.10820`
3. Bosma, W., Cannon, J.: Handbook of Magma Functions. School of Mathematics and Statistics, Univ. of Sydney (1995)
4. Etzion, T., Raviv, N.: Equidistant codes in the Grassmannian. Discret. Appl. Math. **186**, 87–97 (2015)
5. Etzion, T., Vardy, A.: Error-correcting codes in projective space. IEEE Trans. Inf. Theory **57**(2), 1165–1173 (2011)
6. Ghatak, A.: Construction of Singer subgroup orbit codes based on cyclic difference sets (2014), In: Proceedings of the Twentieth National Conference on Communications (NCC 2014), pp. 1-4, Kanpur, India. IEEE (2014)
7. Gluesing-Luerssen, H., Lehmann, H.: Distance distributions of cyclic orbit codes. Des. Codes Cryptogr. **89**, 447–470 (2021)
8. Gluesing-Luerssen, H., Morrison, K., Troha, C.: Cyclic orbit codes and stabilizer subfields. Adv. Math. Commun. **9**(2), 177–197 (2015)
9. Gorla, E., Ravagnani, A.: Equidistant subspace codes. Linear Algebra Appl. **490**, 48–65 (2016)
10. Jungnickel, D.: On automorphism groups of divisible designs. Can. J. Math. **34**(2), 257–297 (1982)
11. Kötter, R., Kschischang, F., R.: Coding for errors and erasures in random network coding. IEEE Trans. Inf. Theory **54**, 3579–3591 (2008)
12. Otal, K., Özbudak, F.: Cyclic subspace codes via subspace polynomials. Des. Codes Cryptogr. **85**(2), 191–204 (2017)
13. Raviv, N., Etzion, T.: Distributed storage systems based on intersecting subspace codes (2015), In: International Symposium on Information Theory, pp. 1462-1466 (2015)
14. Stinson, D., R.: Combinatorial Designs: Constructions and Analysis. Springer, New York (2004)
15. Trautmann, A., L., Manganiello, F., Braun, M., Rosenthal, J.: Cyclic orbit codes. IEEE Trans. Inf. Theory **59**(11), 7386–7404 (2013)
16. Van Lint, J.H., Wilson, R.M.: A Course in Combinatorics 2nd edn. Cambridge University Press, Cambridge (2001)

Weight Distribution of the Binary Reed-Muller Code $\mathcal{R}(4, 9)$

Miroslav Markov and Yuri Borissov

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences
miro@math.bas.bg, youri@math.bas.bg

Abstract. We compute the weight distribution of the $\mathcal{R}(4, 9)$ by combining the approach described in D. V. Sarwate's Ph.D. thesis from 1973 with knowledge on the affine equivalence classification of Boolean functions. To solve this problem posed, e.g., in the book of MacWilliams and Sloane, we apply a refined approach based on the classification of Boolean quartic forms in eight variables due to Ph. Langevin and G. Leander, and recent results on the classification of the quotient space $\mathcal{R}(4, 7)/\mathcal{R}(2, 7)$ due to V. Gillot and Ph. Langevin.

Keywords: binary Reed-Muller code, weight distribution, affine equivalence

1 Introduction

For basic coding theoretical notions, we refer to [13]. All considered codes in this paper are binary, i.e., over the alphabet $\mathbb{F}_2 = \{0, 1\}$.

The binary Reed-Muller codes form one of the oldest studied families of codes invented in 1950s and have an easy-to-implement decoding algorithm based on majority-logic circuits. However, there are few general results about their weight structure, i.e., the weight distributions are known only for:

- the 1^{st} and 2^{nd} order codes of that kind [17] (1970);
- arbitrary order when the weight $< 2d$ [8] (1970), and later on (in 1976) had been extended for weights $< 2.5d$ where d is the minimum weight [9];

Results on the weight spectra of some third order codes are presented in an earlier work [20], and on the spectra of whole families of binary Reed-Muller codes in the very recent work [2]. Some partial results concerning the weight distribution of the third and fourth order Reed-Muller codes are obtained in [15], [9], [18] and [19]. For more information about the weight distributions of binary Reed-Muller codes of particular lengths and orders, the reader is referred to [16].

The weight spectrum of the fourth order Reed-Muller code $\mathcal{R}(4, 9)$ of length 512, has been found in [2] and presented as a numerical example which demonstrates the technique developed there. To our knowledge, there have been very few attempts to determine the (exact) weight distribution of this code, which was listed among the smallest Reed-Muller codes whose weight distributions were unknown (in 1977) (see, [13, p. 447]). Specifically, in the concluding remarks of his Ph.D. thesis [15],

D. V. Sarwate has discussed the applicability of the methods described by him to Reed-Muller codes of lengths larger than 256. He has estimated that there are too many equivalence classes of cosets from the desired type and has come into conclusion that enumerating the $\mathcal{R}(4,9)$ seems out of reach through them. Another promising way to attack the considered problem consists in using the fact that we are dealing with a double-even binary self-dual code and a general form of the weight enumerators of such codes is known from the work of A. M. Gleason (see, e.g., [13, Ch.19]). But, although this second approach has proven itself in the case of shorter codes of that kind and requires modest computational efforts, for its successful application one needs more intrinsic knowledge about the $\mathcal{R}(4,9)$ than those presented in [8] (see, [3, Ch. 11] for details).

This paper is organized as follows. In the next section we give the necessary preliminaries. In Section 3 a refined approach to the problem under consideration enabling one to save computational efforts is exposed. Some conclusions are drawn in the last section.

2 Preliminaries

For basic knowledge on Boolean functions and their applications in Coding Theory and Cryptography, we direct the reader to [1] and [4]. Herein, for the sake of completeness, we recall the classical definition of the binary Reed-Muller code.

Definition 1. *The r^{th} order binary Reed-Muller (or RM) code $\mathcal{R}(r,m)$ of length $n=2^m$, for $0 \leq r \leq m$, is the set of all binary vectors $\mathbf{f}$ of length n which are truth tables of Boolean functions $f(\mathbf{x}), \mathbf{x}=(x_1,\ldots,x_m)$, having algebraic normal forms of degree at most r.*

Henceforth the binary vector $\mathbf{f}$ of length 2^m will be identified with corresponding Boolean function f in m variables.

In order to present our results we need to recall the definition of the weight distribution of a code, i.e., an arbitrary set $\mathbf{C}$ of vectors with fixed length n (this definition holds in particular for cosets of binary linear codes).

Definition 2. *The weight distribution of a code $\mathbf{C}$ of length n is the vector $W(\mathbf{C})=(W_0,\ldots,W_n)$, where W_i denotes the number of codewords with Hamming weight i.*

Accordingly, we recall the definition of the simplest weight enumerator of a code.

Definition 3. *Weight enumerator of a code $\mathbf{C}$ possessing weight distribution $W(\mathbf{C})=(W_0,\ldots,W_n)$ is defined as the following polynomial in the indeterminate z:*

$$\mathcal{W}[z;\mathbf{C}]=\sum_{i=0}^{n} W_i z^i.$$

In this paper, we make use of two facts for the first time exposed in [15] and stated in the next two theorems. (For $0 \leq r \leq m$, the set of all homogeneous polynomials on m binary variables of algebraic degree r adjoined with the 0 is denoted by $\mathcal{H}^{(r)}(m)$.)

Theorem 1. *([15, 5.12]) For $0 \leq r \leq m$, the following holds:*

$$\mathcal{W}[z; \mathcal{R}(r+2, m+2)] = \sum_{p \in \mathcal{H}^{(r+2)}(m+1)} \mathcal{W}^2[z; p + \mathcal{R}(r+1, m+1)].$$

Theorem 2. *([15, 5.13]) Let $p = e + fx_{m+1}$, with given $e \in \mathcal{H}^{(r+2)}(m)$ and $f \in \mathcal{H}^{(r+1)}(m)$. Then the weight enumerator of the coset $\mathcal{C}(p) = p + \mathcal{R}(r+1, m+1)$ equals to:*

$$(*) \quad \sum_{g \in \mathcal{H}^{(r+1)}(m)} \mathcal{W}[z; e + g + \mathcal{R}(r, m)] \cdot \mathcal{W}[z; e + g + f + \mathcal{R}(r, m)].$$

For definition of the general affine group $GA(m)$ and its subgroup the general linear group $GL(m, 2)$, we refer to [13, Ch.13.9]. The action of $A \in GA(m)$ on a Boolean function $f(\mathbf{x})$ will be denoted by $f \circ A$, i.e., $f \circ A = f(A(\mathbf{x}))$. Another necessary definition is that of affine equivalence of two cosets of Reed-Muller code:

Definition 4. *The cosets C_1 and C_2 of $\mathcal{R}(r, m)$ with representatives $f_1 \in C_1$, $f_2 \in C_2$, respectively, are called affine equivalent if there exist a transformation $A \in GA(m)$ such that $f_1 \circ A = f_2$.*

In this article, we extensively use the following well-known property (see, e.g., [7]):
Property $\mathcal{P}$. *The weight enumerators of two affine equivalent cosets of a Reed-Muller code are identical.*

Affine equivalence classification of the cosets of RM codes is useful in studying important coding theoretical and cryptographic properties of Boolean functions comprising them. A strategy how to compute the complete classification of Boolean quartic forms in eight variables, i.e., the classification of the quotient space $\mathcal{R}(4, 8)/\mathcal{R}(3, 8)$ under the action of $GL(8, 2)$, is presented in [12]. Here, just as an extract of this result, we point out that the Boolean quartic forms of eight variables can be classified in 999 (see, as well [7]) linear equivalence classes listed in [10]. Recently, the interest in that topic has been renewed by [5] which (among other things) provides affine equivalence classification of the quotient space $\mathcal{R}(4, 7)/\mathcal{R}(2, 7)$. The authors of [5] and [12] have also outlined applications of their results concerning the covering radii of some RM codes, and Boolean functions in the family of bent ones. In Section 3, we point out yet another application, namely, computing the weight distribution of $\mathcal{R}(4, 9)$.

3 The refined approach

3.1 Rationale

Now, we describe a strategy following which makes feasible the computation of interest.

For $1 \leq r \leq m$, let $n(r, m)$ be the number of $GL(m, 2)-$orbits in the quotient space $\mathcal{R}^*(r, m) = \mathcal{R}(r, m)/\mathcal{R}(r-1, m)$. Also assume that an arbitrary numbering of these orbits (linear equivalence classes) has been fixed.

Corollary 1. *Let $p_i \in \mathcal{H}^{(r+2)}(m+1)$ be a representative of the i^{th} linear equivalence class in $\mathcal{R}^*(r+2, m+1)$ with size L_i. Then the following holds:*

$$\mathcal{W}[z; \mathcal{R}(r+2, m+2)] = \sum_{i=1}^{n(r+2,m+1)} L_i \mathcal{W}^2[z; p_i + \mathcal{R}(r+1, m+1)]. \quad (1)$$

Proof. The assertion is an immediate consequence of Theorem 1 and Property $\mathcal{P}$. □

The above corollary reduces the number of needed weight enumerator computations to the number $n(r+2, m+1)$ which is significantly smaller than the straightforward $|\mathcal{H}^{(r+2)}(m+1)| = 2^{\binom{m+1}{r+2}}$ in Theorem 1. For instance, as it has been already mentioned, $n(4, 8) = 999$ which should be compared with 2^{70}.

Remark 1. Corollary 1 is implicitly used in [15] for shorter RM codes.

Next, we can state another claim that allows further reduction of the cost.

Corollary 2. *For given $e \in \mathcal{H}^{(r+2)}(m)$, let $\mathcal{H}^{(r+1)}(m)$ be partitioned into blocks (subsets) $G_i, 1 \leq i \leq s$ with the property that whenever $g \in G_i$ the enumerator $\mathcal{W}[z; e + g + \mathcal{R}(r, m)]$ is a (distinct) constant polynomial $w_i(z)$. Then the following holds:*

*(**a**) the weight enumerator of the coset $\mathcal{C}(p) = p + \mathcal{R}(r+1, m+1), p = e + f x_{m+1}$ for fixed $f \in \mathcal{H}^{(r+1)}(m)$, can be expressed by*

$$\sum_{i=1}^{s} w_i(z) \left(\sum_{g \in G_i} \mathcal{W}[z; e + g + f + \mathcal{R}(r, m)] \right).$$

*(**b**) the number of polynomial multiplications for computing the aforesaid weight enumerator equals to s, i.e. the number of distinct weight enumerators $\mathcal{W}[z; e + g + \mathcal{R}(r, m)], g \in \mathcal{H}^{(r+1)}(m)$, while that of polynomial additions is $2^{\binom{m}{r+1}} - s$.*

Proof. Rearranging the summands in (*) from Theorem 2 and putting outside of brackets the common multipliers $w_i(z)$ proves (**a**). The claim (**b**) is an immediate consequence of (**a**). □

The affine equivalence classification of $\mathcal{R}(r+2, m)/\mathcal{R}(r, m)$ enables to substantiate the usage of Corollary 2. To see this, let us recall the following definition:

Definition 5. *The subgroup $\mathcal{S}t(e)$ of $GA(m)$ that fixes $e \in \mathcal{H}^{(r+2)}(m)$, i.e., for each $A \in \mathcal{S}t(e)$ it holds: $e \circ A \in e + \mathcal{R}(r+1, m)$, is called stabilizer of e in $GA(m)$.*

For given $e \in \mathcal{H}^{(r+2)}(m)$, the stabilizer $\mathcal{S}t(e)$ partitions the cosets of the form $e + g + \mathcal{R}(r, m)$ where $g \in \mathcal{H}^{(r+1)}(m)$ into disjoint orbits. Denote this partition by $\Delta(e)$. Furthermore, Property $\mathcal{P}$ implies that the enumerator $\mathcal{W}[z; e + g + \mathcal{R}(r, m)]$ is preserved when g runs over an orbit of $\Delta(e)$. The latter permits to constitute efficiently the coarse partition $\Delta'(e) = \{G_i, 1 \leq i \leq s\}$ of $\mathcal{H}^{(r+1)}(m)$ (see, Corollary 2) by merging those orbits possessing identical weight enumerators (the latter ones being computed in advance on chosen orbit representatives).

3.2 Computing $\mathcal{W}[z; \mathcal{R}(4, 9)]$

Our computational work is divided into two main phases: a pre-computing and actual computing.

The aim of pre-computing is to provide tools for efficient computation of the expression (*) in Theorem 2 given a specific e and f, and is carried out following Corollary 2 and the subsequent considerations from the previous subsection.
Let $\mathcal{E}(4, 7)$ be the set of representatives of the twelve linear equivalence classes of $\mathcal{R}^*(4, 7)$ given in [11]. For fixed $e \in \mathcal{E}(4, 7)$, the pre-computing involves the following three tasks:

- $\mathcal{T}1$: Constitute and store the orbits of the partition $\Delta(e)$;
- $\mathcal{T}2$: Compute the weight enumerators of the cosets $e + g + \mathcal{R}(2, 7)$ when g varies over a set of representatives of $\Delta(e)$'s orbits;
- $\mathcal{T}3$: Merge the orbits with identical weight enumerators to obtain the coarse partition $\Delta'(e)$, and make data arrangement permitting for given $f \in \mathcal{H}^{(3)}(7)$ to look up the identifier of a block in $\Delta'(e)$ containing f (respectively, to have direct access to the common weight enumerator).

For all $e \in \mathcal{E}(4, 7)$, we present in **Table 1** of the **Appendix** the sizes of partitions $\Delta(e)$ and $\Delta'(e)$, respectively.

Remark 2. It is worth pointing out that:

- the task $\mathcal{T}1$ is efficiently performed based on the so-called "orbit algorithm" [6] using the set of generators of the stabilizer $\mathcal{S}t(e)$ provided by [11];
- the task $\mathcal{T}2$ can be carried out simultaneously for all representatives by exhaustive generation of the codewords of $\mathcal{R}(2, 7)$ based on some Gray code.

Now, following the strategy described in Section 3.1, we present an algorithm for computing the weight enumerator $\mathcal{W}[z; C(p)]$ of the coset $C(p) = p + \mathcal{R}(3, 8)$ where $p = e + fx_8$ for fixed $e \in \mathcal{E}(4, 7)$ and a given input $f \in \mathcal{H}^{(3)}(7)$. Note that it can be implemented as a subroutine. Recall also that the common weight enumerator $w_i(z)$ corresponding to the block G_i in $\Delta'(e)$ has been already computed in the pre-computing task $\mathcal{T}2$ where $1 \leq i \leq |\Delta'(e)| = s(e)$.

Algorithm 1: Returning the weight enumerator $\mathcal{W}[z; C(p)]$ where $p = e + fx_8$ for fixed e and a given $f \in \mathcal{H}^{(3)}(7)$

```
1 U[z] := 0;
2 for i in [1,s(e)] do
3 |   UU(z) := 0;
4 |   for g in G[i] do
5 |   |   j := FindBlock(g+f);
6 |   |   UU(z) := UU(z) + w[j](z);
7 |   U(z) := U(z) + w[i](z) * UU(z);
8 W[z; C(p)] := U(z);
```

In the actual computing, we apply formula (1) supposing that a set $\mathcal{S}$ of pairs: (representative p_i, orbit size L_i) for the $i-$th class $O_i, 1 \leq i \leq 999$, of the classification

of $\mathcal{R}^*(4,8)$ is available. W.l.o.g., we may assume each p_i is of the form $e+f_i x_8$ for some $e \in \mathcal{E}(4,7)$ and $f_i \in \mathcal{H}^{(3)}(7)$, so the set of classes is naturally partitioned into subsets $\mathcal{O}(e)$ of cardinalities $n(e), e \in \mathcal{E}(4,7)$. (The values $n(e)$ are given in the first column of **Table 2** of the **Appendix**.) Bellow, we present an algorithm for computing the sum in formula (1) and thus $\mathcal{W}[z;\mathcal{R}(4,9)]$. (Note that we call the subroutine $\mathcal{W}[z;C(p)]$.)

Algorithm 2: Computing $\mathcal{W}[z;\mathcal{R}(4,9)]$

```
V(z) := 0;
for e ∈ E(4,7) do
    for j in [1,n(e)] do
        p := Representative(O(e)[j]);
        L := Size(O(e)[j]);
        V(z) := V(z) + L * W^2[z;C(p)];
W[z;R(4,9)] = V(z);
```

The data present in [10] contains information to form a set $\mathcal{S}'$ of kind similar to $\mathcal{S}$. However, the representatives p'_i there are of the form $e' + f'_i x_8$ where e's constitute different set of representatives of the twelve classes of $\mathcal{R}^*(4,7)$, say $\mathcal{E}'(4,7)$. For some elements of $\mathcal{E}(4,7)$ and $\mathcal{E}'(4,7)$, their linear equivalence is evident by eye inspection. For the remaining, we determined those which are linearly equivalent by computing the vectors of invariants of their duals (see, for details [7, pp. 115-117]). The matching found is represented in the rows of **Table 2** where $\overline{\mathcal{E}}(4,7)$ and $\overline{\mathcal{E}}'(4,7)$ are the sets consisting of dual forms of those in $\mathcal{E}(4,7)$ and $\mathcal{E}'(4,7)$, respectively. To find out a nonsingular (7×7) matrix $\mathbf{A}$ with property that $e' \circ \mathbf{A} \in e + \mathcal{R}(3,7)$ for thus determined pairs (e', e), we wrote a simple program in *C* which generates at random such a nonsingular square matrix and then checks the imposed condition. This technique is sufficiently efficient (due to relatively large stabilizers sizes, see, [12, **Table 2**]) and the program finished successfully its work in reasonable time. For similar technique to exploring affine equivalence of Boolean functions, we refer the reader to [14]. The obtained results are presented in the last column of **Table 2** of the **Appendix**. Finally, acting on corresponding $f'_i, 1 \leq i \leq 999$ by the resulting linear transformations (of course, ignoring the terms of degree less than 3), we obtain a type of set required by the **Algorithm 2**. The weight distribution got is presented in **Table 3** of the **Appendix**.

Remark 3. The functions FindBlock$(\cdot)$, Representative$(\cdot)$ and Size$(\cdot)$ have names that are self-explanatory when it comes to their intended purpose.

3.3 Evaluating the computational costs

Following [6] and [11], we estimate that the computational cost of task $\mathcal{T}1$ is $|\mathcal{H}^{(3)}(7)| \times \sum_{e \in \mathcal{E}(4,7)} |Sg(e)| = 2^{35} \times 26 \approx 2^{39.7}$ affine transformations where $Sg(e)$ denotes the set of generators of the stabilizer $\mathcal{S}t(e)$. The computational complexity of task $\mathcal{T}2$ is in total proportional to the product $68443 \times 2^{29} \approx 2^{45.06}$ with the first factor being the number of classes of $\mathcal{R}(4,7)/\mathcal{R}(2,7)$ and the second being the size of $\mathcal{R}(2,7)$. Task $\mathcal{T}3$ can be carried out by applying some sorting technique. In summary, the pre-computing in

case $r = 2$ and $m = 7$ is efficiently performed. In addition, we note that the compressed storing of orbit and data arrangement into RAM needs at most 124 GB of memory.

In the actual computing, for every $e \in \mathcal{E}(4,7)$, **Algorithm 1** requires $|\Delta'(e)|$ multiplications and about 2^{35} additions of degree 128 polynomials with nonnegative integer coefficients. Therefore, **Algorithm 2** requires $\sum_{e\in\mathcal{E}(4,7)} n(e) \times |\Delta'(e)| = 1827252 \approx 2^{20.8}$ multiplications and about $999 \times 2^{35} \approx 2^{45}$ additions of polynomials of that kind; and 999 squarings of degree 256 polynomials and some additional operations with negligible cost, of course.

Remark 4. The straightforward application of Theorem 2 (based on the original partition $\Delta(e)$) will require about 6 times more multiplications of degree 128 polynomials than the actually executed.

Remark 5. Finally, we have two remarks concerning the implementation:

- To meet the memory limitations, we use the appropriate for that aim Delta compression and VByte encoding of the data. These techniques are important to our computer-aided solution, but the details are omitted because of their merely auxiliary role;
- We use the 256-bit CPU registers which ensures that arithmetic operations are performed efficiently and eliminates the need to further estimate the number of processor operations.

4 Conclusion

In this article, thanks to recent advances in the classification of Boolean functions [5],[12] and the utilization of modern high-performance computers, a solution to the problem at hand is obtained. However, we should admit that it may not be doable to push this line of research much further due to the way in which the computational burden increases with code length.

Acknowledgments

We would like to thank the reviewers for their valuable suggestions and comments which significantly improve this paper. We are also grateful to Vladimir Tonchev for pointing out the problem and his stimulating discussions, as well as to Peter Boyvalenkov for his interest and support during the course of this research.

This work was supported, in part, by the Ministry of Education and Science of Bulgaria under the Grant No. DO1-325/01.12.2023 "National Centre for High-performance and Distributed Computing" (NCHDC). The authors acknowledge the provided access to the e-infrastructure of NCHDC.

Appendix

Table 1. Sizes of partitions $\Delta(e)$ and $\Delta'(e)$

$e \in \mathcal{E}(4,7)$: ANF's according to ([11])	$\lvert\Delta(e)\rvert$	$\lvert\Delta'(e)\rvert$
0	12	12
4567	63	52
1235+1345+1356+1456+2346+2356+2456	130	112
2367+4567	289	182
1237+4567	480	306
1257+1367+4567	730	395
1237+1247+1357+2367+4567	204	157
1236+1257+1345+1467+2347+2456+3567	1098	675
1236+1356+1567+2357+2467+2567+3456	1340	811
1367+2345+2356+3456+4567	6449	2170
1234+1237+1267+1567+2345+3456+4567	23988	3377
1236+1367+1567+2345+3456+3457+3467	33660	4636

Table 2. The matching between $\mathcal{E}'(4,7)$ and $\mathcal{E}(4,7)$

Distribution of $n(e)$	$\overline{\mathcal{E}}'(4,7)$	$\overline{\mathcal{E}}(4,7)$	Linear transition transformation
3	0	0	[1000000 0100000 0010000 0001000 0000100 0000010 0000001]
2	123	123	[1000000 0100000 0010000 0001000 0000100 0000010 0000001]
21	127+136+145	137+147+157+237+247+267+467	[0011001 0011110 0100110 1011000 1111010 1001100 0001100]
15	125+134	123+145	[1000000 0100000 0001000 0000100 0010000 0000010 0000001]
89	126+345	123+456	[1000000 0100000 0001000 0000100 0000010 0010000 0000001]
56	126+135+234	123+245+346	[0100000 0010000 0001000 0000010 0000100 1000000 0000001]
10	135+146+235+236+245	123+145+246+356+456	[1000000 0000010 0001000 0010000 0000100 0100000 0000001]
7	127+136+145+234	124+137+156+235+267+346+457	[0110001 1011001 0110011 0111010 1100101 0010111 1001011]
502	125+134+135+167+247+357	127+134+135+146+234+247+457	[0001000 0010000 0000001 0000100 0100000 0000010 1100110]
1	123+247+356	123+127+147+167+245	[0010000 0110011 1010000 0001110 0000001 0010011 0000100]
1	147+156+237+246+345	123+127+167+234+345+456+567	[0101010 1001010 1001001 1111111 0011000 0100010 1001011]
292	127+146+236+345	125+126+127+167+234+245+457	[0100111 0001110 0110110 1011000 0000010 0000100 0010110]

Table 3. Weight Distribution of the [512,256,32] Reed-Muller code

Weight	Number of codewords
0 512	1
32 480	52955952
48 464	919315326720
56 456	271767121346560
60 452	860689275027456
64 448	89163020044002040
68 444	1777323352931696640
72 440	64959328938397057024
76 436	2094952122987829002240
80 432	86129855718211879936768
84 428	3718387228743293604986880
88 424	216407674400647746861465600
92 420	15958945395035022932054114304
96 416	1570964763114053055495174389136
100 412	207755244457303752035637154283520
104 408	34164336816436357675455725024378880
108 404	5992987676360073735151889707696128000
112 400	983217921810034263357552475089021004288
116 396	140881159168600922710983130625456163782656
120 392	17178463264607761296016540993629780705771520
124 388	1770270551281316280504947079180771901717872640
128 384	154198773988541804525321284585063483246993999900
132 380	113804373667128124744559508641773260684479892029 44
136 376	7137934452988742116078397968797161061857152802160 64
140 372	381616600344013129894862647690541247659597966711193 60
144 368	17440779964066130420170168634612348393067326120770585 60
148 364	6832093649302361264113692814929677508406436591321481 2160
152 360	22997442048004658024533166375957838291089128020282067517 44
156 356	66674424868716978552789375387240003239187186349775851094016
160 352	16685597009641605873508056645831229244989283581517157330074 08
164 348	36117082274027891545154187373048131661136552390031364702863360
168 344	67748359898954710779361510124773951426962118474135604146110464 0
172 340	11032441933713096201663286389373184730113421621201515757397082112
176 336	15622509549761981330767923193778086142683556715677647652508417766 4
180 332	19266675322170971615767029917766543442504401756881968874572795084 80
184 328	20723534026876536792281002394151796205045793736436788802938336133120
188 324	19467144274183785293997555336377185623484125923840436555628706529280 0
192 320	15990449901813409988192707661615966056925120850571707914776940752826 32
196 316	11498415685246302189888474222781442491860129957714864173250891967627264
200 312	724594675707436038193788127187724975408707704846264948389597262678097 92
204 308	40054993226393655422034298725822449978056412171282746567439522386149376 0
208 304	19440716119784239090594261981448498630646086750443974295489951775717324 80
212 300	82912118532783785444361572212137368354501088010426952045243530869735424 00
216 296	31095502600701130763682713427899390240950550846409105550583369693522427904
220 292	10262265243551021935495943795989790043448061584592614216685442619215865446 4
224 288	29820628130211072662300075044545013251288181062960712347847355409523781096 0
228 284	76339691963166668867675510699680388300388184743872831189110938463079759872 0
232 280	17224527761762198963574524869345731758046653437351694799190878995825516871 68
236 276	34267504602573059044705476415066421758676994653154784033541236313665086423 04
240 272	60131635994896839993127999354917771797727242479988779533784429205014179338 24
244 268	93095513202488540513326927728892454124955629888945474125328180450571164057 60
248 264	12718986044129514620716674156341900030463015021774940408815989741288144568320
252 260	15336997499945305387056357527918950456934399969250231086077675815418680311808
256	16324199909251682000435577287934368523097397692548071777837483832108326674502

References

1. C. Carlet, *Boolean Functions for Cryptography and Coding Theory*, Cambridge University Press, Cambridge, 2021.
2. C. Carlet and P. Solé, "The weight spectrum of two families of Reed-Muller codes", Discrete Mathematics, **346(10)**, 113568, 2023.
3. P. Charpin, "Open problems on cyclic codes", in *Handbook of Coding Theory*, vol. 1, V. S. Pless, C. W. Huffman, editors, R. A. Brualdi, assistant editor, Elsevier, 963-1063, 1998.
4. Th. W. Cusick and P. Stănică, *Cryptographic Boolean functions and Applications*, Academic Press, Amsterdam,..., Tokyo, 2009.
5. V. Gillot and Ph. Langevin, "Classification of some cosets of the Reed-Muller code", Cryptogr. Commun. (2023), available at https://doi.org/10.1007/s12095-023-00652-4.
6. A. Hulpke, "Computing with group orbits", available at https://www.math.colostate.edu/
7. X. -D. Hou, "$GL(m,2)$ acting on $\mathcal{R}(r,m)/\mathcal{R}(r-1,m)$", Discrete Mathematics, **149**, 99-122, 1996.
8. T. Kasami and N. Tokura, "On the weight structure of Reed-Muller codes", IEEE Trans. Info. Theory, 16 , 752-759, 1970.
9. T. Kasami, N. Tokura, S. Azumi, "On the weight enumeration of weights less than 2.5d of Reed-Muller codes", Information and Control, 30, 380-395, 1976.
10. Ph. Langevin, "Classification of Boolean quartic forms in eight variables", available at https://langevin.univ-tln.fr/project/quartics/quartics.html, 2007.
11. Ph. Langevin, "Classification of $RM(4,7)/RM(2,7)$", available at https://langevin.univ-tln.fr/project/rm742/rm742.html, 2012.
12. Ph. Langevin and G. Leander, "Classification of Boolean quartic forms in eight variables", in *Boolean Functions in Cryptology and Information Security*, B. Preneel and O. A. Logachev (Eds.), IOS Press, 139-147, 2008.
13. F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland Publishing Company, Amsterdam, New York, Oxford, 1977.
14. Q. Meng et al., "Analysis of affinely equivalent Boolean functions", Science in China Series F: Information Sciences., vol. 50, no. 3, 299-306, 2007.
15. D. V. Sarwate, *Weight Enumeration of Reed-Muller Codes and Cosets*, Ph.D., Dep. Elec. Eng., Princeton Univ., Princeton, N.J., Sept. 1973, Advisors: E. R. Berlekamp and J. D. Ullman.
16. N. J. A. Sloane, "On-line Encyclopedia of Integer Sequences". available at https://oeis.org/wiki/List of weight distributions
17. N. J. A. Sloane, E. R. Berlekamp, "Weight enumerator for second-order Reed-Muller codes", IEEE Trans. Info. Theory, 16, 745-751, 1970.
18. Ts. Sugita, T. Kasami, and T. Fujiwara, "The weight distribution of the third-order Reed-Muller code of length 512", IEEE Trans. Info. Theory, 42, No. 5, 1622-1625, 1996.
19. M. Sugino, Y. Tokura and T. Kasami, "Weight distribution of (128,64) Reed-Muller Code", IEEE Trans. Info. Theory, 17, 627-628, 1971.
20. H. C. A. van Tilborg, "Weights in the third-order Reed-Muller codes", JPL Technical Report 32-1526, vol.IV, 86-92, 1971.

Iterative decoding of skew constacyclic codes

E.K. Nouetowa and I. Pogildiakov

Univ Rennes, CNRS, IRMAR - UMR 6625, Rennes Cedex, France
kayode-epiphane.nouetowa@univ-rennes.fr, ivan.pogildiakov@gmail.com

Abstract. The aim of this note is to design an iterative decoding algorithm for skew constacyclic codes defined over finite fields, which is inspired from [2] and [3]. We analyse the algorithm in the single error case, and use computer simulations in the general one.

Acknowledgments.

This work was conducted within the the France 2030 framework porgramme, Centre Henri Lebesgue ANR−11−LABX−0020 − 01.

1 Introduction

Skew cyclic codes are a subclass of linear codes containing the cyclic codes. These codes and their decoding algorithms have been the subject of several works [4,8,10]. Recently, M. Bossert proposed an iterative decoding algorithm for binary cyclic codes [2] using the minimum weight codewords of "dual" codes. Later, M. Bossert et al. extended that work to non-binary cyclic codes [3]. The aim of this note is to adapt these algorithms to skew constacyclic codes using Euclidean duals.

The text is organized as follows. In Section 2, we recall the definition of skew constacyclic codes and a characterization of their Euclidean duals. In Section 3, we give our decoding strategy and show the link with the strategy applied in [3]. In Section 4, we initialize an analysis of the resulting iterative decoding algorithm and provide a condition under which the algorithm always fails.

2 Some generalities on skew constacyclic codes

Recall that a **linear code** $\mathcal{C}$ over a finite field $\mathbb{F}_q$ of length n and dimension k is a k-dimensional subspace of $\mathbb{F}_q^n$. The **Euclidean dual** $\mathcal{C}^\perp$ of $\mathcal{C}$ is defined as

$$\mathcal{C}^\perp = \{v \in \mathbb{F}_q^n \mid \langle v, c\rangle = 0 \text{ for all } c \in \mathcal{C}\},$$

where $\langle x, y\rangle = \sum_{i=0}^{n-1} x_i y_i$ is the Euclidean scalar product of $x = (x_0, \ldots, x_{n-1}) \in \mathbb{F}_q^n$ and $y = (y_0, \ldots, y_{n-1}) \in \mathbb{F}_q^n$. The **minimum distance** d of $\mathcal{C}$ is the smallest

of the (Hamming) weights of the non-zero codewords. Furthermore, a linear code $\mathcal{C}$ is called cyclic if for all $c = (c_0, \ldots, c_{n-1})$ in $\mathcal{C}$ the vector $(c_{n-1}, c_0, \ldots, c_{n-2})$ also belongs to $\mathcal{C}$.

Let θ be an automorphism of $\mathbb{F}_q$, and let ε be a non-zero element of $\mathbb{F}_q$. Consider the map

$$\phi_\varepsilon : \begin{cases} \mathbb{F}_q^n & \longrightarrow \mathbb{F}_q^n, \\ (a_0, ..., a_{n-1}) & \longmapsto (\varepsilon\theta(a_{n-1}), \theta(a_0), \ldots, \theta(a_{n-2})). \end{cases}$$

Skew constacyclic codes are defined as follows.

Definition 1 (Definition 1 of [6]). *A (θ, ε)-**constacyclic code** $\mathcal{C}$ is a linear code over $\mathbb{F}_q$ such that for any $c = (c_0, \ldots, c_{n-1}) \in \mathbb{F}_q^n$,*

$$c \in \mathcal{C} \Rightarrow \phi_\varepsilon(c) \in \mathcal{C}.$$

*If $\varepsilon = 1$, then $\mathcal{C}$ is called θ-**cyclic**. If $\varepsilon = -1$, then $\mathcal{C}$ is called θ-**negacyclic**.*

Note that if θ is the identity, then a θ-cyclic code (resp. θ-negacyclic) is a cyclic (resp. negacyclic) code.

The skew polynomial ring $(R, +, \cdot)$, or Ore ring [11], is the set of polynomials $\mathbb{F}_q[x; \theta]$ over $\mathbb{F}_q$ equipped with the usual component-wise addition '+' and where the multiplication '$\cdot$' is defined by the rule

$$x \cdot a = \theta(a)x \text{ for all } a \in \mathbb{F}_q.$$

Clearly, the ring R is non-commutative if θ is different from the identity. It is well known that R is a left and right Euclidean ring. Moreover, the center of R is the commutative polynomial ring $\mathbb{F}_q^\theta[x^{|\theta|}]$, where $\mathbb{F}_q^\theta$ stands for the subfield of $\mathbb{F}_q$ fixed by θ, and $|\theta|$ is the order of θ.

In this text we use the conventional representation of the elements $c = (c_0, \ldots, c_{n-1})$ of $\mathbb{F}_q^n$ as skew polynomials $c(x) = c_0 + c_1 x + \ldots + c_{n-1}x^{n-1}$ of degree less than n. Under this correspondence, a (θ, ε)-constacyclic code $\mathcal{C}$ can be viewed as a left R-submodule $Rg(x)/R(x^n - \varepsilon)$ of $R/R(x^n - \varepsilon)$ [6], where $g(x)$ is a monic skew right divisor of $x^n - \varepsilon$ called **skew generator polynomial** of the code $\mathcal{C}$. The dimension of $\mathcal{C}$ is $k = n - \deg(g(x))$. One has

$$\mathcal{C} = \{m(x)g(x) \mid m(x) \in R, \deg(m(x)) < k\}.$$

We write $\mathcal{C} = (g)_{n,\theta}^{\varepsilon}$. To characterize the Euclidean dual of $\mathcal{C}$ we recall below the definition of the skew reciprocal polynomial of a skew polynomial in R.

Definition 2 (Definition 3 of [6]). *Let $h = \sum_{i=0}^{k} h_i x^i \in R$ be a skew polynomial of degree k. The **skew reciprocal polynomial** $h^*(x)$ of $h(x)$ is the skew polynomial*

$$h^*(x) = \sum_{i=0}^{k} x^{k-i} \cdot h_i = \sum_{i=0}^{k} \theta^i(h_{k-i})x^i.$$

The **left monic skew reciprocal polynomial** $h^\natural(x)$ *of* $h(x)$ *is*

$$h^\natural(x) = \frac{1}{\theta^{k-m}(h_m)} h^*(x),$$

where $m = \min\{i \mid h_i \neq 0\}$.

The following proposition gives the Euclidean dual of a (θ, ε)-constacyclic code. Notice that we extend the automorphism θ of $\mathbb{F}_q$ to the automorphism $\theta\colon R \to R$ of R by linearity, i.e. by the rule $\sum_{i=0}^{k} a_i x^i \mapsto \sum_{i=0}^{k} \theta(a_i) x^i$.

Proposition 1 (Theorem 1 of [6]). *Let* $\mathcal{C} = (g)^{\varepsilon}_{n,\theta}$ *be a* (θ, ε)*-constacyclic code over* $\mathbb{F}_q$*. The Euclidean dual* $\mathcal{C}^\perp$ *of* $\mathcal{C}$ *is a* $(\theta, 1/\varepsilon)$*-constacyclic code over* $\mathbb{F}_q$ *defined as* $\mathcal{C}^\perp = (h^\natural)^{1/\varepsilon}_{n,\theta}$*, where* h *is the monic skew polynomial such that*

$$x^n - \varepsilon = \theta^n(h(x))g(x).$$

The skew polynomial $h(x)$ *is called* **skew check polynomial** *of the code* $\mathcal{C}$*.*

Example 1 *Let* $\mathbb{F}_4 = \mathbb{F}_2(a)$ *with* $a^2 + a + 1 = 0$*,* $\theta\colon u \mapsto u^2 \in Aut(\mathbb{F}_4)$*, and* $R = \mathbb{F}_4[x; \theta]$*. Let* $n = 12$*. Consider* $g(x) = x^8 + a^2x^6 + x^5 + x^4 + x^3 + x + a \in R$*. One has* $x^{12} - 1 = h(x)g(x)$*, where* $h(x) = x^4 + a^2x^2 + x + a^2$*. The skew polynomial* $g(x)$ *generates a* θ*-cyclic code* $\mathcal{C}$ *of length* 12*, dimension* 4 *and minimum distance* 7 *over* $\mathbb{F}_4$*.*

We have $h^*(x) = a^2x^4 + x^3 + a^2x^2 + 1$ *and* $h^\natural(x) = x^4 + ax^3 + x^2 + a$*. Therefore, the dual* $\mathcal{C}^\perp$ *of* $\mathcal{C}$ *is the* θ*-cyclic code over* $\mathbb{F}_4$ *with skew generator polynomial* $h^\natural(x)$ *having length* 12*, dimension* 8*, and minimum distance* 4*.*

The following technical lemma finds its use in next section.

Lemma 1 (Lemma 1 of [6] and Lemma 7 of [5]). *Let* $f(x)$ *and* $g(x)$ *be skew polynomials in* R*, and let* $k = \deg(f)$*. One has*

1. $(f(x)g(x))^* = \theta^k(g(x)^*)f(x)^*$.
2. *If the constant coefficient of* f *is nonzero, then* $(f(x)^*)^* = \theta^k(f(x))$.
3. *If* $f(x)g(x)$ *is central, then* $f(x)g(x) = g(x)f(x)$.

3 Decoding strategy

Let $\mathcal{C}$ be a (θ, ε)-constacyclic code over $\mathbb{F}_q$ of length n with skew generator polynomial $g(x)$ and skew check polynomial $h(x)$. In this section, we assume that ε belongs to the field $\mathbb{F}_q^\theta$ fixed by θ and that the order $|\theta|$ of θ divides the length n of the code $\mathcal{C}$. Therefore, $x^n - \varepsilon$ is a central polynomial, $R/(x^n - \varepsilon)$ is a left principal ideal ring and $\mathcal{C}$ is a left ideal $(g(x))/(x^n - \varepsilon)$. Furthermore, according to Point 3 of Lemma 1, we have

$$x^n - \varepsilon = h(x)g(x) = g(x)h(x). \tag{1}$$

Lemma 2. *Given $u \in \mathcal{C}^\perp$ and $c \in \mathcal{C}$, the following holds*:

$$c(x)\theta^{-\ell}(u(x)^*) \equiv 0 \pmod{x^n - \varepsilon},$$

where $\ell = \deg(u(x))$.

Proof. Let $c \in \mathcal{C} = (g)^{\varepsilon}_{n,\theta}$ and $u \in \mathcal{C}^\perp$. Consider $m(x)$ and $v(x)$ in R such that $c(x) = m(x)g(x)$ and $u(x) = v(x)h(x)^*$. According to Lemma 1, we have

$$\begin{aligned} u(x)^* &= \theta^{\deg(v(x))}((h(x)^*)^*)v(x)^* && \text{(Point 1. of Lemma 1)} \\ &= \theta^{\deg(v(x))+\deg(h(x))}\left(h(x)\right)v(x)^* && \text{(Point 2. of Lemma 1)} \\ &= \theta^{\ell}\left(h(x)\right)v(x)^* && \text{(because } \deg(h^*(x)) = \deg(h(x))\text{)}. \end{aligned}$$

Therefore, we have $\theta^{-\ell}\left(u(x)^*\right) = h(x)\theta^{-\ell}\left(v(x)^*\right)$ and

$$\begin{aligned} c(x)\theta^{-\ell}(u(x)^*) &= m(x)g(x)h(x)\theta^{-\ell}(v(x)^*) \\ &= m(x)(x^n - \varepsilon)\theta^{-\ell}(v(x)^*) && \text{(according to (1))} \\ &= m(x)\theta^{-\ell}(v(x)^*)(x^n - \varepsilon) && \text{(because } x^n - \varepsilon \text{ is central)} \\ &\equiv 0 \pmod{x^n - \varepsilon}. \end{aligned}$$

■

Remark 1. The polynomials $u(x)$ and $\theta^{-\ell}(u(x)^*)$ have the same Hamming weight.

Definition 3. *Two non-zero words c_1 and $c_2 \in \mathbb{F}_q^n$ are called (θ, ε)-cyclically equivalent if there exist b in $\mathbb{F}_q$ and i in $\mathbb{N}$ such that $c_2 = b\phi_\varepsilon^i(c_1)$ which means that $c_2(x) = bx^i c_1(x) \bmod (x^n - \varepsilon)$. In this case we write $c_1 \sim_{\theta,\varepsilon} c_2$.*

One can show that $\sim_{\theta,\varepsilon}$ is an equivalence relation. Each class of (θ, ε)-cyclically equivalent words in $\mathbb{F}_q^n$ contains a monic representative (considered as a polynomial).

Now we pick the following two sets of words playing an important role in the decoding strategy. Let w be a positive integer. We define

- $\mathcal{B}_w = \{\text{all monic representatives in } \mathcal{C}^\perp/\sim_{\theta,\frac{1}{\varepsilon}} \text{ of Hamming weight } w\}$,
- $\overline{\mathcal{B}}_w = \{\theta^{-\ell}(u(x)^*) \mid u \in \mathcal{B}_w \text{ and } \ell = \deg(u(x))\}$.

Remark 2. Clearly, the set $\mathcal{B}_w$ is not unique. But, however, one can show that the choice of one or the other has no influence on the construction of the frequency matrix that we detail in what follow.

The elements of $\mathcal{B}_w$ are monic skew polynomials. Therefore, the skew reciprocal polynomial of each element of $\mathcal{B}_w$ has constant coefficient one. According to Remark 1, each element of $\overline{\mathcal{B}}_w$ has Hamming weight w. Thus, the elements of $\overline{\mathcal{B}}_w$ are of the form

$$1 + \lambda_{\beta_1}x^{\beta_1} + \cdots + \lambda_{\beta_{w-1}}x^{\beta_{w-1}} \in \mathbb{F}_q[x;\theta],$$

where $\beta_1, \dots \beta_w$ are distinct elements of $\{1, \dots, n-1\}$, and $\lambda_{\beta_1}, \dots, \lambda_{\beta_{w-1}}$ are non-zero elements of $\mathbb{F}_q$.

Unless otherwise stated, in the following we denote by $y = c + e$ a received word, where $c \in \mathcal{C}$ is a codeword, $e \in \mathbb{F}_q^n$ is an error of Hamming weight τ at most the half distance bound.

Remark 3. If $u \in \mathcal{C}^\perp$, i.e. $u(x) = v(x)h(x)^*$ with $\deg(v(x)) < n - \deg(h(x))$, then according to Lemma 1(see also proof of Lemma 2) $\theta^{-\ell}(u(x)^*) = h(x)\theta^{-\ell}(v(x)^*)$. If θ is the identity and $\varepsilon = 1$, then, for u in $\mathcal{C}^\perp$, $\theta^{-\ell}(u^*)$ is in fact a codeword of the cyclic code of length n and of generator polynomial $h(x)$. This code is called "dual" code in [2,3], and we would like to emphasis that this "dual" code is not the Euclidean dual $\mathcal{C}^\perp$, unless $h(x)$ is self-reciprocal.

Proposition 2. *Consider $f(x) \in \overline{\mathcal{B}}_w$, we have*

$$y(x)f(x) \equiv e(x)f(x) \pmod{x^n - \varepsilon}.$$

Proof. Lemma 2 implies $c(x)f(x) \equiv 0 \pmod{x^n - \varepsilon}$. Therefore, as $y(x) = c(x) + e(x)$, we have $y(x)f(x) = c(x)f(x) + e(x)f(x) \equiv e(x)f(x) \pmod{x^n - \varepsilon}$. ∎

Let $d^\perp$ be the minimum distance of $\mathcal{C}^\perp$. We want to concentrate our attention on the sets $\mathcal{B}_{d^\perp}$ and $\overline{\mathcal{B}}_{d^\perp}$.

Example 2 (Example 1, continued) *We have $d^\perp = 4$. Using* Magma *[1], one finds: $\overline{\mathcal{B}}_{d^\perp} = \{1 + x^3 + x^6 + x^9, 1 + a^2x + x^2 + ax^4, 1 + a^2x^2 + a^2x^4 + x^9, 1 + x + a^2x^3 + ax^8, 1 + x^4 + x^6 + x^{10}, 1 + a^2x^3 + ax^4 + x^5, 1 + x + x^6 + x^7, 1 + ax + x^4 + ax^9\}$.*

Let $f(x) = 1 + \lambda_{\beta_1}x^{\beta_1} + \cdots + \lambda_{\beta_{d^\perp-1}}x^{\beta_{d^\perp-1}}$ be an element of $\overline{\mathcal{B}}_{d^\perp}$. Consider the following skew polynomial in R:

$$\begin{aligned}\omega_f^0(x) &= y(x)f(x) \bmod (x^n - \varepsilon)\\ &= e(x)f(x) \bmod (x^n - \varepsilon)\\ &= e(x) + e(x)\lambda_{\beta_1}x^{\beta_1} + \cdots + e(x)\lambda_{\beta_{d^\perp-1}}x^{\beta_{d^\perp-1}} \bmod (x^n - \varepsilon).\end{aligned}$$

Note that $\omega_f^0(x)$ is the sum of the error $e(x)$ and its shifts by the skew monomials $\lambda_{\beta_i}x^{\lambda_{\beta_i}}$, $i \in \{1, \dots, d^\perp - 1\}$. Therefore, the degree of each monomial in $\omega_f^0(x)$ is an error position or the sum of an error position and a non-zero β_i. Given $i \in \{1, \dots, d^\perp - 1\}$, we define $\omega_f^i(x)$ as follows:

$$\omega_f^i(x) = \omega_f^0(x)\varepsilon^{-1}\theta^{-\beta_i}\left(\lambda_{\beta_i}^{-1}\right)x^{n-\beta_i} \bmod (x^n - \varepsilon), \tag{2}$$

where $\varepsilon^{-1}\theta^{-\beta_i}\left(\lambda_{\beta_i}^{-1}\right)x^{n-\beta_i}$ is the inverse of the skew monomial $\lambda_{\beta_i}x^{\beta_i}$ modulo $x^n - \varepsilon$. Notice that $\omega_f^i(x)$ is the sum of the error $e(x)$ and its shifts.

We need the following list of vectors:

$$\mathcal{L} = \left[\omega_f^i \mid f(x) \in \overline{\mathcal{B}}_{d^\perp}, i \in \{0, \dots, d^\perp - 1\}\right]. \tag{3}$$

Remark that there are exactly $|\overline{\mathcal{B}}| \times d^{\perp}$ elements in $\mathcal{L}$.

Now, we order the elements of the ground finite field $\mathbb{F}_q$ as $\{\sigma_0, \ldots, \sigma_{q-1}\}$, where $\sigma_0 = 0$. Let us build a $q \times n$ frequency matrix as follows:

$$\mathcal{T} = \begin{bmatrix} \mathcal{T}_{\sigma_0,0} & \mathcal{T}_{\sigma_0,1} & \ldots & \mathcal{T}_{\sigma_0,n-1} \\ \mathcal{T}_{\sigma_1,0} & \mathcal{T}_{\sigma_1,1} & \ldots & \mathcal{T}_{\sigma_1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{T}_{\sigma_{q-1},0} & \mathcal{T}_{\sigma_{q-1},1} & \ldots & \mathcal{T}_{\sigma_{q-1},n-1} \end{bmatrix},$$

where

$$\mathcal{T}_{\sigma_\mu,j} = \left|\{\omega_f^i \in \mathcal{L} \mid (\omega_f^i)_j = \sigma_\mu\}\right|. \tag{4}$$

Note that the rows of $\mathcal{T}$ are enumerated by the elements of $\mathbb{F}_q$. Given $\mu \in \{0, \ldots, q-1\}$ and $j \in \{0, \ldots, n-1\}$, the entry $\mathcal{T}_{\sigma_\mu,j}$ is the number of skew polynomials $\omega_f^i(x)$ in $\mathcal{L}$ having coefficient $\sigma_\mu \in \mathbb{F}_q$ at the position j. Withing our approach, we use $\mathcal{T}$ to take a decision on the error position and the occurrence at this position in the error vector.

The idea of the decoding strategy is based on the fact that the number of ω_f^i in $\mathcal{L}$ having a coefficient equal to $\sigma_0 = 0$ at the error positions is expected to be low (see Section 4 for more analysis). We define:

$$\nu = \min\{\mathcal{T}_{\sigma_0,j} \mid j \in \{0, \ldots, n-1\}\}, \quad \mathcal{P}_\nu = \{j \in \{0, \ldots, n-1\} \mid \mathcal{T}_{\sigma_0,j} = \nu\}. \tag{5}$$

The elements of $\mathcal{P}_\nu$ are taken as the possible error positions. The sum of the elements of each column of $\mathcal{T}$ is the same, i.e. $\sum_{\mu=0}^{q-1} \mathcal{T}_{\sigma_\mu,j} = |\mathcal{L}| = |\overline{\mathcal{B}}| \times d^{\perp}$. Given a position $j \in \mathcal{P}_\nu$, we determine the largest element

$$\mathcal{M}_j = \max\{\mathcal{T}_{\sigma_\mu,j} \mid \mu \in \{1, \ldots, q-1\}\} \tag{6}$$

of the column j. Therefore, we can delete from $y(x)$ the error magnitudes at the identified erroneous positions by replacing $y(x)$ with $y(x) - \sum_{j \in \mathcal{P}_\nu} \sigma_{\mu_j} x^j$, where $\mu_j \in \{1, \ldots, q-1\}$ are such that $\mathcal{T}_{\sigma_{\mu_j},j} = \mathcal{M}_j$. One checks if $y(x)$ is a code-word. Otherwise, one starts again by the calculation of another $\mathcal{T}$ with the new $y(x)$.

The decoding algorithm is summarized in Algorithm 1.

Let us finish this section by demonstrating the work of Algorithm 1 on a concrete example.

Example 3 (Example 1, continued) *Suppose that*

$$y(x) = x^{11} + x^{10} + ax^9 + ax^8 + a^2x^7 + a^2x^5 + a^2x^3 + ax^2 + 1.$$

One can verify that $y(x)$ is not a codeword of the code $\mathcal{C}$ by dividing $y(x)$ on the right by $g(x)$. Therefore, we are going to apply the decoding strategy to $y(x)$ hopping to recover a corrupted codeword.

Consider the following order in the base finite field $\mathbb{F}_4$:

$$\sigma_0 = 0, \sigma_1 = 1, \sigma_2 = a, \sigma_3 = a^2.$$

Algorithm 1 Decoding algorithm for $\mathcal{C} = (g)^{\varepsilon}_{n,\theta}$

Require: $\overline{\mathcal{B}}_{d^\perp}, y(x) = c(x) + e(x)$, where $c \in \mathcal{C}$; i_{max}
Ensure: $c(x)$ or Failure
1: $i = 0$;
2: **while** $y(x) \notin \mathcal{C}$ and $i \leq i_{max}$ **do**
3: Construct $\mathcal{L}$ defined by (3)
4: Construct $\mathcal{T}$ defined by (4)
5: Determine ν and $\mathcal{P}_\nu$ as in (5)
6: Refine $y(x) = y(x) - \sum_{j \in \mathcal{P}_\nu} \sigma_{\mu_j} x^j$
7: $i \leftarrow i + 1$;
8: **end while**
9: **if** $y(x) \in \mathcal{C}$ **then**
10: **return** $y(x)$
11: **else**
12: **return** Failure
13: **end if**

We put $y_0(x) = y(x)$. At the i-th iteration the decoding algorithm constructs a matrix $\mathcal{T} = \left(\mathcal{T}_{\sigma_\mu, j}\right)_{\mu,j}$ from the polynomial $y_{i-1}(x)$ and the list $\overline{\mathcal{B}}_{d^\perp}$.

Iteration 1. *One finds*

$$\mathcal{T} = \begin{bmatrix} 11 & 11 & 3 & 13 & 11 & 8 & 14 & \mathbf{2} & 10 & 14 & 12 & 3 \\ 7 & 5 & 17 & 9 & 9 & 6 & 6 & 4 & 8 & 6 & 6 & 5 \\ 7 & 11 & 7 & 5 & 7 & 8 & 6 & \boxed{18} & 6 & 6 & 8 & 7 \\ 7 & 5 & 5 & 5 & 5 & 10 & 6 & 8 & 8 & 6 & 6 & 17 \end{bmatrix},$$

$$\nu = 2, \quad \mathcal{P}_\nu = 7, \quad \mathcal{M}_7 = 18, \quad \sigma_{\mu_7} = a.$$

The algorithm takes the following desicion: most probably there is an error in $y_0(x)$ at position 7 and most probably the corresponding entry in the error vector at position 7 is $\sigma_2 = a$. Therefore, we put

$$y_1(x) = y_0(x) - ax^7 = x^{11} + x^{10} + ax^9 + ax^8 + x^7 + a^2x^5 + a^2x^3 + ax^2 + 1.$$

One can verify that $y_1(x)$ does not belong to $\mathcal{C}$, and therefore we proceed with the algorithm.

Iteration 2. *One computes*

$$\mathcal{T} = \begin{bmatrix} 16 & 19 & \mathbf{2} & 18 & 16 & 14 & 19 & 18 & 14 & 19 & 19 & \mathbf{2} \\ 6 & 5 & \boxed{22} & 8 & 6 & 2 & 5 & 8 & 10 & 5 & 5 & 2 \\ 4 & 3 & 6 & 2 & 4 & 6 & 3 & 2 & 6 & 3 & 3 & 6 \\ 6 & 5 & 2 & 4 & 6 & 10 & 5 & 4 & 2 & 5 & 5 & \boxed{22} \end{bmatrix},$$

$$\nu = 2, \quad \mathcal{P}_\nu = \{2, 11\}, \quad \mathcal{M}_2 = \mathcal{M}_{11} = 22, \quad \sigma_{\mu_2} = 1 \text{ and } \sigma_{\mu_{11}} = a^2.$$

Note that the minimum value of the first row of $\mathcal{T}$ occurs several times in the row. Therefore, the algorithm decides that most probably there are errors in $y_1(x)$ at positions 2 and 11, and most probably the corresponding entries in the error vector are $\sigma_1 = 1$ and $\sigma_3 = a^2$. Hence we put

$$y_2(x) = y_1(x) - (a^2x^{11} + x^2) = ax^{11} + x^{10} + ax^9 + ax^8 + x^7 + a^2x^5 + a^2x^3 + a^2x^2 + 1.$$

One verifies that $y_2(x)$ is an element of the code $\mathcal{C}$. Thus, the decoding is successfully done and we have just managed to find out that $y(x) = c(x) + e(x)$, where

$$c(x) = ax^{11} + x^{10} + ax^9 + ax^8 + x^7 + a^2x^5 + a^2x^3 + a^2x^2 + 1 \in \mathcal{C},$$
$$e(x) = a^2x^{11} + ax^7 + x^2.$$

Let us also notice that there were only three errors, which coincides with the error capacity of the code $\mathcal{C}$, and the recovered message $c(x)$ is unique.

4 Plausibility analysis of Algorithm 1

A plausibility analysis of an iterative decoding algorithm for binary cyclic codes was given in [2]. It was improved and completed for non-binary cyclic codes in [9]. This analysis (left column of Page 655) can be adapted to our situation.

In what follows we present a conjecture (Conjecture 1) on the failure of the algorithm. This conjecture is motivated by Example 4 and proved when the error weight is one (Lemma 3).

Let us first introduce a new set. Given $f(x) = 1 + \lambda_{\beta_1}x^{\beta_1} + \ldots + \lambda_{\beta_{d^\perp-1}}x^{\beta_{d^\perp-1}}$ in R, the support S_f^0 of $f(x)$ is

$$S_f^0 := \{0, \beta_1, \ldots, \beta_{d^\perp-1}\},$$

and for $0 < i < d^\perp$ we denote by S_f^i the support of $f(x)\varepsilon^{-1}\theta^{-\beta_i}\left(\lambda_{\beta_i}^{-1}\right)x^{n-\beta_i} \bmod (x^n - \varepsilon)$:

$$S_f^i := \{(v - \beta_i) \bmod n,\ v \in S_f^0\}.$$

Consider the intersection $\mathcal{I}$ of the supports S_f^i:

$$\mathcal{I} = \bigcap_{i=0, f\in\overline{\mathcal{B}}_{d^\perp}}^{d^\perp-1} S_f^i.$$

We have the following conjecture.

Conjecture 1. If $\mathcal{I} \neq \{0\}$, then the set $\mathcal{P}_\nu$ constructed at the first stage of the iterative decoding Algorithm 1 contains non-erroneous positions, and, therefore, the algorithm returns Failure.

This means that the decoding can not be done with the dual codewords of weight $d^\perp$. The following lemma gives a proof of Conjecture 1 when the error weight is equal to 1.

Lemma 3. *Conjecture 1 is true if the error weight is equal to* 1.

PROOF. Assume that $e(x) = e_{\gamma_1} x^{\gamma_1}$. Each $\omega_f^i \in \mathcal{L}$ is of the form

$$\omega_f^i(x) = e_{\gamma_1} x^{\gamma_1} + e_{\gamma_1} \theta^{\gamma_1}(\lambda_{\beta_1}) x^{\gamma_1+\beta_1} + \cdots + e_{\gamma_1} \theta^{\gamma_1}(\lambda_{\beta_{d^\perp-1}}) x^{\gamma_1+\beta_{d^\perp-1}} \bmod x^n - \varepsilon$$

for all $f(x) \in \overline{\mathcal{B}}_{d^\perp}$ and $i \in \{0, \ldots, d^\perp - 1\}$. Therefore, at the position γ_1 of the vector ω_f^i, we have the error magnitude e_{γ_1}. According to (5), $\nu = \mathcal{T}_{\sigma_0,\gamma_1}$ is zero. Algorithm 1 succeeds in this case if and only if $\mathcal{P}_\nu = \{\gamma_1\}$. This is equivalent to $\mathcal{I} = \{0\}$. Namely, one has the following equivalences:

$$\begin{aligned}
\mathcal{P}_\nu \neq \{\gamma_1\} &\Leftrightarrow \exists \beta \in \{1, \ldots, n-1\},\ (\beta + \gamma_1) \bmod n \in \mathcal{P}_\nu \\
&\Leftrightarrow \exists \beta \in \{1, \ldots, n-1\}, \forall i, f,\ (\beta + \gamma_1) \bmod n \in Supp(\omega_f^i(x)) \\
&\Leftrightarrow \exists \beta \in \{1, \ldots, n-1\}, \forall i, f,\ \beta \in S_f^i \\
&\Leftrightarrow \{1, \ldots, n-1\} \cap \mathcal{I} \neq \emptyset \\
&\Leftrightarrow \mathcal{I} \neq \{0\}.
\end{aligned}$$

■

The following example illustrates Conjecture 1 in the case when Algorithm 1 fails for sample of errors of weight bigger than 1.

Example 4 *Consider the* $[54, 19, 21]_9$ *θ-cyclic code* $\mathcal{C} = (g)_{54,\theta}$ *defined over* $\mathbb{F}_9 = \mathbb{F}_3(a)$, *where* $a^2 = a+1$, *and* $g(x) = a^2x^{35} + a^7x^{34} + x^{33} + a^7x^{32} + a^3x^{31} + 2x^{30} + a^2x^{28} + a^7x^{27} + a^7x^{26} + a^2x^{24} + ax^{23} + a^5x^{22} + a^3x^{21} + x^{20} + a^2x^{19} + a^2x^{18} + ax^{17} + a^5x^{15} + ax^{12} + a^3x^{11} + x^{10} + a^7x^9 + a^3x^7 + x^6 + a^2x^5 + a^6x^4 + a^2x^2 + ax + 1$.

The skew check polynomial of $\mathcal{C}$ *is* $h(x) = a^2x^{19} + ax^{18} + a^5x^{17} + a^5x^{15} + a^3x^{14} + a^7x^{13} + ax^{12} + a^7x^{11} + 2x^{10} + a^3x^9 + x^8 + ax^7 + ax^6 + a^2x^5 + a^2x^4 + a^5x^3 + a^7x^2 + ax + 2$.

The dual code of $\mathcal{C}$ *is a* $[54, 35, 6]_9$ *θ-cyclic code with skew generator polynomial* $h^*(x)$. *We have* $x^{54} - 1 = (x^2 - 1)^{27} = g(x)h(x) = h(x)g(x)$ *and one can check that* $h^*(x)$ *divides* $(x^2 - 1)^{18}(x + 1) = (x^{36} + x^{18} + 1)(x + 1)$.

The set $\overline{\mathcal{B}}_6$ *is obtained by considering all the multiples of* $(x^{36}+x^{18}+1)(x+1)$ *of weight* 6*:*

$$\begin{aligned}
\overline{\mathcal{B}}_6 = \{(x^{36} + x^{18} + 1)u \mid u \in \{&x + 1, 2x^2 + 1, x^3 + 1, 2x^4 + 1, \\
&2x^6 + 1, x^7 + 1, 2x^8 + 1, x^9 + 1\}\}.
\end{aligned}$$

The set $\mathcal{I}$ *is therefore equal to* $\{0, 18, 36\}$. *We considered a few of hundreds of thousands of samples, and we found no error vectors of weights up to* 14, *which Algorithm* 1 *can correct.*

For this reason, we opted to work with $\overline{\mathcal{B}}_{13}$ *(cf. the table below).*

In order to overcome the fact that $\mathcal{I}$ may be distinct from $\{0\}$, one chooses to replace $\overline{\mathcal{B}}_{d^\perp}$ with $\overline{\mathcal{B}}_w$, $w \geq d^\perp$, in the entry of Algorithm 1. We have implemented the algorithm in C and present some preliminary computational results in the table bellow.

Table 1. Results of computer simulations of Algorithm 1.

Codes	$[54, 27, 18]_9$		$[54, 19, 21]_9$		$[62, 26, 19]_4$	
Duals	$[54, 27, 18]_9$		$[54, 35, 6]_9$		$[62, 36, 13]_4$	
$d^\perp, w$	$18, 18$		$6, 13$		$13, 13$	
Success rate	$\tau \leq 7$	1	$\tau \leq 12$	1	$\tau \leq 7$	1
	$\tau = 8$	$0,9923$	$\tau = 13$	$0,9999$	$\tau = 8$	$0,9878$
	$\tau = 9$	$0,7532$	$\tau = 14$	$0,9918$	$\tau = 9$	$0,8346$

5 Conclusion

In this text we provide a generalization of the iterative decoding of [2] to the class of skew constacyclic codes that are ideal codes. We have initiated a preliminary analysis of our algorithm, and we aim at providing a more accurate analysis of its success rate. We implemented an improved version of Algorithm 1 both in *Magma* and in *C*, and conducted multiple experiments on several skew constacyclic codes over small finite fields. In the sequel, we would like also to compare our algorithm to other decoding algorithms (designed for skew BCH codes, for example).

References

1. W. Bosma, J. Cannon, C. Playoust: The Magma algebra system. I. The user language. Journal of Symbolic Computation, **24**, 3-4, 235–265, 1997 .
2. M. Bossert: On decoding using codewords of the dual code. arXiv preprint, 2001.02956, Jan. 2020.
3. J. Xing, M. Bossert, S. Bitzer, L. Chen: Iterative decoding of non-binary cyclic codes using minimum-weight dual codewords. in Proc. IEEE Int. Symp. Inform. Theory (ISIT), CA, U.S.A, pp. 333-337, June 2020.
4. D. Boucher, W. Geiselmann, F. Ulmer: Skew-cyclic codes. Appl. Algebra Engin. Commun. Comp., **18**, 379–389, 2007.
5. D. Boucher, F. Ulmer: Coding with skew polynomial rings. J. Symb. Comp., **44**, 1644–1656, 2009.
6. D. Boucher, F. Ulmer: A note on the dual codes of module skew codes. Liqun Chen. Cryptography and coding: 13th IMA international conference, IMACC 2011, Oxford, UK, December 12-15, 2011.
7. D. Boucher, F. Ulmer: Self-dual skew codes and factorization of skew polynomials. J. Symb. Comp., **60**, 47–61, 2014.
8. L. Chaussade, P. Loidreau, F. Ulmer: Skew codes of prescribed distance or rank. Designs, Codes and Cryptography, 50 (3), pp.267-284, 2009.
9. L. Chen, J. Xing, J. Yuan: Plausibility analysis of Shift-Sum decoding for cyclic codes 2021 IEEE International Symposium on Information Theory (ISIT), Melbourne, Australia, pp. 652-657, 2021.
10. J. Gómez-Torrecillas, F. J. Lobillo, G. Navarro: Peterson-Gorenstein-Zierler algorithm for skew RS codes. Linear and Multilinear Algebra, 66, 3, 469–487, 2018.
11. O. Ore: Theory of Non-commutative Polynomials, The annals of Mathematics, 2nd Ser, 34(3), 480-508, 1933.

Introducing locality in some generalized AG codes

Bastien Pacifico[1]

LIRMM, Université de Montpellier
Montpellier, France
bastien.pacifico@gmail.com

Abstract. In 1999, Xing, Niederreiter and Lam introduced a generalization of AG codes using the evaluation at non-rational places of a function field. In this paper, we show that one can obtain a locality parameter r in such codes by using only non-rational places of degree at most r. This is, up to the author's knowledge, a new way to construct locally recoverable codes (LRCs). We give an example of such a code reaching the Singleton-like bound for LRCs. We then investigate similarities with certain concatenated codes. Contrary to previous methods, our construction allows one to obtain directly codes whose dimension is not a multiple of the locality. Finally, we give an asymptotic study using the Garcia–Stichtenoth tower of function fields, for both our construction and a construction of concatenated codes. We give explicit infinite families of LRCs with locality 2 over any finite field of cardinality greater than 3 following our new approach.

1 Introduction

Locally Recoverable Codes (LRCs) are a popular topic lately, in particular for their potential applications in distributed storage [12]. The locality consists in the possibility of recovering one corrupted symbol using a small amount of other symbols. More precisely, a code is said to have locality r if any symbol of a codeword can be obtained using at most r other symbols [15, 12]. It was proven in [12] that an $[n, k, d]$ linear code with locality r verifies

$$d \leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2. \tag{1}$$

A LRC is said to be optimal when the equality is reached in this bound. There exist several constructions of optimal LRCs. The first ones were given in [15, 19, 22], but they required to use an alphabet of exponential size compared to the code length. The construction of Tamo and Barg [5] provides optimal codes of length upper bounded by the size of the alphabet and moreover with constraints on the locality due to the existence of good polynomials. There also exist codes that reach the bound (1) and have length greater than the size of the alphabet. Such constructions can be obtained using for instance algebraic surfaces [4, 18]. In

fact, it has been proven in [14] that, for minimum distance $d \geq 5$, the length of an optimal LRC is at most $\mathcal{O}(dq^3)$, where q is the size of the alphabet. On the other side, a classical problem in coding theory is to study what can be obtained for a fixed alphabet size. Fundamental works on this topic were done in [7, 21]. They gave tight bounds and achievability results, such as a Gilbert–Varshamov bound for LRCs. Several constructions of families of codes were then given, for instance using concatenated codes [7], or in [6, 17, 16, 23, 9]. Other considerations on the topic are the correction of multiple erasures and the correction from multiple recovery sets. Details can be found for instance in [5, 6].

In this paper, we consider the generalized AG-codes (GAG-codes) introduced in [24] by Xing, Niederreiter and Lam. The well known AG-codes defined by Goppa in [13] are given by the evaluation at rational places (i.e. places of degree 1) of functions of an algebraic function field defined over $\mathbb{F}_q$. The generalization of Xing et al. consists in using not only the evaluation at rational places, but at places of higher degrees. In fact, the evaluation at a place of degree d is an element in the residue class field, that is isomorphic to $\mathbb{F}_{q^d}$. It follows that a codeword composed by some evaluations at places of different degrees would be polyalphabetic. To address this difficulty, the solution proposed in [24] is to encode the evaluation at non rational places with a $\mathbb{F}_q$-linear code.

The key observation behind this document is as follows: if we apply the construction of [24] using non-rational places of degree at most $r > 1$, we can obtain linear codes with locality r. It turns out that some of these codes have good or optimal parameters with respect to the Singleton bound for LRCs (1). There are similarities between our codes and some concatenated codes, especially those introduced by Cadambe and Mazumbar in [7, Section VI. A.]. In order to make a comparison and investigate their differences, we give a construction of LRCs obtained by concatenation using an AG-code as outer code. More precisely, a construction from [7] uses a RS code as the outer code. We consider a similar construction by using an AG outer code and show its parameters are similar to those of our new construction using GAG-codes. Using the recursively defined tower of function fields of Garcia and Stichtenoth [11], we give an asymptotic study of both our new construction using generalized AG-codes and the construction of concatenated codes. An important difference is that our new approach from generalized AG-codes allows to construct directly codes whose dimension is not a multiple of the locality, contrary to the one using concatenated codes or the best-known constructions (e.g. [6]).

The paper is organized as follows. In Section 2, we recall the basics of LRCs and concatenated codes. In Section 3, we give the definitions and results of function field theory that we shall use, and present the generalization of AG-codes of [24]. In Section 4, we explain how one can obtain locality in these codes, and give an optimal example. In Section 4, we give two explicit families of LRCs, one using concatenated codes and the second with our new approach using GAG-codes. We give an asymptotic study. In particular, we show the existence of an infinite family of LRCs with locality 2 over finite fields of cardinality greater than 3 and give their parameters.

2 Locally Recoverable Codes (LRCs)

2.1 Generalities

In what follows, we denote by $[n, k, d]$ a linear code over $\mathbb{F}_q$ with length n, dimension k and minimum distance d. Throughout this paper, we focus on the notion of locally recoverable codes (LRCs).

Definition 2.1. *Let $\mathcal{C} \subset \mathbb{F}_q^n$ be a $\mathbb{F}_q$-linear code. The code $\mathcal{C}$ is locally recoverable with locality r if every symbol of a codeword can be recovered using a subset of at most r other symbols. The smallest such r is called the locality of the code.*

There exists a Singleton-like bound for LRCs [12] and an upper bound for the rate of codes with locality r, given in [5, Theorem 2.1], that we recall here.

Theorem 2.1. *Let $\mathcal{C}$ be a q-ary linear code with parameters $[n, k, d]$ and locality r. The rate of $\mathcal{C}$ verifies*

$$\frac{k}{n} \leq \frac{r}{r+1}.$$

The minimum distance d of $\mathcal{C}$ verifies

$$d \leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2.$$

However, these results do not take into account the size of the alphabet, i.e. the cardinality of the base field. There are several bounds considering this constraint. In [7] and [21, Theorem 5.1], the authors gave a Gilbert–Varshamov-type bound for LRCs, see (2), where $R_q(r, \delta)$ denotes the asymptotic bound on the rate of q-ary locally recoverable codes with locality r and relative minimum distance δ.

$$\begin{aligned} R_q(r, \delta) \geq 1 - \min_{0 \leq s \leq 1} \Big[\frac{1}{r+1} \log_q \big((1 + (q-1)s)^{r+1} \\ + (q-1)(1-s)^{r+1} \big) - \delta \log_q s \Big]. \end{aligned} \tag{2}$$

The construction of Tamo-Barg-Vladuts [6] is known to improve upon this bound. This result is also obtained by [17, 9]. Some achievability results considering a fixed alphabet size have been obtained via concatenated codes [7, 9].

2.2 Concatenated codes

Concatenated codes were introduced by Forney [10] in 1965. This name comes from the idea of successively applying two encoders. It consists in first using an outer code over a large alphabet, then using an inner code to encode the codeword symbols of the outer code. In our framework, a concatenated code can be defined as follows.

Definition 2.2. *Let $\mathcal{C}_{\text{out}}$ be a $q^{k'} - ary$ linear code of parameters $[n, k, d]$ and $\mathcal{C}_{\text{in}}$ be a $q - ary$ linear code of parameters $[n', k', d']$ such that*

$$\mathcal{C}_{\text{out}}(m) = (c_1, , \ldots, c_n),$$

where $m \in \mathbb{F}_{q^{k'}}^{k}$ and $c_1, \ldots, c_n \in \mathbb{F}_{q^{k'}}$. Then the concatenated code $\mathcal{C}_{\text{conc}}$ of $\mathcal{C}_{\text{out}}$ and $\mathcal{C}_{\text{in}}$ is defined by

$$\mathcal{C}_{\text{conc}}(m) = (\mathcal{C}_{\text{in}}(c_1) \ \mid \ \cdots \ \mid \ \mathcal{C}_{\text{in}}(c_n)) .$$

Note that the locality of a concatenated code is given by the one of the inner code [9, Theorem 4.1]. Recall also that such a code verifies the following properties.

Proposition 2.1. *The code $\mathcal{C}_{\text{conc}}$ is an $[nn', kk', \geq dd']$ linear code over $\mathbb{F}_q$.*

In [7, Theorem 2], the authors used concatenated codes to obtain asymptotic achievability results on binary LRCs. More precisely, they used an outer random $q^r - ary$ linear code and the $q - ary$ single parity check code of length $r + 1$ as the inner code, and proved the existence of an infinite family of $[n, k, d]_q$ linear codes with locality r. In [9], the authors used concatenated codes to obtain some dimension-optimal locally repairable codes. Moreover, they also used some shortening techniques to obtain dimension-optimal LRCs whose dimension is not a multiple of the locality.

3 Generalized AG-Codes

3.1 Algebraic function fields

Let $\mathbb{F}_q$ be the field with q elements, and let $F/\mathbb{F}_q$ be an algebraic function field over $\mathbb{F}_q$ of genus $g = g(F)$. For $\mathcal{O}$ a valuation ring, a place P is defined to be $\mathcal{O} \setminus \mathcal{O}^{\times}$. The evaluation of a function at P is an element of the residue class field F_P, that is isomorphic to $\mathbb{F}_{q^d}$, where d is the degree of the place. A rational place is a place of degree 1. A divisor $\mathcal{D}$ is defined as a formal sum of places, and we denote by $Supp(\mathcal{D})$ the support of $\mathcal{D}$ and $\mathcal{L}(\mathcal{D})$ the corresponding Riemann-Roch space. Details about algebraic function fields can be found in [20].

In the following, obtaining infinite families of codes with our construction relies on the existence of families of function fields with a large number of places of a given degree. In this context, let us introduce the Drinfeld–Vladut Bound at order r, such as stated in [2, Definition 1.3].

Definition 3.1 (Drinfeld–Vladut Bound of order r). *Let $F/\mathbb{F}_q$ be a function field over $\mathbb{F}_q$ and let $B_r(F/\mathbb{F}_q)$ denote its number of places of degree r. Let*

$$B_r(q, g) = \max\{B_r(F/\mathbb{F}_q) \ \mid \ F/\mathbb{F}_q \text{ is a function field over } \mathbb{F}_q \text{ of genus } g\}.$$

Then,

$$\limsup_{g \longrightarrow +\infty} \frac{B_r(q, g)}{g} \leq \frac{1}{r}(q^{\frac{r}{2}} - 1).$$

For $r = 1$, this gives the usual Drinfeld–Vladut Bound on the number of rational places. There exist several families of function fields reaching this bound, such as the Garcia–Stichtenoth tower of function fields [11]. Such towers are recalled and used in Section 5.

3.2 Generalized AG-codes (GAG-codes)

Let $F/\mathbb{F}_q$ be an algebraic function field defined over $\mathbb{F}_q$ of genus g. In [24], the authors introduced a generalization of AG-codes by using non-rational places. Following their work, we use the notations:

- $P_1, \ldots, P_s$ are s distinct places of F,
- G is a divisor of F such that $Supp(G) \bigcap \{P_1, \ldots, P_s\} = \emptyset$,

and for $1 \leq i \leq s$:

- $k_i = \deg(P_i)$ is the degree of P_i,
- C_i is an $[n_i, k_i, d_i]_q$ linear code,
- π_i is a fixed $\mathbb{F}_q$-linear isomorphism mapping $\mathbb{F}_{q^{k_i}}$ to C_i.

Consider the application

$$\alpha \ : \begin{array}{l} \mathcal{L}(G) \longrightarrow \mathbb{F}_q^n \\ f \quad\ \longmapsto (\pi_1(f(P_1)), \ldots, \pi_s(f(P_s))) \end{array},$$

where $n = \sum_{i=1}^{s} n_i$.

Definition 3.2. *The image of α is called a Generalized Algebraic-Geometric code (GAG-code), denoted by $C(P_1, \ldots, P_s : G : C_1, \ldots, C_s)$.*

Such a code is well-defined if the application α is injective, that is the case if $\deg(G) < \sum_{i=1}^{s} k_i$ [24, Lemma 3.1]. Furthermore, the authors give a lower bound on the minimum distance and the dimension of these codes in the following theorem [24, Theorem 3.2].

Theorem 3.1. *Under the same notations, if $\deg(G) < \sum_{i=1}^{s} k_i$, then the dimension k and the minimum distance d of the code defined by α verify*

- $k \geq \deg(G) - g + 1$, *with equality if* $\deg(G) \geq 2g - 1$,
- $d \geq \sum_{i=1}^{s} d_i - \deg(G) - \max_R \left\{\sum_{i \in R}(d_i - k_i)\right\}$,

where the maximum is extended over all subsets R of $1, \ldots, s$ and an empty sum is defined to be 0.

Remark 3.1. In [24], the authors wrote that this construction is inspired by concatenated codes, with the difference that several distinct "inner" codes can be used. There is another fundamental difference : the dimension of a GAG-code is given by the one of the "outer" code only. On the other hand, the dimension of a concatenated code is given by the product of the dimensions of the inner code and the outer code.

4 Locality in Generalized AG-codes

The main observation behind this paper is the following: if $k_1 = \cdots = k_s = r$, the code defined has locality r. More formally,

Proposition 4.1. *Let F be an algebraic function field defined over $\mathbb{F}_q$ of genus g. Let $P_1, \ldots, P_s$ be places of F of degree $k_i = \deg(P_i)$ respectively, and let G be a divisor such that $\deg(G) < \sum_{i=1}^{s} k_i$. For $1 \leq i \leq s$, let C_i be an $[n_i, k_i, d_i]$ $\mathbb{F}_q$-linear code with locality at most k_i. Let $\mathcal{C} = C(P_1, \ldots, P_s : G : C_1, \ldots, C_s)$ be a generalized AG-code as in Definition 3.2. If there exists $r \in \mathbb{N}$ such that for all $1 \leq i \leq s$, we have $1 < k_i \leq r$ and $n_i > \deg(P_i)$, then $\mathcal{C}$ has locality r.*

In order to obtain codes with a given locality r, it makes sense to use places $P_1, \ldots, P_s$ of degree r, and encode the evaluations at each P_i using the same code $\mathcal{C}'$. We obtain the following.

Proposition 4.2. *Let $\mathcal{C} = C(P_1, \ldots, P_s : G : C_1, \ldots, C_s)$ be a generalized AG-code as defined above. Suppose that $\deg P_1 = \cdots = \deg P_s = r$ and $\mathcal{C}' = C_1 = \cdots = C_s$ is an $[n', r, d']$ linear code with locality r. If $2g - 1 \leq \deg(G) < rs$, then $\mathcal{C}$ is an $[sn', deg(G) - g + 1, \geq d'\left(s - \left\lfloor \frac{\deg G}{r} \right\rfloor\right)]$ linear code over $\mathbb{F}_q$ with locality r.*

A specific family of such codes is introduced in Section 5 and its asymptotic properties are studied. For now, let us give an example reaching the Singleton bound for LRCs.

Example 4.1. Let $F = \mathbb{F}_3(x)$ be the rational function field over $\mathbb{F}_3$. It contains 4 rational places : P_0, P_1, P_2 and P_∞, where P_i can be defined by the polynomial $x - i$ for $0 \leq i \leq 2$ and P_∞ is the place at infinity. It also contains three places of degree 2 : P_1^2, P_2^2 and P_3^2, that can be defined by the irreducible polynomials $P_1^2(x) = x^2 + 2x + 2$, $P_2^2(x) = x^2 + 1$, and $P_3^2(x) = x^2 + x + 2$ respectively. Let $C_1 = C_2 = C_3 = \mathrm{RS}(3, 2) = \{(f(0), f(1), f(2)) \mid f \in \mathbb{F}_3[x]_{<2}\}$. The code $\mathcal{C} = C(P_1^2, P_2^2, P_3^2 : 4P_\infty : \mathrm{RS}(3, 2), \mathrm{RS}(3, 2), \mathrm{RS}(3, 2))$ is a $(9, 5)$ code over $\mathbb{F}_3$ with locality 2. According to Proposition 4.2, the minimum distance of this code is at least 2. Using Magma, we computed that the minimum distance of this code is 3. Consequently, this code is a $[9, 5, 3]$ linear code over $\mathbb{F}_3$ with locality 2, reaching the Singleton-like bound (1).

This example generalizes to any prime power $q \geq 3$. We have the following.

Proposition 4.3. *Let $q \geq 3$ be a prime power. One can similarly define a $[\frac{3}{2}(q^2 - q), q^2 - q - 1, 3]_q$ linear code with locality 2, reaching the Singleton bound.*

In the longer version of the paper, we present the results of our experiments over $\mathbb{F}_3$ using an elliptic curve and the Klein quartic.

5 Some families of LRCs and asymptotic study

Our construction is very close to what can be obtained with concatenated codes. In order to compare both constructions, we introduce a family of concatenated codes and another obtained with our approach.

5.1 Concatenated Construction.

This construction is a generalization to an outer AG-code of the construction of [7, Section VI. A.] that uses an outer extended Reed-Solomon code.

Proposition 5.1 (Concatenated Construction). *Let $F/\mathbb{F}_{q^r}$ be a function field of genus g containing s rational places, denoted by $P_1, \ldots, P_s$. Let $\mathcal{C}_{\text{par}}$ the q-ary single parity check code of length $r+1$ and dimension r, that has minimum distance 2. For $g-1 < k_0 < s-g+1$, let G be a divisor of F of degree k_0+g-1 and $\mathcal{D} = P_1 + \cdots + P_s$. Then, the concatenated code $\mathcal{C}_{\text{conc}}$ defined by the outer code $C(\mathcal{D}, G)$ and the inner code $\mathcal{C}_{\text{par}}$ is a $[n, k, \geq d]$ linear code over $\mathbb{F}_q$ with locality r, such that*

$$n = (r+1)s,$$

$$k = rk_0$$

$$d \geq 2\left(s - \frac{k}{r} - g + 1\right).$$

It follows that the rate of this code verifies

$$\frac{k}{n} \geq \frac{r}{r+1} - \frac{r}{2}\delta - \frac{r(g-1)}{n},$$

where $\delta = \frac{d}{n}$.

Note that in this construction, as well as in the known constructions of [5–7], the dimension is a multiple of the locality. The existence of infinite family of codes defined by this construction is ensured by sequences of function fields reaching the Drinfeld–Vladut bound, such as the recursive tower of function fields defined by Garcia and Stichtenoth [11].

Proposition 5.2. *Let q be a prime power and r an even integer, except $q = r = 2$. Then, the Concatenated Construction provides an infinite family of linear code with locality r verifying*

$$\frac{k}{n} \geq \frac{r}{r+1}\left(1 - \frac{r+1}{2}\delta - \frac{1}{q^{\frac{r}{2}}-1}\right).$$

Example 5.1. For $q = 4$ and $r = 2$, we obtain $\frac{k}{n} \geq \frac{1}{3} - \delta$.

Note that for $r = 2$, this gives the same bound as in [9, Theorem 3.7].

5.2 New construction

Let us introduce a specific family of codes obtained with our new strategy, using Proposition 4.2.

Proposition 5.3 (GAG Construction). *Let $F/\mathbb{F}_q$ be a function field of genus g containing s places of degree $r > 1$, denoted by $P_1, \dots, P_s$. Let $\mathcal{C}_{\text{par}}$ the q-ary single parity check code of length $r+1$ and dimension r, that has minimum distance 2. For $g-1 < k < rs-g+1$, let G be a divisor of F of degree $k+g-1$. Then, the code $C(P_1, \dots, P_s : G : \mathcal{C}_{\text{par}}, \dots, \mathcal{C}_{\text{par}})$ is an $[n, k, \geq d]$ linear code over $\mathbb{F}_q$ with locality r, such that*

$$n = (r+1)s,$$
$$d \geq 2\left(s - \left\lfloor \frac{k+g-1}{r} \right\rfloor\right).$$

It follows that the rate of this code verifies

$$\frac{k}{n} \geq \frac{r}{r+1} - \frac{r}{2}\delta - \frac{g-1}{n},$$

where $\delta = \frac{d}{n}$.

In [2], Ballet and Rolland studied the descent of the tower $\mathcal{T}/\mathbb{F}_{q^2}$ to the field of constant $\mathbb{F}_q$. The authors also proved that these towers reach the Drinfeld–Vladut bound at order 2 [2, Proposition 3.3]. This allows us to prove the existence of infinite families of linear code with locality 2.

Proposition 5.4. *Let $q > 3$ be a prime power. Then, the GAG Construction provides an infinite family of linear code with locality 2 verifying*

$$\frac{k}{n} \geq \frac{2}{3}\left(1 - \frac{q}{q^2-q-2}\right) - \delta.$$

Remark 5.1. While ℓ is increasing, the rate of the codes defined by the GAG Construction tends to verify $\frac{k}{n} \geq \frac{2}{3}(1 - \frac{3}{2}\delta - \frac{1}{q-1})$. This is exactly the bound obtained for the Concatenated Construction in Proposition 5.2, specialized to locality 2. More generally, according to the Drinfeld–Vladut Bound, the best rate that can be obtained is

$$\frac{n}{k} \geq \frac{r}{r+1}\left(1 - \frac{r+1}{2}\delta - \frac{1}{q^{\frac{r}{2}}-1}\right).$$

Remark 5.2. The GAG Construction requires asymptotically a large number of places of degree r. Such objects can be obtained by the descent to $\mathbb{F}_q$ of function fields defined of $\mathbb{F}_{q^r}$ reaching the Drinfeld–Vladut bound ([8], see [3]). The sequences studied in [2] for $r = 2$ or $q = 2$ and $r = 4$ are convenient for our study.

Remark 5.3. One can construct directly codes where the dimension is not a multiple of the locality with the new GAG Construction, while it is not possible with the Concatenated Construction. In the literature, it is classical to obtain LRCs whose dimension is a multiple of the locality, then some techniques can be used to obtain different dimensions, as in [9].

Remark 5.4. A reasonable question is whether one can build codes with locality 2 of any dimension $k \in \mathbb{N}$ using the GAG Construction. It is not possible with the tower used previously, but it might be possible using the densified version of the tower introduced in [2].

Remark 5.5. Although it was quite natural to consider only places of a fixed degree r, one can extend our new construction to places of smaller degree, provided that we combine them to obtain spaces of dimension r. Moreover, one can also consider generalized evaluation maps, and for instance use the local expansion at order r at rational places.

Remark 5.6. The construction introduced in this document might be generalized in order to obtain codes with hierarchical locality [1].

6 Aknowledgments

The author is deeply grateful to the French ANR BARRACUDA (ANR-21-CE39-0009-BARRACUDA) for its support, and to several of its members for the many inspiring and helpful discussions. The author furthermore thanks anonymous referees for their valuable remarks and comments.

References

1. Sean Ballentine, Alexander Barg, and Serge Vlăduţ. Codes with hierarchical locality from covering maps of curves. *IEEE Transactions on Information Theory*, 65(10):6056–6071, 2019.
2. Stéphane Ballet and Robert Rolland. Families of curves over any finite field attaining the generalized Drinfeld-Vlăduţ bound. *Publications Mathématiques de Besançon, Algèbre et Théorie des Nombres*, pages 5–18, 2011.
3. Stéphane Ballet, Jean Chaumine, Julia Pieltant, Matthieu Rambaud, Hugues Randriambololona, and Robert Rolland. On the tensor rank of multiplication in finite extensions of finite fields and related issues in algebraic geometry. *Uspekhi Mathematichskikh Nauk, 76:1(457), 31–94*, 2021.
4. Alexander Barg, Kathryn Haymaker, Everett W. Howe, Gretchen L. Matthews, and Anthony Várilly-Alvarado. Locally recoverable codes from algebraic curves and surfaces. In Everett W. Howe, Kristin E. Lauter, and Judy L. Walker, editors, *Algebraic Geometry for Coding Theory and Cryptography*, pages 95–127, Cham, 2017. Springer International Publishing.
5. Alexander Barg and Itzhak Tamo. A family of optimal locally recoverable codes. *IEEE Transactions on Information Theory*, 60(8):4661–4676, 2014.
6. Alexander Barg, Itzhak Tamo, and Serge Vlăduţ. Locally recoverable codes on algebraic curves. *IEEE Transactions on Information Theory*, 63(8):4928–4939, 2017.
7. Viveck R. Cadambe and Arya Mazumdar. Bounds on the size of locally recoverable codes. *IEEE Transactions on Information Theory*, 61(11):5787–5794, 2015.
8. Ignacio Cascudo, Ronald Cramer, Chaoping Xing, and An Yang. Asymptotic bound for multiplication complexity in the extensions of small finite fields. *IEEE Transactions on Information Theory*, 58(7):4930–4935, 2012.

9. Jin Yi Chen, Shu Liu, Liming Ma, Ting-Yi Wu, and Chaoping Xing. Optimal and asymptotically good locally repairable codes via propagation rules. *IEEE Transactions on Communications*, 71(10):5623–5632, 2023.
10. G.D. Forney. *Concatenated Codes*. M.I.T. Press research monographs. M.I.T. Press, 1966.
11. Arnaldo Garcia and Henning Stichtenoth. A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vlăduţ bound. *Inventiones Mathematicae*, 121:211–222, 1995.
12. Parikshit Gopalan, Cheng Huang, Huseyin Simitci, and Sergey Yekhanin. On the locality of codeword symbols. *IEEE Transactions on Information Theory*, 58(11):6925–6934, 2012.
13. V. D. Goppa. Codes on algebraic curves. *Dokl. Akad. Nauk SSSR*, 259(6):1289–1290, 1981.
14. Venkatesan Guruswami, Chaoping Xing, and Chen Yuan. How long can optimal locally repairable codes be? *IEEE Transactions on Information Theory*, 65(6):3662–3670, 2019.
15. Cheng Huang, Minghua Chen, and Jin Li. Pyramid codes: Flexible schemes to trade space for access efficiency in reliable data storage systems. In *Sixth IEEE International Symposium on Network Computing and Applications (NCA 2007)*, pages 79–86, 2007.
16. Gaojun Luo and Xiwang Cao. Constructions of optimal binary locally recoverable codes via a general construction of linear codes. *IEEE Transactions on Communications*, 69(8):4987–4997, 2021.
17. Liming Ma and Chaoping Xing. Constructive asymptotic bounds of locally repairable codes via function fields. *IEEE Transactions on Information Theory*, 66(9):5395–5403, 2020.
18. Cecilia Salgado, Anthony Várilly-Alvarado, and Jose Felipe Voloch. Locally recoverable codes on surfaces. *IEEE Transactions on Information Theory*, 67(9):5765–5777, 2021.
19. Natalia Silberstein, Ankit Singh Rawat, O. Ozan Koyluoglu, and Sriram Vishwanath. Optimal locally repairable codes via rank-metric codes. In *2013 IEEE International Symposium on Information Theory*, pages 1819–1823, 2013.
20. Henning Stichtenoth. *Algebraic Function Fields and Codes*. Number 254 in Graduate Texts in Mathematics. Springer-Verlag, second edition, 2008.
21. Itzhak Tamo, Alexander Barg, and Alexey Frolov. Bounds on the parameters of locally recoverable codes. *IEEE Transactions on Information Theory*, 62(6):3070–3083, 2016.
22. Itzhak Tamo, Dimitris S. Papailiopoulos, and Alexandros G. Dimakis. Optimal locally repairable codes and connections to matroid theory. In *2013 IEEE International Symposium on Information Theory*, pages 1814–1818, 2013.
23. Pan Tan, Cuiling Fan, Cunsheng Ding, Chunming Tang, and Zhengchun Zhou. The minimum locality of linear codes. *Des. Codes Cryptography*, 91(1):83–114, 2023.
24. Chaoping Xing, Harald Niederreiter, and Kwok-Yan Lam. A generalization of algebraic-geometry codes. *IEEE Trans. Inf. Theory*, 45(7):2498–2501, 1999.

On rotation-symmetric Boolean bent functions outside the $\mathcal{M}^{\#}$ class

Alexandr Polujan[1], Sadmir Kudin[2], and Enes Pasalic[2]

[1] Otto-von-Guericke-Universität, Universitätsplatz 2, 39106, Magdeburg, Germany
alexandr.polujan@gmail.com
[2] University of Primorska, FAMNIT & IAM, Glagoljaška 8, 6000 Koper, Slovenia
sadmir.kudin@famnit.upr.si, enes.pasalic6@gmail.com

Abstract. Rotation-symmetric bent functions, being invariant under the action of the cyclic group, attracted a lot of attention in the last three decades due to their applications in cryptography. Finding new constructions of such functions is a well-known difficult problem [4, Open Problem 17]. Most of the known constructions of rotation-symmetric bent functions are based on applying equivalence mappings to special Maiorana-McFarland bent functions in such a way, that a resulting function is invariant under the action of cyclic group. Finding rotation-symmetric bent functions not of this type (thus, those which do not belong to the completed Maiorana-McFarland class $\mathcal{M}^{\#}$) is a very challenging problem [24, p. 27], to which no solutions are currently known. In this paper, we provide for the first time a solution to this problem, by showing that an infinite family of rotation-symmetric bent functions [21] does not belong to $\mathcal{M}^{\#}$, for all $n \geq 8$.

Keywords: Bent function, Rotation-Symmetry, Maiorana-McFarland class, EA-equivalence, Classification.

1 Introduction

Let $\mathbb{F}_2^n$ be the vector space of dimension n over $\mathbb{F}_2 = \{0, 1\}$. A mapping $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ is called a *Boolean function* in n variables, and the set of all Boolean functions in n variables is denoted by $\mathcal{B}_n$. The *Walsh transform* $W_f\colon \mathbb{F}_2^n \to \mathbb{Z}$ of $f \in \mathcal{B}_n$ is defined by $W_f(a) = \sum_{x\in\mathbb{F}_2^n}(-1)^{f(x)+x\cdot a}$ for $a \in \mathbb{F}_2^n$, where $x \cdot a = x_1a_1 + x_2a_2 + \cdots + x_na_n$. A Boolean function $f \in \mathcal{B}_n$ with $n = 2m$ is called *bent* if $|W_f(a)| = 2^{n/2}$ for all $a \in \mathbb{F}_2^n$. For a Boolean bent function $f \in \mathcal{B}_n$, the Boolean function $\tilde{f} \in \mathcal{B}_n$ defined for any $a \in \mathbb{F}_2^n$ by $W_f(a) = 2^{\frac{n}{2}}(-1)^{\tilde{f}(a)}$, is also bent and is called the *dual* of f. A Boolean function on $\mathbb{F}_2^n$ can be uniquely expressed as a polynomial in $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1+x_1^2, \ldots, x_n+x_n^2)$. This representation is called the *algebraic normal form* (*ANF*, for short), that is, $f(x) = \sum_{v\in\mathbb{F}_2^n} c_v \left(\prod_{i=i}^{n} x_i^{v_i}\right)$, where $x = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$, $c_v \in \mathbb{F}_2$ and $v = (v_1, \ldots, v_n) \in \mathbb{F}_2^n$. The *algebraic degree* of a Boolean function f, denoted by $\deg(f)$, is the algebraic degree of its ANF. On the set of all Boolean functions one can introduce an equivalence relation in the following way: two functions $f, f' \in \mathcal{B}_n$ are called *extended-affine*

equivalent (EA-equivalent), if there exist a non-degenerate affine transformation $A \in AGL(n,2)$ and an affine function $l(x) = a \cdot x + b$ on $\mathbb{F}_2^n$, where $a \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2$, such that $f'(x) = f(xA) + l(x)$ holds for all $x \in \mathbb{F}_2^n$.

The *Maiorana-McFarland class* $\mathcal{M}$ is the set of n-variable ($n = 2m$) Boolean bent functions of the form

$$f(x,y) = x \cdot \pi(y) + h(y), \text{ for all } x, y \in \mathbb{F}_2^m,$$

where π is a permutation on $\mathbb{F}_2^m$, and $h \in \mathcal{B}_m$ is an arbitrary Boolean function. The set of bent functions $f \in \mathcal{B}_n$ which are equivalent to at least one function in $\mathcal{M}$ is called the *completed Maiorana-McFarland class* and is denoted by $\mathcal{M}^{\#}$. Note that for $n = 2, 4, 6$, all bent functions in $\mathcal{B}_n$ are members of $\mathcal{M}^{\#}$, see [9, p. 214].

With the following criterion of Dillon, one can show that a given Boolean bent function $f \in \mathcal{B}_n$ is (not) a member of the completed Maiorana-McFarland class.

Lemma 1. *[10, p. 102] Let $n = 2m$. A Boolean bent function $f \in \mathcal{B}_n$ belongs to $\mathcal{M}^{\#}$ if and only if there exists an m-dimensional linear subspace V of $\mathbb{F}_2^n$ such that, for any $a, b \in V$,*

$$D_a D_b f(x) = f(x) + f(x+a) + f(x+b) + f(x+a+b) = 0, \text{ for all } x \in \mathbb{F}_2^n.$$

Following [19], we call a subspace U of $\mathbb{F}_2^n$ an *$\mathcal{M}$-subspace* of $f \in \mathcal{B}_n$, if for all $a, b \in U$ we have that $D_{a,b} f = 0$.

In order to introduce rotation-symmetric Boolean bent functions, which are the main subject of this paper, we give the following notation. Let $(x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n$. For $1 \leq k \leq n$, we define the *cyclic shift* to the right by k positions

$$\rho_n^k(x_i) = \begin{cases} x_{i+k} & \text{if } i+k \leq n \\ x_{i+k-n} & \text{if } i+k > n \end{cases}. \tag{1}$$

We extend the definition of ρ to tuples by $\rho_n^k(x_1, \ldots, x_n) = \left(\rho_n^k(x_1), \ldots, \rho_n^k(x_n)\right)$ and to monomials by $\rho^k(x_{i_1} \cdots x_{i_\ell}) = \rho^k(x_{i_1}) \cdots \rho^k(x_{i_\ell})$.

Definition 1. *A Boolean function $f \in \mathcal{B}_n$ is rotation-symmetric (RotS) if for any vector $(x_1, \ldots, x_n) \in \mathbb{F}_2^n$, the equality $f(\rho_n^k(x_1, \ldots, x_n)) = f(x_1, \ldots, x_n)$ holds for any $1 \leq k \leq n$.*

A Boolean function $f \in \mathcal{B}_n$ is called *rotation-symmetric bent*, if it is bent and rotation-symmetric. Sometimes, it is more convenient to represent such functions with the help of *short algebraic normal form (SANF)*, which is defined as follows.

Definition 2. *Let $f \in \mathcal{B}_n$ be a rotation-symmetric function. We define*

$$G_n(x_1, \ldots, x_n) = \left\{\rho_n^k(x_1, \ldots, x_n), \text{ for } 1 \leq k \leq n\right\},$$

that is, the orbit of $(x_1, \ldots, x_n)$ *under the action of* $\rho_n^k, 1 \leq k \leq n$. *The function* f *can be written as*

$$a_0 + a_1 x_1 + \sum_{j=1}^{n} a_{1j} x_1 x_j + \cdots + a_{12\ldots n} x_1 x_2 \cdots x_n,$$

where the coefficients $a_0, a_1, a_{1j}, \ldots, a_{12\ldots n} \in \mathbb{F}_2$, *and the existence of a representative term* $x_1 x_{i_2} \ldots x_{i_\ell}$ *implies the existence of all terms from* $G_n(x_1 x_{i_2} \ldots x_{i_l})$ *in the ANF. This representation of* f *is called the short algebraic normal form (SANF) of* f.

In the following example, we illustrate the connection between SANF and ANF of a RotS function.

Example 1. Let f be a cubic RotS function in $n = 4$ variables, whose SANF is given by $x_1x_2 + x_1x_2x_3$. Then, its ANF is given by $x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1 + x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_1 + x_4x_1x_2$.

One of the central problems in the theory of rotation-symmetric bent functions is the construction of new infinite families [4, Open Problem 17]. The first examples of RotS bent functions were obtained with a computer in [20], where the question about theoretical constructions of RotS bent functions was raised. First theoretical constructions of RotS bent functions of degree 2 and 3 (in the $\mathcal{M}^\#$ class) were obtained in [13]. A problem of finding RotS bent functions of higher degrees was solved later in [22], where the first general construction of rotation symmetric bent functions in $\mathcal{B}_{2m}$ for any m with an arbitrary degree in the range from 2 to m was proposed using the $\mathcal{M}^\#$ class. Later, several other constructions of bent functions of higher degrees were proposed, see, e.g., [7,21,24,25]. Due to the importance of the $\mathcal{M}^\#$ class in the construction methods of RotS bent functions, the following open problem was suggested by Zhao, Zheng and Zhang in [24, p. 27]:

Open Problem 1 *How to construct RotS bent functions which do not belong to the Maiorana-McFarland class?*

The main aim of this paper is to provide the first solution to this open problem by analyzing an infinite family of RotS bent functions of the maximum algebraic degree constructed by Su [21].

The rest of the article is organized in the following way. In Section 2, we extend the computational investigation of RotS bent functions, originally initiated in [20]. Particularly, we enumerate and classify all rotation-symmetric cubic bent functions in ten variables, and indicate that some of them do not belong, up to equivalence, to the completed Maiorana-McFarland class $\mathcal{M}^\#$. In Section 3, we show that an infinite family of RotS bent functions of Su [21], does not belong to $\mathcal{M}^\#$, for all even $n \geq 8$. In this way, we provide a solution to Open Problem 1. The paper is concluded in Section 4.

2 Classification and enumeration of cubic RotS bent functions in 10 variables

In this section, we classify and enumerate rotation-symmetric cubic bent functions in 10 variables, thus extending the computational results in [20]. The original motivation of restricting ourselves to the functions of degree 3 was to find more examples of cubic bent functions outside $\mathcal{M}^{\#}$, of which only a few examples and constructions are known [19]. Since our classification approach is based on the use of combinatorial designs obtained from linear codes, we give the following definition first.

Definition 3. *Let n be even and let $f \in \mathcal{B}_n$ be a Boolean bent function. Let the linear code $\mathcal{C}_f$ over $\mathbb{F}_2$ be defined as the row space of the $(n+2) \times 2^n$-matrix over $\mathbb{F}_2$ with columns $(1, x, f(x))_{x \in \mathbb{F}_2^n}^T$. The incidence structure $\mathbb{D}(f)$ supported by the codewords of the minimum weight $w = 2^{n-1} - 2^{n/2-1}$ in $\mathcal{C}_f$ is a 2-$(2^n, 2^{n-1} - 2^{n/2-1}, 2^{n-2} - 2^{n/2-1})$ design, which is called an addition design of f.*

It is well-known, that Boolean bent functions $f, f' \in \mathcal{B}_n$ are EA-equivalent if and only if their linear codes $\mathcal{C}_f$ and $\mathcal{C}_{f'}$ are permutation equivalent [12, Theorem 9]. In turn, permutation equivalence of codes $\mathcal{C}_f$ and $\mathcal{C}_{f'}$ can be reduced to isomorphism of addition designs $\mathbb{D}(f)$ and $\mathbb{D}(f')$, since an incidence matrix of $\mathbb{D}(f)$ is a generator matrix of $\mathcal{C}_f$, for details we refer to [11,18]. Using this approach (reducing equivalence of Boolean and vectorial functions to isomorphism of well-defined incidence structures), several important classes of certain well-defined mappings $F \colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ were recently classified [1,16,17].

Now, we describe the main idea behind the used methodology and give the main steps of the used approach. For all the following steps we assume that $n = 10$.

1. Construct all cubic RotS cubic functions on $\mathbb{F}_2^n$ using the modular invariant theory approach [6,8,20].
   1.1 Construct the ring of polynomials $K[x_1, \ldots, x_n]^{C_n}$ of degree 2 and 3, containing the polynomials that are invariant under the action of the cyclic group C_n. As explained in [8, p. 129], for $n = 10$, we have that $\dim(K) = g_{n,2} + g_{n,3} = \frac{n}{2} + \frac{(n-1)\cdot(n-2)}{6} = 17$, where $g_{n,w}$ is the number of distinct cycles of weight w. The terms of degree 1 are not included in the definition of K, since for an affine function $l \in \mathcal{B}$, Boolean functions f and $f + l$ on $\mathbb{F}_2^n$ are EA-equivalent.
   1.2 The ring $K[x_1, \ldots, x_n]^{C_n}$ can be constructed with, e.g., Magma [2], and SANFs of generators of K are given in the following list:

$$I = \{x_1x_2, x_1x_3, x_1x_4, x_1x_5, x_1x_6, x_1x_2x_3, x_1x_2x_4, x_1x_2x_5, x_1x_2x_6 \\ x_1x_2x_7, x_1x_2x_8, x_1x_2x_9, x_1x_3x_5, x_1x_3x_6, x_1x_3x_7, x_1x_3x_8, x_1x_4x_7\}.$$

2. Determine all cubic bent functions K per definition. In total, we got 1572 cubic bent functions.

3. Split the collection to the *"preclasses"* using invariants (a preclass is understood as a subset of functions with the same value of some invariant). Here we use the "distribution of the first-order derivatives", which is a multiset, that counts how many first-order derivatives $D_a f(x) = f(x+a) + f(x)$ are affine (0), ..., semi-bent (8). Here, the value in brackets is simply the rank over $\mathbb{F}_2$ of the matrix $A + A^T$, where A is an upper triangular matrix defined by the quadratic form $x \mapsto D_a f(x) = xAx^T + l(x)$, where l is an affine function. We do not give a proof that "distribution of the first-order derivatives" is an invariant, however it is clear that two EA-equivalent functions f and f' have the same collections of derivatives up to EA-equivalence.
   Using this invariant, we got 6 preclasses of RotS cubic bent functions. The corresponding values can be found in Table 2.
4. Classify the functions using designs from linear codes (see Definition 3).
   In total, we got 8 EA-equivalence classes. The representatives given by SANF are given in Table 1.
5. Check, which EA-equivalence classes (do not) belong to the $\mathcal{M}^\#$ class using [19, Algorithm 1].
   Note that if $f \in \mathcal{M}^\#$, then for any f' EA-equivalent to f it holds that $f' \in \mathcal{M}^\#$. In this way, it is enough to check membership of an arbitrary representative of a given equivalence class w.r.t. to inclusion in $\mathcal{M}^\#$. As one can see from Table 2, there is only one EA-equivalence class C_8 of RotS cubic bent functions outside $\mathcal{M}^\#$. This is the first example of a rotation-symmetric bent function outside $\mathcal{M}^\#$ in the literature (to the best of our knowledge). The SANF of its representative $f_8 \in C_8$ is given in Table 1.

We summarize all these observations in the following proposition.

Proposition 1. *On $\mathbb{F}_2^{10}$, there exist 1572 rotation-symmetric cubic bent functions, which are divided into 8 EA-equivalence classes. Among them, there is one class outside $\mathcal{M}^\#$.*

We give representatives of the obtained equivalence classes in Table 1 and summarize their inclusion in $\mathcal{M}^\#$ together with the distribution of derivatives in Table 2.

Table 1. RotS cubic bent functions in 10 variables: representatives f_i of extended-affine equivalence classes C_i together with their cardinalities

$f_i \in C_i$	SANF of a representative $f_i \in C_i$	$\lvert C_i \rvert$
f_1	$x_1x_6 + x_1x_2x_5 + x_1x_2x_7 + x_1x_2x_8$	384
f_2	$x_1x_6 + x_1x_2x_4 + x_1x_2x_5 + x_1x_2x_6 + x_1x_2x_7 + x_1x_2x_8 + x_1x_2x_9$	24
f_3	$x_1x_6 + x_1x_2x_3 + x_1x_2x_4 + x_1x_2x_6 + x_1x_2x_7 + x_1x_2x_9 + x_1x_3x_7$	72
f_4	$x_1x_2 + x_1x_6 + x_1x_2x_4 + x_1x_2x_7 + x_1x_2x_9$	384
f_5	$x_1x_2 + x_1x_6 + x_1x_2x_4 + x_1x_2x_6 + x_1x_2x_7 + x_1x_2x_9 + x_1x_3x_8$	384
f_6	$x_1x_6 + x_1x_2x_3 + x_1x_2x_5 + x_1x_2x_7 + x_1x_2x_9 + x_1x_3x_5$	192
f_7	$x_1x_6 + x_1x_2x_3 + x_1x_2x_5 + x_1x_2x_8 + x_1x_3x_7$	36
f_8	$x_1x_6 + x_1x_2x_5 + x_1x_2x_9 + x_1x_3x_5 + x_1x_3x_8$	96
Total	—	1572

Table 2. Equivalence classes of RotS cubic bent functions in 10 variables

C_i	C_i in $\mathcal{M}^\#$?	Distr. of der. of C_i
C_1	✓	$\{* \, 2^{10}, 4^{101}, 6^{560}, 8^{352} \, *\}$
C_2		$\{* \, 0^1, 2^{510}, 8^{512} \, *\}$
C_3		
C_4		$\{* \, 4^{31}, 8^{992} \, *\}$
C_5		
C_6		$\{* \, 2^{15}, 4^{256}, 6^{240}, 8^{512} \, *\}$
C_7		$\{* \, 0^{31}, 2^{480}, 4^{512} \, *\}$
C_8	×	$\{*2^{30}, 4^{257}, 6^{480}, 8^{256} \, *\}$

Remark 1. If a multiset in Table 2 does not contain 0, it means that all elements of the corresponding equivalence class do not contain affine derivatives. We notice that in general only a few infinite families of such functions are known, see [3,14].

Since even rotation-symmetric bent functions of low algebraic degree can be outside the $\mathcal{M}^\#$ class (as the findings of this section indicate), it is reasonable to focus on the analysis of infinite families of rotation-symmetric bent of the maximum algebraic degree w.r.t. to their inclusion in $\mathcal{M}^\#$. This problem will be considered in detail in the following section.

3 A family of rotation-symmetric bent functions outside the $\mathcal{M}^\#$ class

As indicated in Introduction, most of the known infinite families of rotation-symmetric bent functions belong to $\mathcal{M}^\#$. In this section, we show that two families of bent functions constructed by Su in [21] do not belong to the $\mathcal{M}^\#$ class. For convenience, we enumerate in this section coordinates of a vector $x \in \mathbb{F}_2^n$ starting from 0, that is $x = (x_0, \ldots, x_{n-1})$.

The Construction. The following two families of bent functions were constructed in [21]. For any even integer $n = 2m \geq 4$, a construction of n-variable rotation-symmetric bent function with maximal algebraic degree m is given as

$$f\left(x_0, x_1 \cdots, x_{n-1}\right) = \sum_{i=0}^{m-1} \left(x_i x_{m+i}\right) + \sum_{i=0}^{n-1} \left(x_i x_{i+1} \cdots x_{i+m-2} \overline{x_{i+m}}\right), \qquad (2)$$

whose dual function is

$$\widetilde{f}\left(x_0, x_1 \cdots, x_{n-1}\right) = \sum_{i=0}^{m-1} \left(x_i x_{m+i}\right) + \sum_{i=0}^{n-1} \left(x_i x_{i+1} \cdots x_{i+m-2} \overline{x_{i+n-2}}\right), \qquad (3)$$

where $\overline{x_i} = x_i + 1$ and the subscript of x is modulo n.

To prove that bent functions defined by (2) and (3) are outside $\mathcal{M}^\#$, it is enough to show that only $f \notin \mathcal{M}^\#$, as [15, Remark 3.1] indicates. For the completeness, we add its proof since we are not aware of any explicitly stated in the literature.

Lemma 2. *Let $f \in \mathcal{B}_n$ be bent. Then, $f \in \mathcal{M}^\#$ if and only if $\tilde{f} \in \mathcal{M}^\#$.*

Proof. W.l.o.g, we assume that f is given in bivariate form $f : \mathbb{F}_2^{n/2} \times \mathbb{F}_2^{n/2} \to \mathbb{F}_2$. Suppose that $f \in \mathcal{M}$. That is, there exists a permutation π of $\mathbb{F}_2^{n/2}$ and a Boolean function h in $n/2$ variables such that $f(x, y) = x \cdot \pi(y) + h(y)$ for $x, y \in \mathbb{F}_2^{n/2}$. The dual $\tilde{f}$ of f (see e.g. [5]) is then defined by $\tilde{f}(x, y) = y \cdot \pi^{-1}(x) \oplus h(\pi^{-1}(x))$ for $x, y \in \mathbb{F}_2^{n/2}$, where π^{-1} is the inverse permutation of π. It is clear that $\tilde{f} \in \mathcal{M}$.

Now suppose that $f \in \mathcal{M}^\#$ is EA-equivalent to some function $f' \in \mathcal{M}$. That is, for some affine permutation ψ of $\mathbb{F}_2^{n/2} \times \mathbb{F}_2^{n/2}$ and some elements $\lambda, \mu \in \mathbb{F}_2^{n/2}, \varepsilon \in \mathbb{F}_2$, we have that $f(x, y) = f'(\psi(x, y)) + (\lambda, \mu) \cdot (x, y) + \varepsilon$. The dual $\tilde{f}$ of f is defined by

$$\begin{aligned}
(-1)^{\tilde{f}(x,y)} &= 2^{-n/2} W_f(x, y) \\
&= 2^{-n/2} \sum_{a,b \in \mathbb{F}_2^{n/2}} (-1)^{f'(\psi(a,b)) + (a,b)\cdot(\lambda,\mu) + (a,b)\cdot(x,y) + \varepsilon} \\
&= 2^{-n/2} (-1)^{\varepsilon} \sum_{a,b \in \mathbb{F}_2^{n/2}} (-1)^{f'(\psi(a,b)) + (a,b)\cdot(x+\lambda, y+\mu)} \\
&= 2^{-n/2} (-1)^{\varepsilon} W_{f' \circ \psi}(x + \lambda, y + \beta) \\
&= 2^{-n/2} (-1)^{\varepsilon} \cdot 2^{n/2} (-1)^{\widetilde{(f' \circ \psi)}(x+\lambda, y+\mu)} \\
&= (-1)^{\widetilde{(f' \circ \psi)}(x+\lambda, y+\mu) + \varepsilon}.
\end{aligned}$$

Thus, $\tilde{f}(x, y) = \widetilde{(f' \circ \psi)}(x + \lambda, y + \mu) + \varepsilon$. Since ψ is an affine permutation of $\mathbb{F}_2^{n/2} \times \mathbb{F}_2^{n/2}$, we know that $(f' \circ \psi)(a, b) = f'(\psi(a, b)) = f'(a', b')$ for some $a', b' \in \mathbb{F}_2^{n/2}$. Because $f' \in \mathcal{M}$, it follows that $f' \circ \psi \in \mathcal{M}$. Consequently, we then have that $\widetilde{(f' \circ \psi)} \in \mathcal{M}^\#$. In other words, $\tilde{f} \in \mathcal{M}^\#$. Thus, we conclude that if $f \in \mathcal{M}^\#$ then also $\tilde{f} \in \mathcal{M}^\#$. Furthermore, as $\tilde{f}$ is also a bent function, we have that if $\tilde{f} \in \mathcal{M}^\#$ then $\tilde{\tilde{f}} = f \in \mathcal{M}^\#$. This concludes the proof. □

We will also need the following technical result about derivatives of the indicator of a flat.

Lemma 3. *Let E be a subspace of $\mathbb{F}_2^n$, and let $a \in \mathbb{F}_2^n$ and $\mathbb{1}_{a+E} : \mathbb{F}_2^n \to \mathbb{F}_2$ be the indicator of $a + E$. Let $v_1, \ldots, v_k$ be a set of vectors in $\mathbb{F}_2$. If $v_1, \ldots, v_k$ are linearly independent in the quotient space $\mathbb{F}_2^n / E$ then*

$$D_{v_1} \ldots D_{v_k} \mathbb{1}_{a+E} = \mathbb{1}_{a+E'},$$

where E' is the subspace $E' = \langle E, v_1, \ldots, v_k \rangle$. Otherwise, if $v_1, \ldots, v_k$ are linearly dependent in $\mathbb{F}_2^n / E$, then $D_{v_1} \ldots D_{v_k} \mathbb{1}_{a+E} = 0$.

Proof. The proof is by induction on k. Assume $k = 1$. Then, the statement v_1 is linearly independent in the quotient space $\mathbb{F}_2^n / E$ simply means that $v_1 \notin E$.

Hence, if B is a basis for E, then $B \cup \{v_1\}$ is a basis for $E' = \langle E, v_1, \rangle$. We compute

$$D_{v_1}\mathbb{1}_{a+E}(x) = \mathbb{1}_{a+E}(x) + \mathbb{1}_{a+E}(x+v_1) = \mathbb{1}_{a+E}(x) + \mathbb{1}_{a+v_1+E}(x),$$

for all $x \in \mathbb{F}_2^n$. If $x \in a + E'$, then $x = a + e + cv_1$, for some unique $e \in E$ and $c \in \mathbb{F}_2$. If $c = 0$, then $\mathbb{1}_{a+E}(x) = 1$ and $\mathbb{1}_{a+v_1+E}(x) = 0$, and if $c = 1$, then $\mathbb{1}_{a+E}(x) = 0$ and $\mathbb{1}_{a+v_1+E}(x) = 1$. Hence, in any case, $D_{v_1}\mathbb{1}_{a+E}(x) = 1$. If $x \in \mathbb{F}_2^n \setminus a + E'$, then $\mathbb{1}_{a+E}(x) = \mathbb{1}_{a+v_1+E}(x) = 0$, hence $D_{v_1}\mathbb{1}_{a+E}(x) = 0$. We conclude that $D_{v_1}\mathbb{1}_{a+E}(x) = \mathbb{1}_{a+E'}(x)$, for all $x \in \mathbb{F}_2^n$. Otherwise, if v_1 is linearly dependent in the quotient space $\mathbb{F}_2^n/E$, then $v_1 \in E$, hence $v_1 + E = E$, so $D_{v_1}\mathbb{1}_{a+E}(x) = 0$.

Assume now that the statement is true for $k-1$, and that $v_1, \ldots, v_k$ are linearly independent in $\mathbb{F}_2^n/E$. Then, $v_2, \ldots, v_k$ are linearly independent in $\mathbb{F}_2^n/E$, and since the result is true for $k-1$, we have that

$$D_{v_2} \ldots D_{v_k}\mathbb{1}_{a+E} = \mathbb{1}_{a+E''},$$

where E'' is the subspace $E'' = \langle E, v_2, \ldots, v_k \rangle$. Consequently, from the already proved $k = 1$ case, we compute

$$D_{v_1} \ldots D_{v_k}\mathbb{1}_{a+E} = D_{v_1}\mathbb{1}_{a+E''} = \mathbb{1}_{a+E'},$$

where E' is the subspace $E' = \langle E, v_1, \ldots, v_k \rangle$. Otherwise, assume that $v_1, \ldots, v_k$ are linearly dependent in $\mathbb{F}_2^n/E$. If $v_2, \ldots, v_k$ are linearly dependent in $\mathbb{F}_2^n/E$, then since the statement is true for $k-1$, we have $D_{v_2} \ldots D_{v_k}\mathbb{1}_{a+E} = 0$. If $v_2, \ldots, v_k$ are linearly independent in $\mathbb{F}_2^n/E$, then again,

$$D_{v_2} \ldots D_{v_k}\mathbb{1}_{a+E} = \mathbb{1}_{a+E''},$$

where E'' is the subspace $E'' = \langle E, v_2, \ldots, v_k \rangle$. Since $v_1, \ldots, v_k$ are linearly dependent in $\mathbb{F}_2^n/E$, we deduce that $v_1 \in E''$, and from the $k = 1$ case, we have

$$D_{v_1}D_{v_2} \ldots D_{v_k}\mathbb{1}_{a+E} = D_{v_1}\mathbb{1}_{a+E''} = 0.$$

Hence, the statement is also true for k, and this concludes the proof. □

For the indicator of $\{0_n\}$, i.e. $\delta_0 \colon \mathbb{F}_2^n \to \mathbb{F}_2$, whose ANF is given by $\delta_0(x_1, \ldots, x_n) = \prod_{i=1}^n (x_i + 1)$, we get the following corollary.

Corollary 1. *For any two distinct nonzero vectors $a, b \in \mathbb{F}_2^n$, the algebraic degree of $D_aD_b\delta_0$ is $n-2$.*

Definition 4. *For a Boolean function $f \in \mathcal{B}_n$, we define its 2-rank as follows*

$$2\text{-rank}(f) := \text{rank}_{\mathbb{F}_2}\left(f(x+y)\right)_{x,y\in\mathbb{F}_2^n}. \tag{4}$$

In [23], it was shown that for a Boolean function f on $\mathbb{F}_2^n$ with $\deg(f) \geq 2$, the 2-rank is an invariant under EA-equivalence. Remarkably, using the notion of 2-rank, the authors also showed that any bent function $f \in \mathcal{B}_{2m}$ from the $\mathcal{M}^\#$ must satisfy the following necessary condition.

Theorem 2 (The $\mathcal{M}^{\#}$-Bound). *[23] Let $f \in \mathcal{B}_{2m}$ be bent such that $f \in \mathcal{M}^{\#}$. Then,* 2-rank$(f) \leq 2^{m+1} - 2$.

With the notion of 2-rank, Corollary 1 and Theorem 2, we are ready to prove the main result of this section.

Theorem 3. *For $m \geq 4$, the rotation-symmetric bent function $f \in \mathcal{B}_{2m}$ defined by (2) and its dual $\tilde{f} \in \mathcal{B}_{2m}$ defined by (3) are outside the $\mathcal{M}^{\#}$ class.*

Proof. By Lemma 2, is enough to show that only $f \notin \mathcal{M}^{\#}$, since $\tilde{f} \notin \mathcal{M}^{\#}$ if and only if $f \notin \mathcal{M}^{\#}$.

For $m = 2, 3$, the function f belongs to $\mathcal{M}^{\#}$ since all bent functions in 4 and 6 variables are members of the $\mathcal{M}^{\#}$ class, as already mentioned in the introduction. For $m = 4, 5, 6$, we compute the value of the 2-rank of f in Table 3, and compare it with the upper bound of the 2-rank of bent function in $2m$ variables in $\mathcal{M}^{\#}$, which is equal to $2^{m+1} - 2$ according to Theorem 2. As one can see from Table 3, we have that 2-rank$(f) > 2^{m+1} - 2$, from what follows that for these dimensions the function f is outside $\mathcal{M}^{\#}$.

Table 3. The value of 2-rank(f) for the bent function $f \in \mathcal{B}_{2m}$ defined by (2)

m	4	5	6
2-rank(f)	42	112	286
$\mathcal{M}^{\#}$-Bound	30	62	126

Now we proceed with the general case $m \geq 7$. Set $n = 2m$ and let E be an arbitrary m-dimensional subspace of $\mathbb{F}_2^n$. We will show that we can find two vectors $a, b \in E$ such that the algebraic degree of $D_a D_b f$ is $m - 2$.

We can find an (at least) $(m-4)$-dimensional subspace W of E such that for all $w = (w_0, \ldots, w_{n-1}) \in W$ we have $w_0 = w_{m-1} = w_m = w_{n-1} = 0$. To see this, take a basis B for E. If there are no elements w in the basis such that $w_0 = 0$, set $B_1 = B$. If $w_0 = 1$, for some elements of the basis, take one such element w, and add it to the other elements of the basis with the first coordinate equal to 1, and call the new basis B'. Set $B_1 = B' \setminus w$. In any case, the subspace generated by B_1 will have dimension at least $m - 1$, and all vectors in it will have the first coordinate equal to 0. Continue this process for the rest of the considered coordinates.

Define the mapping $L : W \to \mathbb{F}_2^{m-2}$ by $L(w_0, \ldots, w_{n-1}) = (w_1, \ldots, w_{m-2})$, for all $(w_0, \ldots, w_{n-1}) \in W$. Since L is linear, by the rank-nullity theorem, we have:

$$\dim(W) = \dim(Ker(L)) + \dim(Im(L)). \tag{5}$$

If $\dim(Im(L)) \geq 2$, there are two vectors in $a, b \in W$ such that $(a_1, \ldots, a_{m-2})$ and $(b_1, \ldots, b_{m-2})$ are two linearly independent vectors in $\mathbb{F}_2^{m-2}$. Consequently, from Corollary 1, we get that the algebraic degree of

$$D_a D_b(x_0 x_1 \cdots x_{m-2}\overline{x_m})$$

is $m-2$. Since a, b are in W, we have $a_0 = a_m = b_0 = b_m = 0$, hence every term of degree $m-2$ of $D_a D_b(x_0 x_1 \cdots x_{m-2}\overline{x_m})$ contains $x_0 x_m$. From the definition (2)

of f, we deduce that the only other term of f whose ANF has terms containing x_0x_m is $x_mx_{m+1}\cdots x_{2m-2}\overline{x_0}$. Hence, the only part of D_aD_bf that can cancel the $m-2$ degree terms of $D_aD_b(x_0x_1\cdots x_{m-2}\overline{x_m})$ is $D_aD_b(x_mx_{m+1}\cdots x_{2m-2}\overline{x_0})$. But every $m-2$ degree term of

$$D_aD_b(x_mx_{m+1}\cdots x_{2m-2}\overline{x_0})$$

contains $m-4$ variables from $x_{m+1},\ldots,x_{2m-2}$, and since $x_0x_1\cdots x_{m-2}\overline{x_m}$ does not contain those variables, the $m-2$ degree terms of $D_aD_b(x_mx_{m+1}\cdots x_{2m-2}\overline{x_0})$ cannot cancel the $m-2$ degree terms of $D_aD_b(x_0x_1\cdots x_{m-2}\overline{x_m})$. We conclude that, if $\dim(Im(L))\geq 2$, there are two vectors a,b in W such that the algebraic degree of D_aD_bf is $m-2$.

If $\dim(Im(L))\leq 1$, then from the equation (5) we get $\dim(Ker(L))\geq m-4-1=m-5\geq 2$, since $m\geq 7$. Let $a,b\in W$, be two linearly independent vectors in $Ker(L)$. From the definitions of W and L, we deduce that $(a_{m+1},\ldots,a_{2m-2})$ and $(b_{m+1},\ldots,b_{2m-2})$ are two linearly independent vectors in $\mathbb{F}_2^{m-2}$, and that the rest of the coordinates of a and b are equal to 0. From Corollary 1 we get that the algebraic degree of

$$D_aD_b(x_mx_{m+1}\cdots x_{2m-2}\overline{x_0})$$

is $m-2$. Similarly to the case $\dim(Im(L))\geq 2$, the only part of D_aD_bf that can cancel the $m-2$ degree terms of $D_aD_b(x_mx_{m+1}\cdots x_{2m-2}\overline{x_0})$ is

$$D_aD_b(x_0x_1\cdots x_{m-2}\overline{x_m}).$$

Since the corresponding coordinates of the vectors a and b are equal to 0, we have that $D_aD_b(x_0x_1\cdots x_{m-2}\overline{x_m})=0$. Consequently, the $m-2$ degree terms of $D_aD_b(x_mx_{m+1}\cdots x_{2m-2}\overline{x_0})$ in D_aD_bf are not canceled, hence the algebraic degree of D_aD_bf is $m-2$.

We conclude that we can always find $a,b\in W\subseteq E$, such that the algebraic degree of D_aD_bf is $m-2$. Since E was an arbitrary m-dimensional subspace of $\mathbb{F}_2^{2m}$, f is outside $\mathcal{M}^{\#}$ by Lemma 1. □

4 Conclusion

In this paper, we showed the existence of rotation-symmetric bent functions outside the $\mathcal{M}^{\#}$ class. This result indicates that just pursuing construction of bent functions with nice symmetries, it is possible to construct functions, not stemming from the well-known Maiorana-McFarland construction, opposite to many previously known results. Therefore, we think that in general it is an interesting research problem to construct bent functions in n variables symmetric w.r.t. the action of a certain given group $G<S_n$, since such functions can indeed induce functions that do not belong to the well-known powerful constructions.

Acknowledgements

Sadmir Kudin and Enes Pasalic are partly supported by the Slovenian Research Agency (research program P1-0404 and research projects N1-0159, J1-2451 and J1-4084).

References

1. Bapić, A., Pasalic, E., Polujan, A., Pott, A.: Vectorial Boolean functions with the maximum number of bent components beyond the Nyberg's bound. Designs, Codes and Cryptography (2023). p. 4.
2. Bosma, W., Cannon, J., Playoust, C.: The Magma algebra system. I. The user language. J. Symbolic Comput. **24**(3-4), 235–265 (1997). p. 4.
3. Canteaut, A., Charpin, P.: Decomposing bent functions. IEEE Transactions on Information Theory **49**(8), 2004–2019 (2003). p. 6.
4. Carlet, C.: Open Problems on Binary Bent Functions, pp. 203–241. Springer International Publishing, Cham (2014). pp. 1 and 3.
5. Carlet, C.: Boolean Functions for Cryptography and Coding Theory. Cambridge University Press (2021). p. 7.
6. Charnes, C., Rötteler, M., Beth, T.: Homogeneous bent functions, invariants, and designs. Designs, Codes and Cryptography **26**(1), 139–154 (2002). p. 4.
7. Chen, X., Su, S.: Constructions of rotation symmetric bent functions and bent idempotent functions. Advances in Mathematics of Communications (2023). p. 3.
8. Cusick, T.W., Stanica, P.: Cryptographic Boolean Functions and Applications (Second edition). Academic Press (2017). p. 4.
9. Dillon, J.F.: A survey of bent functions. NSA Technical Journal **Special Issue**, 191–215 (1972). p. 2.
10. Dillon, J.F.: Elementary Hadamard Difference Sets. Ph.D. thesis, University of Maryland (1974). p. 2.
11. Ding, C., Munemasa, A., Tonchev, V.D.: Bent vectorial functions, codes and designs. IEEE Transactions on Information Theory **65**(11), 7533–7541 (2019). p. 4.
12. Edel, Y., Pott, A.: On the equivalence of nonlinear functions. In: Enhancing Cryptographic Primitives with Techniques from Error Correcting Codes, pp. 87–103 (2009). p. 4.
13. Gao, G., Zhang, X., Liu, W., Carlet, C.: Constructions of quadratic and cubic rotation symmetric bent functions. IEEE Transactions on Information Theory **58**(7), 4908–4913 (2012). p. 3.
14. Mandal, B., Gangopadhyay, S., Stănică, P.: Cubic Maiorana-McFarland bent functions with no affine derivative. International Journal of Computer Mathematics: Computer Systems Theory **2**(1), 14–27 (2017). p. 6.
15. Pasalic, E., Bapić, A., Zhang, F., Wei, Y.: Explicit infinite families of bent functions outside the completed Maiorana-McFarland class. Designs, Codes and Cryptography **91**(7), 2365–2393 (2023). p. 6.
16. Polujan, A., Pott, A.: On design-theoretic aspects of Boolean and vectorial bent functions. IEEE Transactions on Information Theory **67**(2), 1027–1037 (2021). p. 4.
17. Polujan, A., Pott, A.: Towards the classification of quadratic vectorial bent functions in 8 variables. In: The 7th international workshop on Boolean functions and their applications (2022). p. 4.

18. Polujan, A.: Boolean and vectorial functions: A design-theoretic point of view. Ph.D. thesis, Otto-von-Guericke-Universität Magdeburg (2021). p. 4.
19. Polujan, A.A., Pott, A.: Cubic bent functions outside the completed Maiorana-McFarland class. Designs, Codes and Cryptography **88**(9), 1701–1722 (2020). pp. 2, 4, and 5.
20. Stănică, P., Maitra, S.: Rotation symmetric boolean functions—count and cryptographic properties. Discrete Applied Mathematics **156**(10), 1567–1580 (2008). pp. 3 and 4.
21. Su, S.: A new construction of rotation symmetric bent functions with maximal algebraic degree. Advances in Mathematics of Communications **13**(2), 253–265 (2019). pp. 1, 3, and 6.
22. Tang, C., Qi, Y., Zhou, Z., Fan, C.: Two infinite classes of rotation symmetric bent functions with simple representation. Applicable Algebra in Engineering, Communication and Computing **29**(3), 197–208 (Jun 2018). p. 3.
23. Weng, G., Feng, R., Qiu, W.: On the ranks of bent functions. Finite Fields and Their Applications **13**(4), 1096–1116 (2007). pp. 8 and 9.
24. Zhao, Q., Zheng, D., Zhang, W.: Constructions of rotation symmetric bent functions with high algebraic degree. Discrete Applied Mathematics **251**, 15–29 (2018). pp. 1 and 3.
25. Zhou, J., Li, N., Zeng, X., Yunge, X.: A generic construction of rotation symmetric bent functions. Advances in Mathematics of Communications **15**(4), 721–736 (2021). p. 3.

On the maximum weight codewords of linear rank-metric codes

Olga Polverino, Paolo Santonastaso, and Ferdinando Zullo

Università degli Studi della Campania "Luigi Vanvitelli"
{olga.polverino,paolo.santonastaso,ferdinando.zullo}@unicampania.it

Abstract. Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be a linear rank-metric code. In this paper, we investigate the problem of determining the number $M(\mathcal{C})$ of codewords in $\mathcal{C}$ with maximum weight, that is $\min\{m, n\}$, and to characterize codes attaining the maximum value for $M(\mathcal{C})$.

Keywords: rank-metric codes · weight distribution · q-system

1 Introduction

Rank-metric codes have attracted a lot of attention recently since their numerous applications and interesting mathematical connections. The origin of rank-metric codes dates back to Delsarte [8] in 1978, some years later they were rediscovered by Gabidulin in [9] and Roth in [18]. In the next, we mainly focus on *linear* codes. The **rank (weight)** $w(v)$ of a vector $v = (v_1, \dots, v_n) \in \mathbb{F}_{q^m}^n$ is the dimension of the vector space generated over $\mathbb{F}_q$ by its entries, i.e, $w(v) = \dim_{\mathbb{F}_q}(\langle v_1, \dots, v_n \rangle_{\mathbb{F}_q})$.

A **(linear vector) rank metric code** $\mathcal{C}$ is an $\mathbb{F}_{q^m}$-subspace of $\mathbb{F}_{q^m}^n$ endowed with the rank distance defined as

$$d(x, y) = w(x - y),$$

for any $x, y \in \mathbb{F}_{q^m}^n$.

If $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ is a rank-metric code, we write that $\mathcal{C}$ is an $[n, k, d]_{q^m/q}$ code (or $[n, k]_{q^m/q}$ code) if k is the $\mathbb{F}_{q^m}$-dimension of $\mathcal{C}$ and d is its minimum distance, that is

$$d = \min\{d(x, y) \colon x, y \in \mathcal{C}, x \neq y\}.$$

The parameters of a rank-metric code are related by a Singleton-like bound.

Theorem 1. *[8] Let $\mathcal{C}$ be an $[n, k, d]_{q^m/q}$ code. Then*

$$mk \leq \max\{m, n\}(\min\{n, m\} - d + 1). \tag{1}$$

An $[n, k, d]_{q^m/q}$ code is called **Maximum Rank Distance code** (or shortly **MRD code**) if its parameters attains the bound (1). We say that two $[n, k]_{q^m/q}$

codes $\mathcal{C}$ and $\mathcal{C}'$ are **equivalent** if $\mathcal{C}' = \mathcal{C}\,A = \{vA : v \in \mathcal{C}\}$, for some element $A \in \mathrm{GL}(n,q)$. In the next, without losing of generality, we only consider *non-degenerate* codes, i.e. codes that cannot be isometrically embedded in a smaller space. More precisely, an $[n,k]_{q^m/q}$ rank-metric code $\mathcal{C}$ is said to be **non-degenerate** if the columns of any generator matrix of $\mathcal{C}$ are $\mathbb{F}_q$-linearly independent.

Our goal is to provide information on the number of codewords of maximum weight of a code. For some classes of rank metric codes, the weight distribution is known, such as for MRD codes or classes of few weight codes, but in general very few is known and giving information on the weight distribution is hard. For non-degenerate rank-metric codes, in [2, Proposition 3.11] it is established the existence of at least one codeword of maximum weight. Maximum weight codewords are also intriguing in connection with the rank-metric version of the Critical problem by Crapo and Rota (cf. [3], and also [11]). This interest is further heightened due to the connection with q-polymatroids, as explored in [10]. For an $[n,k]_{q^m/q}$ code $\mathcal{C}$, we define $M(\mathcal{C})$ as the number of its codewords with weight $\min\{m,n\}$. In this next, we investigate the following two problems:

Problem 1. To determine upper and lower bounds on $M(\mathcal{C})$.

Problem 2. To characterize the extremal cases in the obtained bounds on $M(\mathcal{C})$.

To address such problems, we mainly employ tools from combinatorics: we use the projective version of systems, namely linear sets, which are point sets in projective spaces.

The paper is structured as follows. In Section 2, we describe the geometric correspondence between rank-metric codes and systems/linear sets. Then we deal with upper and lower bounds on $M(\mathcal{C})$ using geometric arguments. Section 3 is devoted to the case of equality in the upper bounds: the geometry in this case is either related to canonical subgeometries or to linear sets with minimum size.
Part of the results are taken from [16], whereas the last part is original and it has been developed by the second author.

2 Geometric interpretation of the number of maximum weight codewords

2.1 Geometric description of rank-metric codes

The geometric counterpart of rank-metric codes are the systems, see [2, 17].

Definition 1. *An $[n,k,d]_{q^m/q}$ **system** U is an $\mathbb{F}_q$-subspace of $\mathbb{F}_{q^m}^k$ of dimension n, such that $\langle U\rangle_{\mathbb{F}_{q^m}} = \mathbb{F}_{q^m}^k$ and*

$$d = n - \max\left\{\dim_{\mathbb{F}_q}(U \cap H) \mid H \text{ is an } \mathbb{F}_{q^m}\text{-hyperplane of } \mathbb{F}_{q^m}^k\right\}.$$

Moreover, two $[n,k,d]_{q^m/q}$ *systems* U *and* U' *are* ***equivalent*** *if there exists an invertible matrix* $A \in \mathrm{GL}(k, \mathbb{F}_{q^m})$ *such that*

$$UA = U'.$$

Rank-metric codes and systems are related in the following way. Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code and G be an its generator matrix. Then the $\mathbb{F}_q$-subspace U obtained as the $\mathbb{F}_q$-span of the columns of G is called a **system associated with** $\mathcal{C}$. Viceversa, let U be an $[n,k]_{q^m/q}$ system. Define G as the matrix whose columns are an $\mathbb{F}_q$-basis of U and let $\mathcal{C}$ be the code generated by G. $\mathcal{C}$ is called a **code associated with** U.

It is possible to prove that two codes $\mathcal{C}$ and $\mathcal{C}'$ are equivalent if and only if their associated systems are equivalent.

Moreover, we recall how the support of a codeword is related to the intersections with a system associated with the code.

Let $G \in \mathbb{F}_{q^m}^{k\times n}$ such that its columns are $\mathbb{F}_q$-linearly independent and let U be the $\mathbb{F}_q$-span of the columns of G. Define the map

$$\begin{aligned} \psi_G : \mathbb{F}_q^n &\longrightarrow U \\ \lambda &\longmapsto \lambda G^\top, \end{aligned}$$

which turns out to be an $\mathbb{F}_q$-linear isomorphism.

Let $c = (c_1, \ldots, c_n) \in \mathbb{F}_{q^m}^n$ and $\Gamma = (\gamma_1, \ldots, \gamma_m)$ be an ordered $\mathbb{F}_q$-basis of $\mathbb{F}_{q^m}$. The **(rank) support** of c is defined as the column span of $\Gamma(c)$, where $\Gamma(c) \in \mathbb{F}_q^{n\times m}$, is the matrix defined by

$$c_i = \sum_{j=1}^{m} \Gamma(c)_{ij}\gamma_j, \qquad \text{for all } i \in \{1, \ldots, n\}.$$

As proved in [2, Proposition 2.1], the support does not depend on the choice of Γ and we can talk about the support of a vector without mentioning Γ.

Theorem 2. *([17] and [14, Theorem 3.1]) Let* $\mathcal{C}$ *be a non-degenerate* $[n,k,d]_{q^m/q}$ *code and let* G *be a generator matrix. Let* $U \subseteq \mathbb{F}_{q^m}^k$ *be the* $\mathbb{F}_q$*-span of the columns of* G*. Then, for every* $x \in \mathbb{F}_{q^m}^k$

$$\psi_G^{-1}(U \cap x^\perp) = \mathrm{supp}(xG)^\perp,$$

where $\mathrm{supp}(xG)^\perp$ *denotes the orthogonal complement of* $\mathrm{supp}(xG)$ *with respect to the standard scalar product in* $\mathbb{F}_q^n$ *and* $x^\perp$ *denotes the orthogonal complement of* $\langle x\rangle_{\mathbb{F}_{q^m}}$ *with respect to the standard scalar product in* $\mathbb{F}_{q^m}^k$*. In particular, the rank weight of an element* $xG \in \mathcal{C}$*, with* $x = (x_1, \ldots, x_k) \in \mathbb{F}_{q^m}^k$ *is*

$$w(xG) = n - \dim_{\mathbb{F}_q}(U \cap x^\perp). \tag{2}$$

As a consequence,

$$d = n - \max\left\{\dim_{\mathbb{F}_q}(U \cap H) \colon H \text{ is an } \mathbb{F}_{q^m}\text{-hyperplane of } \mathbb{F}_{q^m}^k\right\}. \tag{3}$$

The above argument allows to estabilish a one-to-one correspondence between equivalence classes of $[n,k,d]_{q^m/q}$ systems and $[n,k,d]_{q^m/q}$ codes.

Generalized rank weights have been introduced several times with different definitions, see e.g. [12], and they have been used also as a tool for the inequivalence of families of codes as was done in [4].

We will deal with the definition given in [17] and more precisely to the equivalent one given in [2, Theorem 3.14], directly connected with the systems.

Definition 2. *Let $\mathcal{C}$ be a $[n,k,d]_{q^m/q}$ rank metric code and let U be an associated system. For any $r \in \{1,\dots,k\}$, the r-**th generalized rank weight** is*

$$d_r^{\mathrm{rk}}(\mathcal{C}) = n - \max\left\{\dim_{\mathbb{F}_q}(U \cap H) \colon H \text{ is an } \mathbb{F}_{q^m}\text{-subspace of codim. } r \text{ of } \mathbb{F}_{q^m}^k\right\}. \tag{4}$$

Note that when $r = 1$, in the above definition we obtain the minimum distance.

For our aims, it will often be useful to look at the systems projectively via the notion of linear sets. Let $V = V(k, q^m)$ be a k-dimensional vector space over $\mathbb{F}_{q^m}$ and let $\Lambda = \mathrm{PG}(V, \mathbb{F}_{q^m}) = \mathrm{PG}(k-1, q^m)$. For an $\mathbb{F}_q$-subspace U of V of dimension n, the set of points

$$L_U = \{\langle u \rangle_{\mathbb{F}_{q^m}} : u \in U \setminus \{0\}\} \subseteq \Lambda$$

is said to be an $\mathbb{F}_q$-**linear set of rank** n. Let $\Omega = \mathrm{PG}(W, \mathbb{F}_{q^m})$ be a projective subspace of Λ. The **weight of** Ω in L_U is defined as

$$w_{L_U}(\Omega) = \dim_{\mathbb{F}_q}(U \cap W).$$

We have the following upper bound on the number of points of a linear set:

$$|L_U| \leq \frac{q^n - 1}{q - 1}.$$

Moreover, if $L_U \neq \emptyset$, then

$$|L_U| \equiv 1 \pmod{q}, \tag{5}$$

and if $\langle L_U \rangle = \mathrm{PG}(k-1, q^m)$ then

$$|L_U| \geq \frac{q^k - 1}{q - 1}. \tag{6}$$

The above bound can be improved if some assumptions are added.

Theorem 3. *([7, Theorem 1.2] and [5, Lemma 2.2]) If L_U is an $\mathbb{F}_q$-linear set of rank n, with $1 < n \leq m$ on $\mathrm{PG}(1, q^m)$, and L_U contains at least one point of weight 1, then $|L_U| \geq q^{n-1} + 1$.*

Recently, extending the results in [7], in [1] more lower bounds on the size of a linear set have been proved, see [1].

2.2 Geometric dual of a linear set and of a rank-metric code

We recall the notion of the dual of an $\mathbb{F}_q$-subspace of the k-dimensional $\mathbb{F}_{q^m}$-vector space $V = V(k, q^m)$. Let $\sigma\colon V \times V \to \mathbb{F}_{q^m}$ be a nondegenerate reflexive bilinear form defined on V and consider

$$\begin{aligned} \sigma' : V \times V &\longrightarrow \mathbb{F}_q \\ (x, y) &\longmapsto \mathrm{Tr}_{q^m/q}(\sigma(x, y)). \end{aligned}$$

In this way, σ' turns out to be a nondegenerate reflexive bilinear form on V seen as an $\mathbb{F}_q$-vector space of dimension km. Consider $\perp$ and $\perp'$ as the orthogonal complement maps defined by σ and σ', respectively. For an $\mathbb{F}_q$-subspace U of V, the **dual subspace of** U with respect to σ' is the $\mathbb{F}_q$-subspace $U^{\perp'}$ in V. The definition of the dual of an $\mathbb{F}_q$-subspace in V does not depend on the choice of σ. Indeed, if $\perp'_1$ and $\perp'_2$ are orthogonal complement maps as above, then the dual subspaces $U^{\perp'_1}$ and $U^{\perp'_2}$ of U are $\mathrm{GL}(k, q^m)$-equivalent, cf. [15, Proposition 2.5].

Moreover, for any $\mathbb{F}_q$-subspace U of $V = V(k, q^m)$, we have that $\dim_{\mathbb{F}_q}(U^{\perp'}) = km - \dim_{\mathbb{F}_q}(U)$ and for any $\mathbb{F}_{q^m}$-subspace of V holds $W^{\perp'} = W^{\perp}$. As a consequence, the following relation holds.

Proposition 1. *[15, Property 2.6] Let U be an $\mathbb{F}_q$-subspace of V and W be an $\mathbb{F}_{q^m}$-subspace of V. Then*

$$\dim_{\mathbb{F}_q}(U^{\perp'} \cap W^{\perp}) = \dim_{\mathbb{F}_q}(U \cap W) + \dim_{\mathbb{F}_q}(V) - \dim_{\mathbb{F}_q}(U) - \dim_{\mathbb{F}_q}(W).$$

For more details on duality operation, see also [13]. By using the dual of a subspace, recently in [6], an operation has been introduced on rank metric codes called *geometric dual*. This operation takes an $[n, k, d]_{q^m/q}$ code and gives another code with parameters $[mk - n, k, d']_{q^m/q}$.

Definition 3. *Let $\mathcal{C}$ be a non-degenerate $[n, k, d]_{q^m/q}$ and let U be a system associated with $\mathcal{C}$. Suppose also that $d^{\mathrm{rk}}_{k-1}(\mathcal{C}) \geq n - m + 1$. Then a **geometric dual** $\mathcal{C}^{\perp_{\mathcal{G}}}$ of $\mathcal{C}$ (with respect to $\perp'$) is defined as $\mathcal{C}'$, where $\mathcal{C}'$ is any code associated with the system $U^{\perp'}$, where $\perp'$ is defined as before.*

This definition is justified by the following result.

Theorem 4. *[6, Theorem 3.4] Let $\mathcal{C}$ be an $[n, k, d]_{q^m/q}$ code, and let U be a system associated with $\mathcal{C}$. Suppose also that $d^{\mathrm{rk}}_{k-1}(\mathcal{C}) \geq n - m + 1$. Then, up to equivalence, a geometric dual $\mathcal{C}^{\perp_{\mathcal{G}}}$ of $\mathcal{C}$ does not depend on the choice of the associated system and on the choice of a code in the equivalence class of $\mathcal{C}$, hence $\perp_{\mathcal{G}}$ is well-defined.*

2.3 General bounds on $M(\mathcal{C})$

We provide upper and lower bounds on $M(\mathcal{C})$, by using geometric arguments and bounds on the size of a linear set.

We start by describing the geometric meaning of $M(\mathcal{C})$.

Proposition 2. *[16, Proposition 3.6 and Proposition 3.12] Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code and let U be any associated system. Then, if $n \leq m$,*

$$M(\mathcal{C}) = (q^m-1)|\{H \text{ hyperplane of } \mathrm{PG}(k-1,q^m)\colon H \cap L_U = \emptyset\}|;$$

and if $n > m$,

$$\begin{aligned} M(\mathcal{C}) &= (q^m-1)|\{H \text{ hyperplane of } \mathrm{PG}(k-1,q^m)\colon w_{L_U}(H) = n-m\}| \\ &= (q^m-1)|\mathrm{PG}(k-1,q^m) \setminus L_{U^{\perp'}}|. \end{aligned}$$

Proof. Let G be a generator matrix of $\mathcal{C}$ such that the $\mathbb{F}_q$-span of its columns is U. Assume first that $n \leq m$. Then by (2), a codeword $c = uG$ has maximum weight n if and only if

$$w_{L_U}(u^\perp) = n - w(uG) = 0,$$

and hence the assertion follows. If $n > m$, again by (2), a codeword $c = uG$ has maximum weight if and only if

$$w_{L_U}(u^\perp) = n - w(uG) = n - m,$$

and hence $M(\mathcal{C}) = (q^m-1)|\{H \text{ hyperplane of } \mathrm{PG}(k-1,q^m)\colon w_{L_U}(H) = n-m\}|$. The second one follows by considering the linear sets associated with $U^{\perp'}$ and by applying Proposition 1.

First, we concentrate on the case $n \leq m$. By using some geometric results, cf, [16, Lemma 3.7 and Lemma 3.8], we are able to prove the following.

Theorem 5. *[16, Theorem 3.9 and Theorem 3.11] Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code. Assume that $n \leq m$. Then*

$$\frac{q^{mk}-1}{q^m-1} - \frac{q^n-1}{q-1}\frac{q^{(k-1)m}-1}{q^m-1} + q\beta \leq \frac{M(\mathcal{C})}{q^m-1} \leq \prod_{i=1}^{n-1}(q^m-q^i) \tag{7}$$

where

$$\beta = \frac{q^{mk}-1}{q^m-1} - \prod_{i=1}^{k-1}(q^m-q^i) - \frac{q^k-1}{q-1}\prod_{i=1}^{k-2}(q^m-q^i).$$

Moreover, if $n-e$ is the second maximum weight of $\mathcal{C}$. Then

$$q^{m(k-1)} - q^{m(k-2)+n-e} - q^{m(k-2)}\left(\frac{q^{n-e}-1}{q^e-1}\right) \leq \frac{M(\mathcal{C})}{q^m-1} \leq q^{m(k-1)} - q^{m(k-2)+n-e},$$

i.e., $m(k-2)+n = \lfloor \log_q(q^{m(k-1)} - \frac{M(\mathcal{C})}{q^m-1}) \rfloor + e$.

The latter bound of the above theorem depends on the second maximum weight and the possible values of $M(\mathcal{C})$ are in disjoint intervals (according to e). Moreover, once the parameters m, n, q of the code $\mathcal{C}$ are known, by using the value $M(\mathcal{C})$, then one can directly determine the second largest weight $n-e$ of the code $\mathcal{C}$.

We can now derive bounds on $M(\mathcal{C})$ in the case $n > m$ by making use of the bounds on the number of points of linear sets, cf. (3).

Theorem 6. *[16, Theorem 3.14] Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be an $[n,k]_{q^m/q}$ code. Assume that $m \leq n$ and $d_{k-1}^{\mathrm{rk}}(\mathcal{C}) \geq n-m+1$. Then*

$$\frac{q^{km}-1}{q^m-1} - \frac{q^{km-n}-1}{q-1} \leq \frac{M(\mathcal{C})}{q^m-1} \leq \frac{q^{km}-1}{q^m-1} - \frac{q^k-1}{q-1}. \tag{8}$$

In particular, if the second maximum weight of $\mathcal{C}$ is $m-e$,

$$\frac{q^{km}-1}{q^m-1} - \frac{q^{km-n}-1}{q^e-1} \leq \frac{M(\mathcal{C})}{q^m-1} \leq \frac{q^{km}-1}{q^m-1} - \left(q^{km-n-e} + \frac{q^{k-1}-1}{q-1}\right), \tag{9}$$

i.e. $km-n = \lfloor \log_q(\frac{q^{mk}-1}{q^m-1} - \frac{M(\mathcal{C})}{q^m-1}) \rfloor + e$.

As for the case $n \leq m$, the possible values of $M(\mathcal{C})$ are in disjoint intervals, according to the second largest weight $m-e$ of the code.

The above upper bound (9) can be improved with a more involved condition. This result relies on recent bounds on the size of linear sets proved in [1].

Theorem 7. *[16, Theorem 3.16] Let $\mathcal{C}$ be a $[n,k]_{q^m/q}$ code and assume that $m \leq n$ and $d_{k-1}^{\mathrm{rk}}(\mathcal{C}) \geq n-m+1$. Let G' be any of generator matrix of $\mathcal{C}^{\perp_\mathcal{G}}$. Suppose there exist $r \geq 1$ codewords $c_1, \ldots, c_r \in \mathcal{C}^{\perp_\mathcal{G}}$ $\mathbb{F}_{q^m}$-linearly independent such that the $\mathbb{F}_q$-subspace*

$$W = \psi_{G'}\left(\bigcap_{i=1}^{r} \mathrm{supp}(c_i)^\perp\right)$$

satisfies $\dim_{\mathbb{F}_q}(W) = \dim_{\mathbb{F}_{q^m}}(\langle W \rangle_{\mathbb{F}_{q^m}}) = k-r$. Then

$$\frac{M(\mathcal{C})}{q^m-1} \leq \frac{q^{km}-1}{q^m-1} - \left(q^{km-n-1} + \ldots + q^{km-n-k+r} + \frac{q^r-1}{q-1}\right). \tag{10}$$

Remark 1. In [16], there are proved some refinements of the above bounds when $k=2$.

3 Classification results based on maximum weight codewords

In this section, we study codes attaining the upper bound on $M(\mathcal{C})$, also providing classification results.

3.1 Equality in the upper bounds

Without conditions on the weight distribution on $\mathcal{C}$, the maximum for $M(\mathcal{C})$ is assumed if and only if either $\mathcal{C}$ or its geometric dual is the entire space.

Theorem 8. *[16, Theorem 5.1] Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code and assume that $d^{\mathrm{rk}}_{k-1}(\mathcal{C}) \geq n-m+1$.*

- *If $n < m$ then $M(\mathcal{C})$ is maximum with respect to (7) if and only if $n = k$ and $\mathcal{C} = \mathbb{F}_{q^m}^k$.*
- *If $n \geq m$ then $M(\mathcal{C})$ is maximum with respect to (8) if and only if $n = mk-k$ and $\mathcal{C}^{\perp_{\mathcal{G}}} = \mathbb{F}_{q^m}^k$.*

We can also characterize the case of equality in Theorem 7, when $r > 1$.

Proposition 3. *[16, Proposition 5.2] Let $\mathcal{C}$ be a non-degenerate $[n,k]_{q^m/q}$ code and assume that $m \leq n$ and $d^{\mathrm{rk}}_{k-1}(\mathcal{C}) \geq n-m+1$. Let G' be any of generator matrix of $\mathcal{C}^{\perp_{\mathcal{G}}}$. Suppose there exist $r > 1$ codewords $c_1, \ldots, c_r \in \mathcal{C}^{\perp_{\mathcal{G}}}$ $\mathbb{F}_{q^m}$-linearly independent such that*

$$W = \psi_{G'}\left(\bigcap_{i=1}^{r} \mathrm{supp}(c_i)^{\perp}\right)$$

satisfies $\dim_{\mathbb{F}_q}(W) = \dim_{\mathbb{F}_{q^m}}(\langle W \rangle_{\mathbb{F}_{q^m}}) = k - r$ and

$$M(\mathcal{C}) = q^{km} - 1 - (q^m - 1)\left(q^{km-n} + \ldots + q^{km-n-k+r} + \frac{q^r - 1}{q-1}\right).$$

Then $n = mk - k$ and $\mathcal{C}^{\perp_{\mathcal{G}}} = \mathbb{F}_{q^m}^k$.

When $r = 1$, we have a different scenario that we will describe in the next. We start by describing some constructions for codes satisfying the assumptions of Theorem 7 with $r = 1$.

Let $\lambda \in \mathbb{F}_{q^m} \setminus \mathbb{F}_q$ be an element generating $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$ and

$$G = \begin{pmatrix} 1 & \lambda & \ldots & \lambda^{t_1-1} & 0 & \ldots & & & & 0 \\ 0 & 0 & \ldots & 0 & 1 & \lambda & \ldots & \lambda^{t_2-1} & 0 & \ldots & 0 \\ \vdots & & & & & & & & \ddots & & \\ 0 & \ldots & & & & & & & 0 & 1 & \ldots & \lambda^{t_k-1} \end{pmatrix} \in \mathbb{F}_{q^m}^{k \times (t_1+\ldots+t_k)}. \quad (11)$$

Let $\mathcal{C}_{\lambda,t_1,\ldots,t_k}$ be the $\mathbb{F}_{q^m}$-linear rank metric code in $\mathbb{F}_{q^m}^{t_1+\ldots+t_k}$ having G as a generator matrix.

We now determine the parameters of these codes.

Theorem 9. *[16, Theorem 5.4] Let $\lambda \in \mathbb{F}_{q^m} \setminus \mathbb{F}_q$ be an element generating $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$ and let $\mathcal{C}_{\lambda,t_1,\ldots,t_k}$ be with $t_1 \leq t_2 \leq \ldots \leq t_k \leq m-1$. Then $\mathcal{C}_{\lambda,t_1,\ldots,t_k}$ is an $[t_1+\ldots+t_k, k, t_1]_{q^m/q}$ code.*

Under certain conditions on the parameters, the codes $\mathcal{C}_{\lambda,t_1,\ldots,t_k}$ satisfies the assumptions of Theorem 7 with $r = 1$ and reaches the maximum for $M(\mathcal{C}_{\lambda,t_1,\ldots,t_k})$ among the $[n,k]_{q^m/q}$ codes satisfying the assumptions of Theorem 7.

Theorem 10. *[16, Theorem 5.8] Let $\lambda \in \mathbb{F}_{q^m} \setminus \mathbb{F}_q$ be an element generating $\mathbb{F}_{q^m}$. Let consider the code $\mathcal{C}_{\lambda,t_1,\ldots,t_k} \subseteq \mathbb{F}_{q^m}^n$, where $n = t_1 + \ldots + t_k$, $m \leq n$ and $km - n \leq m + k$. Let $\{m - t_1 - 1, \ldots, m - t_k - 1\} = \{s_{i_1}, \ldots, s_{i_\ell}\}$, with $s_{i_1} > \ldots > s_{i_\ell}$. Then, if either*

$$k \leq \sum_{j=1}^{\ell} \frac{q^{s_{i_j}} - 2q^{s_{i_j}/2}}{s_{i_j}}$$

or

$$mk - k - t_1 - \ldots - t_k \leq q,$$

the code $\mathcal{C}_{\lambda,t_1,\ldots,t_k}$ satisfies the assumptions of Theorem 7 with $r = 1$ and reaches the maximum for $M(\mathcal{C}_{\lambda,t_1,\ldots,t_k})$ among the $[n,k]_{q^m/q}$ codes satisfying the assumptions of Theorem 7 with $r = 1$.

3.2 Characterization results

The code $\mathcal{C}_{\lambda,t_1,\ldots,t_k}$ has a very nice generator matrix in (11): this description says that the code admits a basis whose supports are in direct sum. When a code satisfy this property, $\mathcal{C}$ is the direct sum of k one-dimensional rank-metric codes. We will see in the next result that a code of dimension k is the direct sum of k one-dimensional rank-metric codes if and only if it is of type $(t_1, \ldots, t_k)$, that is it admits a basis $c_1, \ldots, c_k \in \mathcal{C}$ such that $t_i = w(c_i)$ for any i and

$$t_1 + \ldots + t_k = n.$$

Theorem 11. *Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code. Then the following are equivalent:*

1) *$\mathcal{C}$ is of type $(t_1, \ldots, t_k)$;*
2) *a generator matrix of $\mathcal{C}$ is of the form*

$$G = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1t_1} & 0 & \ldots & & & & 0 \\ 0 & 0 & \ldots & 0 & a_{21} & a_{22} & \ldots & a_{2t_2} & 0 & \ldots & 0 \\ \vdots & & & & & & & & \ddots & & \\ 0 & \ldots & & & & & & & 0 & a_{k1} & \ldots & a_{kt_k} \end{pmatrix} \in \mathbb{F}_{q^m}^{k \times n},$$

that is $\mathcal{C}$ is the direct sum of k rank-metric codes.

Proof. 1) $\Rightarrow$ 2). Assume that $\mathcal{C}$ is of type $(t_1, \ldots, t_k)$. This means that there exist k codewords $c_1, \ldots, c_k \in \mathcal{C}$ such that $t_i = w(c_i)$ for any i and

$$t_1 + \ldots + t_k = n.$$

Let G' be a generator matrix for $\mathcal{C}$ having as rows the $c_i's$. Therefore,

$$c_1 = e_1 G, \ldots, c_k = e_k G \in \mathbb{F}_{q^m}^n,$$

where the e_i's are the vectors of the standard basis of $\mathbb{F}_{q^m}^k$. By using (2), we have that

$$\dim_{\mathbb{F}_q}(U \cap e_i^{\perp}) = n - w(c_i) \text{ and } \sum_{i=1}^{k} \dim_{\mathbb{F}_q}(U \cap e_i^{\perp}) = kn - n.$$

Let consider the dual subspace $U^{\perp'} \subseteq \mathbb{F}_{q^m}^k$ of U. Then $\dim_{\mathbb{F}_q}(U^{\perp'}) = km - n$ and by Proposition 1, we get

$$\sum_{i=1}^{k} \dim_{\mathbb{F}_q}(U^{\perp'} \cap \langle e_i \rangle_{\mathbb{F}_{q^m}}) = km - n.$$

Hence,

$$U^{\perp'} = W_1 \times W_2 \times \ldots \times W_k,$$

for some $W_1, \ldots, W_k$ $\mathbb{F}_q$-subspaces of $\mathbb{F}_{q^m}$, implying that

$$U = U_1 \times U_2 \times \ldots \times U_k,$$

where $U_i = \{a \in \mathbb{F}_{q^m} : \mathrm{Tr}_{q^m/q}(ab) = 0, \text{ for any } b \in W_i\}$ and $\dim_{\mathbb{F}_q}(U_i) = t_i$.
$2) \Rightarrow 1)$. This implication follows by the definition of a code being of type $(t_1, \ldots, t_k)$.

We point out that this approach can be further pursued to get new constructions of rank-metric codes with a large number of maximum weight codewords. This is part of an ongoing project.

References

1. S. Adriaensen and P. Santonastaso. On the minimum size of linear sets. *arXiv preprint math/2301.13001*, 2023.
2. G. N. Alfarano, M. Borello, A. Neri, and A. Ravagnani. Linear cutting blocking sets and minimal codes in the rank metric. *Journal of Combinatorial Theory, Series A*, 192:105658, 2022.
3. G. N. Alfarano and E. Byrne. The critical theorem for q-polymatroids. *arXiv preprint arXiv:2305.07567*, 2023.
4. D. Bartoli, G. Marino, and A. Neri. New MRD codes from linear cutting blocking sets. *Annali di Matematica Pura ed Applicata (1923-)*, pages 1–28, 2022.
5. G. Bonoli and O. Polverino. $\mathbb{F}_q$-linear blocking sets in $\mathrm{PG}(2, q^4)$. *Innovations in Incidence Geometry: Algebraic, Topological and Combinatorial*, 2(1):35–56, 2005.
6. M. Borello and F. Zullo. Geometric dual and sum-rank minimal codes. *arXiv preprint arXiv:2303.07288 to appear in* Journal of Combinatorial Designs *https://doi.org/10.1002/jcd.21934*, 2023.
7. J. De Beule and G. Van de Voorde. The minimum size of a linear set. *Journal of Combinatorial Theory, Series A*, 164:109–124, 2019.
8. P. Delsarte. Bilinear forms over a finite field, with applications to coding theory. *Journal of Combinatorial Theory, Series A*, 25(3):226–241, 1978.

9. E. M. Gabidulin. Theory of codes with maximum rank distance. *Problemy Peredachi Informatsii*, 21(1):3–16, 1985.
10. H. Gluesing-Luerssen and B. Jany. q-polymatroids and their relation to rank-metric codes. *Journal of Algebraic Combinatorics*, pages 1–29, 2022.
11. A. Gruica, A. Ravagnani, J. Sheekey, and F. Zullo. Rank-metric codes, semifields, and the average critical problem. *SIAM Journal on Discrete Mathematics*, 37(2):1079–1117, 2023.
12. R. Jurrius and R. Pellikaan. On defining generalized rank weights. *Advances in Mathematics of Communications*, 11(1):225–235, 2017.
13. G. Lunardon, G. Marino, O. Polverino, and R. Trombetti. Translation dual of a semifield. *Journal of Combinatorial Theory, Series A*, 115(8):1321–1332, 2008.
14. A. Neri, P. Santonastaso, and F. Zullo. The geometry of one-weight codes in the sum-rank metric. *Journal of Combinatorial Theory, Series A*, 194:105703, 2023.
15. O. Polverino. Linear sets in finite projective spaces. *Discrete Mathematics*, 310(22):3096–3107, 2010.
16. O. Polverino, P. Santonastaso, and F. Zullo. Maximum weight codewords of a linear rank metric code. *arXiv preprint arXiv:2302.00979*, 2023.
17. T. H. Randrianarisoa. A geometric approach to rank metric codes and a classification of constant weight codes. *Designs, Codes and Cryptography*, 88:1331–1348, 2020.
18. R. M. Roth. Maximum-rank array codes and their application to crisscross error correction. *IEEE Transactions on Information Theory*, 37(2):328–336, 1991.

Characterization of Some Non-Canonical Minihypers in PG$(r, 3)$ and the Main Problem of Coding Theory

Assia Rousseva[1], Ivan Landjev[2,3], and Emiliyan Rogachev[1]

[1] *Faculty of Mathematics and Informatics, Sofia University, 5 J. Bourchier blvd., 1164 Sofia, Bulgaria*
assia@fmi.uni-sofia.bg, rogachev@uni-sofia.bg

[2] *Bulgarian Academy of Sciences, Institute of Mathematics and Informatics, 8 Acad G. Bonchev str., 1113 Sofia, Bulgaria* ivan@math.bas.bg

[3] *New Bulgarian University, 21 Montevideo str., 1618 Sofia, Bulgaria*
i.landjev@nbu.bg

1 Introduction

The problem of finding the shortest length of an $\mathbb{F}_q$-linear code of fixed dimension k and fixed minimum distance d (denoted by $n_q(k,d)$) is known as the main problem in coding theory (cf. [2]). Codes with parameters $[n,k,d]_q$, where $n = n_q(k,d)$, are said to be optimal (with respect to the length). There exists a natural lower bound on $n_q(k,d)$ – the so-called Griesmer bound:

$$n_q(k,d) \geq \sum_{i=0}^{k-1} \lceil \frac{d}{q^i} \rceil. \tag{1}$$

The right-hand side in the above inequality is usually denoted by $g_q(k,d)$. Linear $[n,k,d]_q$ codes of length $n = g_q(k,d)$ are called Griesmer codes. It is known that for fixed k and q Griesmer codes do exist for all sufficiently large minimum distances. One possible approach to the main problem of coding theory is to determine the exact value of $n_q(k,d)$ for fixed q and k for all d.

For ternary linear codes the exact value of d is known for all $k \leq 5$ and all d. This has been done during the years by various authors (see [13] and the references there). For $k = 6$ the value of $n_3(6,d)$ is known for all but 70 values of d [13].

In this paper we characterize certain minihypers in PG$(r,3)$, $r \leq 5$. Based on the obtained results we rule out the existence of certain hypothetical Griesmer codes with $q = 3$, $k = 6$. As a by-product we prove several general results on minihypers that turn out to be important in the investigation of the main problem of coding theory.

2 Preliminaries

We start with some basic definitions and facts on linear codes and multisets of points in the geometries PG$(k-1,q)$. Since we prefer to keep the letter k for

the dimension of the linear codes, the multisets associated with them will be contained in the $(k-1)$-dimensional projective geometry.

A multiset in $\mathrm{PG}(k-1,q)$ is a mapping $\mathcal{K}\colon \mathcal{P} \to \mathbb{N}_0$, from the pointset $\mathcal{P}$ of $\mathrm{PG}(k-1,q)$ to the non-negative integers. For a subset $\mathcal{Q}$ of $\mathcal{P}$, we define $\mathcal{K}(\mathcal{Q}) = \sum_{P\in\mathcal{Q}} \mathcal{K}(P)$. The integer $\mathcal{K}(\mathcal{Q})$ is called the multiplicity of the subset $\mathcal{Q}$. A point of multiplicity i is called an i-point. Similarly, i-lines, i-planes, i-solids are points, lines, 3-dimensional subspaces of multiplicity i, $i = 0, 1, \ldots$. The integer $\mathcal{K}(\mathcal{P})$ is called the cardinality of the multiset $\mathcal{K}$. A multiset $\mathcal{K}$ in $\mathrm{PG}(k-1,q)$ is called an (n,w)-arc, if: (a) $\mathcal{K}(\mathcal{P}) = n$; (b) $\mathcal{K}(H) \leq w$ for each hyperplane H in $\mathrm{PG}(k-1,q)$, and (c) there is a hyperplane H_0 with $\mathcal{K}(H_0) = w$. In a similar way, we define an (n,w)-minihyper (or (n,w)-blocking set) as a multiset $\mathcal{K}$ in $\mathrm{PG}(k-1,q)$ satisfying: (d) $\mathcal{K}(\mathcal{P}) = n$; (e) $\mathcal{K}(H) \geq w$ for each hyperplane H in $\mathrm{PG}(k-1,q)$, and (f) there is a hyperplane H_0 with $\mathcal{K}(H_0) = w$. The existence of an $[n,k,d]_q$-code C of full length (no coordinate identically zero) is equivalent to that of a $(n, n-d)$-arc in $\mathrm{PG}(k-1,q)$. From any generator matrix G of C one can define a multiset $\mathcal{K}$ with points (with the corresponding multiplicities) the columns of G. This correspondence between $[n,k,d]_q$ codes and $(n,n-d)$-arcs maps isomorphic codes to projectively equivalent arcs and vice versa. The same correspondence maps $[n,k,\delta]_q$-anticodes to $(n,n-\delta)$-minihypers in $\mathrm{PG}(k-1,q)$.

Given an (n,w)-arc $\mathcal{K}$ in $\mathrm{PG}(k-1,q)$, we denote by $\gamma_i(\mathcal{K})$ the maximal multiplicity of an i-dimensional flat in $\mathrm{PG}(k-1,q)$, i.e. $\gamma_i(\mathcal{K}) = \max_\delta \mathcal{K}(\delta)$, $i = 0, \ldots, k-1$, where δ runs over all i-dimensional flats in $\mathrm{PG}(k-1,q)$. If $\mathcal{K}$ is clear from the context we shall write just γ_i. If $\mathcal{K}$ is a (n,w)-arc in $\mathrm{PG}(k-1,q)$ with a maximal point multiplicity γ_0 then $c\,\mathrm{PG}(k-1,q) - \mathcal{K}$ is a $(\gamma_0 v_k - |\mathcal{K}|, \gamma_0 v_{k-1} - w)$-minihyper in $\mathrm{PG}(k-1,q)$, where $v_k = (q^k-1)/(q-1)$.

The integer $\Delta > 1$ is a called a divisor of the linear code C if the weight of every word in C is divisible by Δ. In what follows, we repeatedly make use of the following result [14].

Theorem 1 *[14] Let C be an $[n,k,d]$-code over $\mathbb{F}_p$, p a prime, meeting the Griesmer bound. If $p^e | d$, then p^e is a divisor of C.*

Geometrically, this can be stated as follows.

Theorem 2 *Let $\mathcal{K}$ be a Griesmer (n,w)-arc in $\mathrm{PG}(k-1,p)$, p prime, with $w \equiv n \pmod{p^e}$, $e \geq 1$. Then $\mathcal{K}(H) \equiv n \pmod{p^e}$ for every hyperplane H.*

An $[n,k,d]_q$-code C is called extendable if there exists an $[n+1,k,d+1]_q$ code C' such that C can be obtained from C' by puncturing. An (n,w)-arc $\mathcal{K}$ in $\mathrm{PG}(k-1,q)$ is called extendable if there exists an $(n+1,w)$-arc $\mathcal{K}'$ in $\mathrm{PG}(k-1,q)$ with $\mathcal{K}'(x) \geq \mathcal{K}(x)$ for every point of $\mathrm{PG}(k-1,q)$. Clearly, extendable arcs are associated with extendable codes.

The next extension result about arcs is the geometric version of Hill-Lizak's result for codes [4,5]. Below we state their result in coding-theoretic and geometric form.

Theorem 3 *Let C be an $[n,k,d]_q$-code with $\gcd(n-w,q)=1$ and with all weights congruent to 0 or d modulo q. Then C can be extended to an $[n+1,k,d+1]_q$-code all of whose weights are congruent to 0 or $d+1$ modulo q.*

Theorem 4 *Let $\mathcal{K}$ be an (n,w)-arc in $\mathrm{PG}(k-1,q)$ with $\gcd(n-w,q)=1$. Assume that the multiplicities of all hyperplanes are congruent to n or w modulo q. Then $\mathcal{K}$ can be extended to an $(n+1,w)$-arc.*

The following result by Hitoshi Kanda [7] is slightly different and concerns codes over $\mathbb{F}_3$.

Theorem 5 *[7] Let $\mathcal{C}$ be an $[n,k,d]_3$ code with $(d,3)=1$ whose possible weights of codewords satisfy $A_i=0$ for all $i \not\equiv 0,-1,-2 \pmod 9$. Then $\mathcal{C}$ is extendable.*

In what follows we shall need also the following result which was proved by many authors in a weaker form.

Theorem 6 *[10] If $x \le q-q/p$ then every (xv_t, xv_{t-1})-minihyper in $\mathrm{PG}(t,q)$ is a sum of x hyperplanes.*

3 Two Theorems on Canonical Minihypers

Let d and k be positive integers and let d be written in the following form:

$$d = sq^{k-1} - \lambda_{k-2}q^{k-2} - \ldots - \lambda_1 q - \lambda_0, \tag{2}$$

where $0 \le \lambda_i \le q-1$. It is easily checked that

$$g_q(k,d) = sv_k - \lambda_{k-2}v_{k-1} - \ldots \lambda_1 v_2 - \lambda_0 v_1. \tag{3}$$

The existence of a Grismer code of dimension k and minimum distance is equivalent to that of a minihyper in $\mathrm{PG}(k-1,q)$ with parameters $\left(\sum_{i=0}^{k-2}\lambda_i v_{i+1}, \sum_{i=0}^{k-2}\lambda_i v_i\right)$ and a maximal point multiplicity s. It should be noted that a minihyper with the above parameters can always be constructed as the sum of λ_{k-2} hyperplanes, λ_{k-3} hyperlines, and so on, λ_1 lines, and λ_0 points. Minihypers constructed in this way are called canonical.

In many cases, the following two theorems turn out to be useful.

Theorem 7 *Assume that every minihyper with parameters*

$$\left(\sum_{i=0}^{k-2}\lambda_i v_{i+1}, \sum_{i=0}^{k-2}\lambda_i v_i\right)$$

in $\mathrm{PG}(k-1,q)$ is canonical. Then every minihyper with parameters

$$\left(\sum_{i=0}^{k-2}\mu_i v_{i+1}, \sum_{i=0}^{k-2}\mu_i v_i\right)$$

in $\mathrm{PG}(k-1,q)$, where $\mu_i \le \lambda_i$, is also canonical.

Theorem 8 *Assume that every minihyper with parameters*

$$\left(\sum_{i=0}^{k-2}\lambda_i v_{i+1}, \sum_{i=0}^{k-2}\lambda_i v_i\right)$$

in $\mathrm{PG}(k-1,q)$ *is canonical. Then for every* $r \geq k-1$ *every minihyper in* $\mathrm{PG}(r,q)$ *with the same parameters is also canonical.*

4 Characterization of some Minihypers in PG(3, 3) and PG(4, 3)

All minihypers in the next two sections have a maximal point multiplicity of two, because our ultimate goal is to prove the non-existence of certain hypothetical Griesmer codes that imply this property. It is not difficult, however, to obtain a characterization also without this restriction.

The first four theorems contain characterizations of some minypers in $\mathrm{PG}(3,3)$ that are almost straightforward, but are needed for Theorem 13 and the nonexistence proofs in the next section.

Theorem 9 *A* $(21,6)$*-blocking set in* $\mathrm{PG}(3,3)$ *is one of the following:*

(a) the sum of a plane and two lines;
(b) the sum of a plane and a plane $(9,2)$*-blocking set which is the complement of an oval;*
(c) a blocking set with $\lambda_2 = 1$*,* $a_{12} = 2$ *(see [8]);*
(d) a blocking set with $\lambda_2 = 0$*,* $a_{12} = 1$ *(see [8]).*

Theorem 10 *Every* $(22,6)$*-minihyper in* $\mathrm{PG}(3,3)$ *is either reducible, or the sum of a plane and a plane* $(9,4)$*-minihyper.*

Theorem 11 *A* $(30,9)$*-minihyper in* $\mathrm{PG}(3,3)$ *is one of the following:*

(a) the sum of two planes and a line;
(b) the complement of a cap.

Theorem 12 *A* $(39,12)$*-minihyper in* $\mathrm{PG}(3,3)$ *is the sum of a plane* $(12,3)$*-minihyper and a three-dimensional affine space.*

The $(12,3)$-minihypers in $\mathrm{PG}(2,3)$ are obtained as complements of $(14,5)$-arcs that are known [1].

The next two theorems contain characterizations in minihypers in $\mathrm{PG}(4,3)$. The details of the proofs are contained in [9].

Theorem 13 *Let* $\mathcal{B}$ *be a* $(66,21)$*-blocking set in* $\mathrm{PG}(4,3)$*. Then* $\mathcal{B}$ *is one of the following:*

(a) the sum of a solid and two planes;

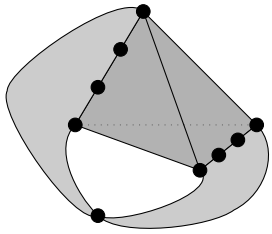


Fig. 1. $(66, 21)$-blocking set of type (a) without 3-points

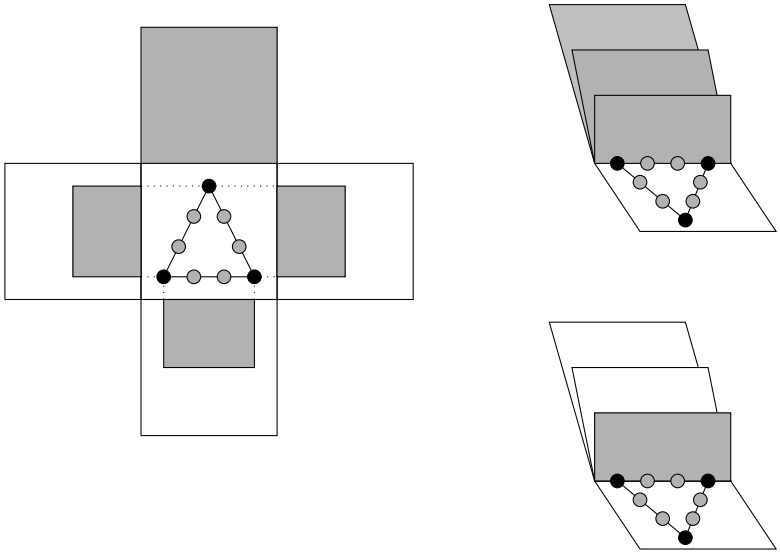


Fig. 2. $(66, 21)$-blocking set of type (b)

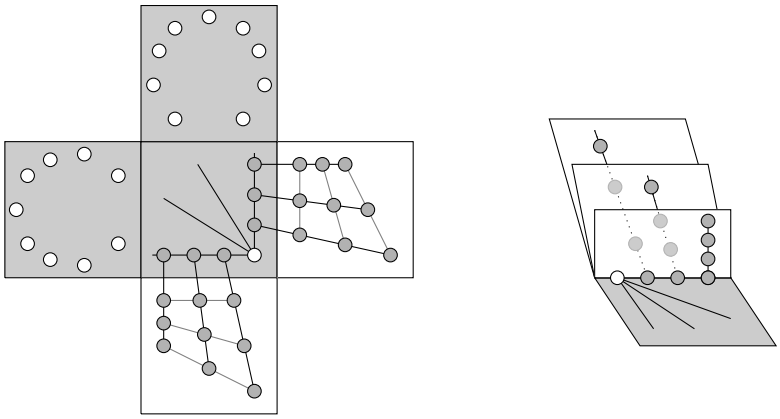


Fig. 3. $(66, 21)$-blocking set of type (c)

(b) the sum of an affine space of dimension 3 and three affine planes contained in the four solids through a common 12-plane which is the sum of three (not necessarily different lines);
(c) the dual of the $(11, 5)$*-arc in* $\mathrm{PG}(4, 3)$

The minihypers from Theorem 13 are presented in the figures below.

Using the extension theorem of Kanda [7], one can prove that every $(68, 21)$-minihyper is reducible to a $(66, 21)$-minihyper.

Theorem 14 *Every* $(68, 21)$*-minihyper in* $\mathrm{PG}(4, 3)$ *is reducible to a* $(66, 21)$*-minihyper.*

5 The non-existence of some Griesmer codes

In this section, we sketch the proofs for the non-existence of certain minihypers in $\mathrm{PG}(5, 3)$. This implies the non-existence of several hypothetical Griesmer codes and leads to improvements in Maruta's tables with exact values of $n_3(6, d)$ [11].

Theorem 15 *There exists no minihyper with parameters* $(207, 67)$ *in* $\mathrm{PG}(3, 5)$ *and with maximal point multiplicity 2. Consequently, there exists no* $[521, 6, 346]_3$ *codes and* $n_3(6, 346) = 522$*.*

Proof. Let $\mathcal{K}$ be a $(207, 67)$-minihyper in $\mathrm{PG}(3, 5)$ with maximal point multiplicity 2. Fix a 4-dimensional subspace Δ_0 of multiplicity 67 and a 21-solid S in Δ_0. Denote by $\Delta_1, \Delta_2, \Delta_3$ the other three 4-dimensional subspaces through S. We have two possibilities:

(A) $\mathcal{K}(\Delta_0) = \mathcal{K}(\Delta_1) = \mathcal{K}(\Delta_2) = 67$, $\mathcal{K}(\Delta_3) = 69$;
(B) $\mathcal{K}(\Delta_0) = \mathcal{K}(\Delta_1) = 67$, $\mathcal{K}(\Delta_2) = \mathcal{K}(\Delta_3) = 68$.

Let us note at first that if there exists a 4-dimensional subspace of multiplicity ≥ 148 then $\mathcal{K}$ is the sum of a hyperplane and a $(86, 27)$-minihyper in $\mathrm{PG}(5, 3)$, which, in turn, is the sum of two solids and a plane $(6, 1)$-minihyper. Now it can be proved that a point of multilpicity 3 cannot be avoided. From now on we shall assume that all hyperplanes have multiplicity <148.

Since the solids in a minimal hyperplane have multiplicity $\equiv 0, 1 \pmod 3$, the admissible multiplicities of a hyperplane are the following:

$$67, \ldots, 72, 94, \ldots, 99, 121, \ldots 126.$$

(B) Let us select Δ_0 to be a 67-hyperplane reducible to a 66-hyperplane of type (a) or (b). Note that there is always such a hyperplane. Select a 12-plane π such that $\mathcal{K}|_\pi$ is a triangle (a sum of three non-concurrent lines). Consider a projection φ from π onto some plane disjoint from π. The image of a 67- or a 68-plane is of type

$$(1)\ \ (18+\varepsilon_1, 18+\varepsilon_2, 9+\varepsilon_3, 9),\ \text{or}\ (2)\ \ (27+\varepsilon_1, 9+\varepsilon_2, 9+\varepsilon_3, 9),$$

where ε_i are non-negative integers with $\varepsilon_1+\varepsilon_2+\varepsilon_3 = 1$ or 2. From now on points of multiplicity $9+\varepsilon, 18+\varepsilon$, or $27+\varepsilon$ will be called $(27+)-, (18+)$ or $(9+)$-points respectively.

If a point in the projection plane is of multiplicity $9+\varepsilon$, $\varepsilon \in \{0, 1, 2\}$, then it is the image of a solid of multiplicity $21+\varepsilon$ that has a plane without 0-points meeting π in one of the sides of the triangle. Denote the three sides of triangle in π by α, β, and γ. With each point in the projection plane of multiplicity 9 we can associate a letter α, β, or γ depending on the line in which the full plane meets the triangle It follows from Theorem 13 that if the image of a hyperplane is a line of type $(18+\varepsilon_1, 18+\varepsilon_2, 9+\varepsilon_3, 9)$ the letters assigned to the two 9's are the same; for a line of type $(27+\varepsilon_1, 9+\varepsilon_2, 9+\varepsilon_3, 9)$ the three letters should be different. Without loss of generality we assign to $\varphi(\pi)$ the letter α.

Apart from this it is easily noted that the projection plane does not contain a line of type $(18+\varepsilon_0, 9+\varepsilon_1, 9+\varepsilon_2, 9+\varepsilon_3)$ since it would be the image of a hyperplane of multiplicity $57+\sum_i \varepsilon_i \leq 63$. Thus we have to rule out five possibilities for the projection $\mathcal{K}^\varphi$: these are the cases where x of the lines $\varphi(\Delta_i)$ are of the first and $4-x$ – of the second type.

Assume $x = 0$. Then the $(27+)$-points have to be collinear. Since all $(9+)$-points different from $\varphi(\pi)$ are assigned β and γ, there exists a line of multiplicity ≤ 68 and type (2) which not assigned all three letters, a contradiction. If $x = 1, 2$, or 4, there exists a line in the projection plane that is of multiplicity ≤ 51, again a contradiction.

It remains to consider the case $x = 3$. The three $(9+)$-points should be collinear, otherwise there is a line of multiplicity ≤ 51. The fourth point on this line should be a $(27+)$-point. Moreover this line should be the image of a hyperplane of multiplicity ≥ 69. Now there exists a line of type (1) which is the image of a hyperplane of multiplicity 67 or 68 and which is assigned two different letters, again a contradiction.

Now we are left with the case when all 67-hyperplanes are of the type (c), which is easily ruled out by considering a projection from a 12-plane contained in a 21-solid.

Theorem 16 *There exists no minihyper with parameters* $(209, 68)$ *in* $\mathrm{PG}(3, 5)$ *and with maximal point multiplicity 2. Consequently, there exists no* $[519, 6, 345]_3$ *codes and* $n_3(6, 345) = 520$.

Proof. Note that in this case the minihyper is divisible cince the complementary arc is divisible (Theorem 2). Hence the only possible hyperplane multiplicities are 68, 95, and 122. (Larger hyperplanes force the existence of 3-points as Theorem 15.) If there exists a hyperplane containing a minihyper of the type (a) or (b), we can proceed as in the proof of Theorem 15. Hence we have to deal with the case when all hyperplanes are of the type described in Theorem 13(c).

Consider a 68-plane of type (c) and fix a 12-plane π contained in a 21-solid S in this hyperplane. We can always select such a plane. Consider a projection from the 12-plane. Then the images of the four hyperplanes through S have type $(18 + \varepsilon_1, 18 + \varepsilon_2, 9 + \varepsilon_3, 9)$, where $\sum_i \varepsilon_i = 2$. Now there is a line in the plane of projection that is of multiplicity at most 53. Hence it is the image of a hyperplane of multiplicity at most 65, a contradiction.

Corollary 17 *Every* $(210, 68)$*-minihyper in* $\mathrm{PG}(5, 3)$ *is reducible. Consequently, there exists no* $[518, 6, 344]_3$*-code and* $n_3(6, 344) = 519$.

Proof. This can be done by using Hill-Lizak's extension theorem in its minihyper version. Hyperplanes of multiplicity ≥ 148 are ruled out since they necessarily contain a 3-point. Hence the possible hyperplane multiplicities are

$$68, \ldots, 75, 94, \ldots, 102, 121, \ldots, 127.$$

It remains to rule out those with multiplicity $\equiv 1 \pmod 3$. Then the minihyper would be reducible to the nonexistent $(209, 68)$-minihyper by Hill-Lizak's theorem.

The obtained results are summarized in the following table. The boldfaced entries are the results announced in this contribution.

d	$g_3(6, d)$	$n_3(6, d)$	$\mathcal{B}$	$\mathcal{B}\|_H$
343	517	517–518	$(211, 68)$	$(68, 21)$
344	518	**519**	$(210, 68)$	$(68, 21)$
345	519	**520**	$(209, 68)$	$(68, 21)$
346	521	**522**	$(207, 67)$	$(67, 21)$

Acknowledgments

The first and the third author were supported by the Bulgarian National Science Research Fund under Grant KP-06-N72/6-2023.

References

1. S. Ball, R. Hill, I. Landjev, H. Ward, On $(q^2+q+2, q+2)$-arcs in the projective plane PG$(2,q)$, *Des. Codes Cryptogr.* 24(2001), 205–224.
2. R. Hill, Optimal linear codes, Cryptography and Coding (ed. Ch. Mitchell), Oxford University Press, 1992, 75–104.
3. J.W.P. Hirschfeld, *Projective Geometries over Finite Fields*, Oxford University Press, 2nd edition, 1998.
4. R. Hill, An extension theorem for linear codes, *Des. Codes and Crypt.* 17(1999), 151–157.
5. R. Hill, P. Lizak, Extensions of linear codes, in: Proc. Int. Symp. on Inf. Theory, Whistler, BC, Canada 1995.
6. T. Honold, I. Landjev, On maximal arcs in projective Hjelmslev planes of even characteristic, *Finite Fields and Their Appl.* **11**(2005), 292-304.
7. H. Kanda, A new extension theorem for ternary linear codes and its application, *Finite Fields and Their Appl.* 64(2020), 101711 https://doi.org/10.1016/j-ffa.2020.101711
8. I. Landjev, A. Rousseva, E. Rogachev, On a Class of Minihypers in the Geometries PG(r,q), LNICST 450, Proceedings of the 18th EAI Int. Conf on CSECS (T. Zlateva and R. Goleva eds.) 2022, 142–153, 2022. https://doi.org/10.1007/978 − 3 − 031 − 17292 − $2_1$2
9. I. Landjev, A. Rousseva, E. Rogachev, Characterization of Some Minihypers of PG$(4,3)$, *Annual of Sofia University*, Faculty Math and Inf. 109(2022), 91-98. DOI: 10.60063/GSU.FMI.109.1-8
10. I. Landjev, P. Vandendriessche, A note on (xv_t, xv_{t-1})-minihypers in PG(t,q), *Journal of Combinatorial Theory* Ser. A 119(2012), 1123–1131. doi: 10.1016/j-jcta.2012.02.009
11. T. Maruta, Griesmer Bound for Linear Codes over Finite Fields, https://mars39.lomo.jp/opu/griesmer.htm
12. A. Rousseva, I. Landjev, Linear codes close to the Griesmer bound and the related geometric structures, *Designs, Codes and Cryptography* 87(2019), 841–854 https://doi.org/10.1007/s10623-018-0565-3
13. T. Sawashima, T. Maruta, Nonexistence of some ternary linear codes with minimum weight -2 modulo 9, *Adv. in Math. of Comm.* 17(6)(2023), 1338–1357. doi: 10.394/amc.2021052
14. H.N. Ward, Divisibility of codes meeting the Griesmer bound, *J. Combin. Theory* Ser. A, **83**(1998), 79–93.

Bounds on Sphere Sizes in the Sum-rank Metric and Coordinate-additive Metrics

Hugo Sauerbier Couvée[1], Thomas Jerkovits[2], and Jessica Bariffi[2,3]

[1] Department of Electrical and Computer Engineering, Technical University of Munich, Germany.
[2] Institute of Communication and Navigation, German Aerospace Center, Germany.
[3] Institute of Mathematics, University of Zurich, Switzerland

Abstract. This paper provides new bounds on the size of spheres in any coordinate-additive metric with a particular focus on improving existing bounds in the sum-rank metric. We derive improved upper and lower bounds based on the entropy of a distribution related to the Boltzmann distribution, which work for any coordinate-additive metric. Additionally, we derive new closed-form upper and lower bounds specifically for the sum-rank metric that outperform existing closed-form bounds.

Keywords: Sum-rank metric · Coordinate-additive metric · Sphere size · Combinatorics · Coding theory · Information theory

1 Introduction

The sum-rank metric [15], Hamming metric [7] and Lee metric [10] are examples of coordinate-additive metrics. Codes with distance properties in such metrics are of particular interest in various applications, such as linear network coding [12], quantum-resistant cryptography [8,17], coding for storage [13], space-time coding [19]. Bounds on the size of an ℓ-dimensional ball or sphere in such metrics are essential for deriving bounds like the sphere-packing bound or the Gilbert–Varshamov bound [4]. An information-theoretic approach for bounding the volume of an ℓ-dimensional ball concerning any coordinate-additive metric, via the entropy of an auxiliary probability distribution, was presented in [11]. Specifically addressing the sum-rank metric, closed-form upper and lower bounds for the sphere size were introduced in [17,16] and further discussed in [6]. However, these bounds are limited in their tightness, particularly noticeable in scenarios involving smaller sizes of the base field q and/or a larger number of blocks ℓ.

The exact value for the size of an ℓ-dimensional sphere $\mathcal{S}_t^\ell$ of radius t in any coordinate-additive metric can be derived by computing all its (ordered) integer partitions, where each part of the partition has at most a part size of the maximal possible weight in the corresponding metric. These will represent the decomposition of the nonzero entries of the elements in the sphere. To get the size of the sphere we sum over all integer partitions adding up the number of elements that have a weight decomposition corresponding to the integer partition. Although this procedure provides the exact value of $\left|\mathcal{S}_t^\ell\right|$, it often doesn't give an

intuitive or practical understanding of the sphere size or how this size changes as the parameters change. For large parameters it is even impractical to compute the size in this way. Hence, the derivation of closed-form bounds on the exact formula are of major interest. A current method of obtaining both upper and lower bounds on $\left|\mathcal{S}_t^\ell\right|$ is, for instance, to consider only the partition attaining the maximum number of elements. This approach is utilized by [16,17,6]. Another method is to bound the size of an ℓ-dimensional ball $\mathcal{B}_t^\ell$ of radius t, since clearly every upper bound on $\left|\mathcal{B}_t^\ell\right|$ is a valid upper bound on $\left|\mathcal{S}_t^\ell\right|$, too. On a complex analytic side, sizes of spheres and balls can be described using generating functions, whose coefficients can be computed using the saddle-point technique and other techniques from analytic combinatorics (see [2,3]). We refer to [18] for a more detailed discussion and proofs of the results presented in this paper.

2 Preliminaries

Let q be a prime power and denote by $\mathbb{F}_q$ the finite field of q elements. The natural numbers $\mathbb{N}$ shall include 0. Given a random variable X over a finite alphabet $\mathcal{A}$ with probability distribution P, we define $P(a) := \mathrm{Prob}(X = a)$ with $a \in \mathcal{A}$. The entropy $H(P)$ of P with respect to the base q is defined as $H(P) := -\sum_{a\in\mathcal{A}, P(a)\neq 0} P(a) \log_q P(a)$.

2.1 Coordinate-Additive Metrics

Let $(\mathcal{A}, +)$ be a finite abelian group with identity element 0 called the **alphabet**. We define a **weight function** $\mathrm{wt}_\mathcal{A} : \mathcal{A} \to \mathbb{N}$ on $\mathcal{A}$ to be a function satisfying for all $a, b \in \mathcal{A}$:

1. $\mathrm{wt}_\mathcal{A}(a) = 0$ if and only if $a = 0$,
2. $\mathrm{wt}_\mathcal{A}(a) = \mathrm{wt}_\mathcal{A}(-a)$,
3. $\mathrm{wt}_\mathcal{A}(a + b) \leq \mathrm{wt}_\mathcal{A}(a) + \mathrm{wt}_\mathcal{A}(b)$.

This function can be extended to a **coordinate-additive weight function** on the cartesian product $\mathcal{A}^\ell$ (with group structure inherited coordinate-wise from $\mathcal{A}$) by defining the weight of an ℓ-tuple to be the sum of the weights of its coordinates, i.e., $\mathrm{wt}_{\Sigma\mathcal{A}}(a_1, \ldots, a_\ell) = \sum_{i=1}^{\ell} \mathrm{wt}_\mathcal{A}(a_i)$. This coordinate-additive weight function naturally induces a **metric** $\mathrm{d}_{\Sigma\mathcal{A}} : \mathcal{A}^\ell \times \mathcal{A}^\ell \to \mathbb{N}$ as $\mathrm{d}_{\Sigma\mathcal{A}}(v, w) := \mathrm{wt}_{\Sigma\mathcal{A}}(v-w)$. Given a coordinate-additive weight function $\mathrm{wt}_{\Sigma\mathcal{A}}$ on $\mathcal{A}^\ell$, we define the ℓ-dimensional **sphere**, respectively **ball**, of radius $t \in \mathbb{N}$ by

$$\mathcal{S}_t^\ell := \{v \in \mathcal{A}^\ell \ : \ \mathrm{wt}_{\Sigma\mathcal{A}}(v) = t\} \quad \text{and} \quad \mathcal{B}_t^\ell := \{v \in \mathcal{A}^\ell \ : \ \mathrm{wt}_{\Sigma\mathcal{A}}(v) \leq t\}.$$

For the special case of the sum-rank metric, let m, η and ℓ be positive integers. Also define $\mu := \min\{m, \eta\}$ and $n := \eta\ell$. We write $\mathbb{F}_q^{m\times\eta\ell}$ for the space of $m\times(\eta\ell)$ matrices over the finite field $\mathbb{F}_q$. Every matrix $M \in \mathbb{F}_q^{m\times\eta\ell}$ is represented as a sequence of ℓ blocks, i.e., $M = (B_1 \,|\, B_2 \,|\, \ldots \,|\, B_\ell)$ with each $B_i \in \mathbb{F}_q^{m\times\eta}$. The **sum-rank weight** of a matrix $M \in \mathbb{F}_q^{m\times\eta\ell}$ is defined as $\mathrm{wt}_{SR}(M) :=$

$\sum_{i=1}^{\ell} \mathrm{rk}_q(B_i)$ where $\mathrm{rk}_q(B_i)$ is the rank of B_i over $\mathbb{F}_q$. Analogously, we define for every $0 \leq t \leq \mu \cdot \ell$, the **sum-rank sphere** of radius t as

$$\mathcal{S}_t^{m,\eta,\ell,q} := \{M \in \mathbb{F}_q^{m \times \eta\ell} \ : \ \mathrm{wt}_{SR}(M) = t\}.$$

For fixed m, η, q, ℓ, the sum-rank sphere sizes $\left|\mathcal{S}_t^{m,\eta,\ell,q}\right|$ can be computed with a dynamic program described in [17].

2.2 Ordinary Generating Functions

The theory of ordinary generating functions (OGFs) is a useful branch of mathematics that lays connections between combinatorics, analysis, number theory, probability theory and other fields. In this paper we restrict ourselves to OGFs corresponding to weights in coordinate-additive metrics, which are polynomials with non-negative coefficients. Consider a finite abelian group $\mathcal{A}$ with weight function $\mathrm{wt}_{\mathcal{A}}$ and induced coordinate-additive weight function $\mathrm{wt}_{\Sigma\mathcal{A}}$ on $\mathcal{A}^\ell$. The OGF corresponding to $\mathrm{wt}_{\Sigma\mathcal{A}}$ is defined as the polynomial

$$F_{\mathcal{A}^\ell}(z) := \textstyle\sum_{v \in \mathcal{A}^\ell} z^{\mathrm{wt}_{\Sigma\mathcal{A}}(v)} = \sum_{i=0}^{\mu\ell} |\mathcal{S}_i^\ell| \ z^i.$$

The OGF for $\mathcal{A} = \mathcal{A}^1$ is denoted by $F_{\mathcal{A}}(z)$. For a polynomial $F(z) = F_0 + F_1 z + \ldots + F_d z^d$ we use the notation $[z^i]F(z)$ to refer to the i-th coefficient F_i of $F(z)$, with $[z^i]F(z) = 0$ for $i > \deg(F)$. The OGF for the sum-rank metric on $\mathbb{F}_q^{m \times \eta\ell}$ is denoted by $\mathcal{S}^{m,\eta,\ell,q}(z) = \sum_{i=0}^{\mu\ell} |\mathcal{S}_i^{m,\eta,\ell,q}| \ z^i$.

Definition 1 (Partial order on polynomials). *Let $F(z), G(z) \in \mathbb{R}[z]$ be two real polynomials. If $[z^i]F(z) \leq [z^i]G(z)$ for all $i \in \mathbb{N}$, we say $F(z)$ is coefficient-wise less-than-or-equal to $G(z)$, denoted as $F(z) \preccurlyeq_c G(z)$.*

Proposition 1 ([2, Theorem I.1]). *Let $\mathcal{A}_1$, $\mathcal{A}_2$ be two finite alphabets with weight functions $\mathrm{wt}_{\mathcal{A}_1}, \mathrm{wt}_{\mathcal{A}_2}$ respectively. Then $\mathrm{wt}_{\mathcal{A}_1 \times \mathcal{A}_2}(a, b) := \mathrm{wt}_{\mathcal{A}_1}(a) + \mathrm{wt}_{\mathcal{A}_2}(b)$ is a weight function on $\mathcal{A}_1 \times \mathcal{A}_2$ and*

$$F_{\mathcal{A}_1 \times \mathcal{A}_2}(z) = F_{\mathcal{A}_1}(z) F_{\mathcal{A}_2}(z).$$

In particular, we have $F_{\mathcal{A}^\ell}(z) = F_{\mathcal{A}}(z)^\ell$, for $\ell \in \mathbb{N}$. Furthermore, the product of real polynomials with non-negative coefficients preserves the partial order: if $F(z) \preccurlyeq_c G(z)$ and $K(z) \preccurlyeq_c L(z)$, then $F(z)K(z) \preccurlyeq_c G(z)L(z)$.

Lemma 1. *Let $F(z)$ be a real polynomial of degree $d > 0$ with non-negative coefficients $F_i \geq 0$ and first derivative $F'(z)$. If $F(z)$ is not a monomial, then the function $G(z) = zF'(z)/F(z)$ is a strictly increasing smooth function on the positive reals $\mathbb{R}_{>0}$. In particular if $F(0) > 0$, which is the case with OGFs of finite alphabets with weight functions, $G(z)$ is a bijection from $[0, \infty)$ to $[0, d)$.*

Proof. Smoothness follows directly from smoothness of $F(z)$ and $1/z$ on $\mathbb{R}_{>0}$. Setting $K(a, b) := bF'(b)F(a) - aF'(a)F(b)$ with $0 < a < b$, we can show that $K(a, b) > 0$, thereby proving $G(z)$ is strictly increasing. Lastly, we have that $\lim_{z\to\infty} F'(z)/z^{d-1} = dF_d$ and $\lim_{z\to\infty} F(z)/z^d = F_d$, so $\lim_{z\to\infty} G(z) = d$. □

3 Information-Theoretic Bounds on Spheres

In [11] an asymptotically tight upper bound on the volume of an ℓ-dimensional ball $\left|\mathcal{B}_t^\ell\right|$ of radius t was introduced. This bound is valid for any arbitrary additive weight function $\mathrm{wt}_\mathcal{A}$ with respect to some finite abelian group $\mathcal{A}$ as described in Section 2.1. The bound was proved to hold for normalized weights ρ with $\rho := t/\ell$ up to the average weight $\overline{w} := |\mathcal{A}|^{-1} \sum_{a\in\mathcal{A}} \mathrm{wt}_\mathcal{A}(a)$ at which the volume of the ball is saturated. We extend the result from [11] to the size of spheres and also prove that the bound holds for $\rho \geq \overline{w}$ up to the maximum possible weight, i.e. $0 < \rho < \mu$ with $\mu := \max_{a\in\mathcal{A}} \{\mathrm{wt}_\mathcal{A}(a)\}$. Note that this notation coincides with $\mu = \min\{m, \eta\}$ for the sum-rank metric. For any $a \in \mathcal{A}$, $\ell \in \mathbb{N}$ and $0 < \rho < \mu$, we define the probability distribution

$$P_\beta(a) := \frac{q^{-\beta\,\mathrm{wt}_\mathcal{A}(a)}}{\mathcal{Z}(\beta)} \tag{1}$$

where β is the unique solution to the weight constraint

$$\sum_{a\in\mathcal{A}} P_\beta(a)\,\mathrm{wt}_\mathcal{A}(a) = \rho \tag{2}$$

and $\mathcal{Z}(\beta)$ is chosen s.t. $\sum_{a\in\mathcal{A}} P_\beta(a) = 1$. Note that the normalized radius ρ and β are in one-to-one correspondence due to the weight constraint (2) and Lemma 1 (cf. (3)). For a $\beta \in \mathbb{R}$, the value ρ determined by this correspondence is denoted $\rho(\beta)$. Let us denote by $H_\rho := H(P_\beta)$ the entropy of the distribution in (1). Then, the following bound was proven in [11].

Theorem 1 ([11]). *For any $0 < \rho \leq \overline{w}$ and $\ell \in \mathbb{N}$ we have*

$$\frac{1}{\ell} \log_q \left|\mathcal{B}_{\rho\ell}^\ell\right| \leq H_\rho.$$

The following is an immediate consequence of Theorem 1 above.

Corollary 1. *For any $0 < \rho < \overline{w}$ and $\ell \in \mathbb{N}$ we have*

$$\frac{1}{\ell} \log_q \left|\mathcal{S}_{\rho\ell}^\ell\right| \leq H_\rho.$$

3.1 Upper Bounds

We show that Corollary 1 also holds for normalized weights s.t. $0 < \rho < \mu$. Recall the OGFs for $\mathcal{A}$ and $\mathcal{A}^\ell$

$$F_\mathcal{A}(z) = \sum_{a\in\mathcal{A}} z^{\mathrm{wt}_\mathcal{A}(a)} \quad \text{and} \quad F_{\mathcal{A}^\ell}(z) = \sum_{v\in\mathcal{A}^\ell} z^{\mathrm{wt}_{\Sigma\mathcal{A}}(v)} = F_\mathcal{A}(z)^\ell.$$

We now can express $\mathcal{Z}(\beta)$, $\rho(\beta)$ and H_ρ in terms of these OGFs, i.e.

$$\mathcal{Z}(\beta) = F_\mathcal{A}\left(q^{-\beta}\right), \quad \rho(\beta) = q^{-\beta}\frac{F'_\mathcal{A}\left(q^{-\beta}\right)}{F_\mathcal{A}\left(q^{-\beta}\right)}, \quad H_\rho = \log_q\left(\frac{F_\mathcal{A}(q^{-\beta})}{(q^{-\beta})^\rho}\right). \tag{3}$$

Due to space constraints, we skip the proof for these equalities. We now make use of a technique explained in [2, Section VIII.2] where Flajolet and Sedgewick

present the saddle-point bound, i.e., an upper bound on the coefficients of a OGF. For any real valued $y > 0$ we have

$$|\mathcal{S}_t^\ell| y^t = \left([z^t] F_{\mathcal{A}^\ell}(z)\right) y^t \leq F_{\mathcal{A}^\ell}(y) = F_{\mathcal{A}}(y)^\ell.$$

We can further rewrite this expression and take the infimum on the right-hand side and obtain

$$\tfrac{1}{\ell} \log_q |\mathcal{S}_t^\ell| \leq \inf_{y>0} \log_q \left(\tfrac{F_{\mathcal{A}}(y)}{y^\rho} \right). \tag{4}$$

We can, moreover, show that a global minimum of $F_{\mathcal{A}}(y)/y^\rho$ exists and therefore the infimum is a minimum: by setting the derivative of $F_{\mathcal{A}}(y)/y^\rho$ to zero and using (3) for ρ, we obtain a local minimum for $y = q^{-\beta}$. Then using Lemma 1, we can show that the derivative of $F_{\mathcal{A}}(y)/y^\rho$ is negative for $0 < y < q^{-\beta}$ and positive for $y > q^{-\beta}$. Therefore, the local minimum is also the global minimum, where the function $\log_q \left(\frac{F_{\mathcal{A}}(y)}{y^\rho} \right)$ takes the value H_ρ (cf. (3)).

To summarize, the saddle-point bound (4) coincides with the entropy bound (see also [5, Theorem 4.1], [1, Theorem IV.9]), but extends the range of ρ to $(0, \mu)$, as stated in the following theorem.

Theorem 2. *For any $0 < \rho < \mu$ and $\ell \in \mathbb{N}$ we have*

$$\frac{1}{\ell} \log_q |\mathcal{S}_{\rho\ell}^\ell| \leq H_\rho.$$

3.2 Lower Bounds

We now derive a lower bound based on the probability distribution in (1). Let $X_\beta, X_{\beta,1}, X_{\beta,2}, X_{\beta,3}, \ldots$ be i.i.d. random variables taking values in $\mathcal{A}$ with probability distribution P_β. Define the function $\varphi_\beta(a) := -\log_q (P_\beta(a))$ for $a \in \mathcal{A}$. As a consequence of Chebyshev's inequality [20], we have for any $\gamma > 0$

$$\text{Prob}\left(\left| \tfrac{1}{\ell} \textstyle\sum_{i=1}^{\ell} \varphi_\beta(X_{\beta,i}) \;-\; H_\rho \right| \geq \gamma \right) \leq \tfrac{\text{Var}(\varphi_\beta(X_\beta))}{\ell\gamma^2} = \tfrac{\beta^2 \,\text{Var}(\text{wt}_{\mathcal{A}}(X_\beta))}{\ell\gamma^2}.$$

By setting $\gamma = |\beta|\delta/\ell$, where δ is chosen for some variable $0 < \varepsilon < 1$ as

$$\delta = \ell^{1/2} \frac{\text{Var}(\text{wt}_{\mathcal{A}}(X_\beta))^{1/2}}{(1-\varepsilon)^{1/2}}, \tag{5}$$

we can derive a lower bound with a similar technique used in [11].

Theorem 3. *Given $t = \ell\rho$ and $0 < \varepsilon < 1$, let β be defined by the weight constraint* (2) *and δ as in* (5)*. Then*

$$\sum_{-\delta<j<\delta,\, j\in\mathbb{Z}} |\mathcal{S}_{t+j}^\ell| \geq \varepsilon \; q^{\ell H(P_\beta) - |\beta|\delta}.$$

Theorem 4 gives an alternative bound using the inequality

$$\max_{-\delta<j<\delta,\, j\in\mathbb{Z}} |\mathcal{S}_{t+j}^\ell| \geq \frac{1}{2\lceil\delta\rceil - 1} \sum_{-\delta<j<\delta,\, j\in\mathbb{Z}} |\mathcal{S}_{t+j}^\ell|.$$

Theorem 4. *Given $t = \ell\rho$ and $0 < \varepsilon < 1$, let β be defined by the weight constraint* (2) *and δ as in* (5)*. Then*

$$\max_{-\delta<j<\delta,\, j\in\mathbb{Z}} \frac{1}{\ell}\log_q |\mathcal{S}^{\ell}_{t+j}| \geq H(P_\beta) - \frac{|\beta|\delta}{\ell} - \frac{1}{\ell}\log_q\left(\frac{2\lceil\delta\rceil - 1}{\varepsilon}\right).$$

Empirically, good bounds seem to be obtained for ε close to 0. Moreover, for constant ε and ρ, the bound coincides asymptotically with Theorem 2 as $\ell \to \infty$ and is therefore asymptotically tight.

4 Bounds on Spheres in the Sum-rank Metric

In this section we derive improved closed-form upper and lower bounds on the size of a sphere in the sum-rank metric. Hence, we fix m,η and q and we use $\mathrm{NM}_q(m,\eta,t)$ to denote the number of matrices of rank t over $\mathbb{F}_q^{m\times\eta}$. For $a, b \in \mathbb{N}$ we define the ***q*-binomial coefficient** as $\begin{bmatrix} a \\ b \end{bmatrix}_q = \prod_{i=1}^{b} \frac{1-q^{a-b+i}}{1-q^i}$. Then, $\mathrm{NM}_q(m,\eta,t) = \begin{bmatrix} m \\ t \end{bmatrix}_q \prod_{i=0}^{t-1}(q^\eta - q^i)$ (see [14]). The ***q*-Pochhammer symbol** is defined as

$$(a;x)_\infty := \textstyle\prod_{i=0}^{\infty}(1-ax^i), \quad \gamma_q := \left(\frac{1}{q};\frac{1}{q}\right)_\infty^{-1}.$$

Let $q \geq 2$, $\mu = \min\{m,\eta\}$, $\mathfrak{M} = \max\{m,\eta\}$ and $0 \leq i \leq \mu$. Then the q-binomial coefficients and q-Pochhammer symbols satisfy the following inequalities, that follow from elementary arguments (see [9, Lemma 2.2])

$$1 + \tfrac{1}{q} \geq \left(\tfrac{1}{q^2};\tfrac{1}{q^2}\right)_\infty^{-1} \quad \text{and} \quad \begin{bmatrix} \mu \\ i \end{bmatrix}_q \geq \begin{cases} (1+\frac{1}{q})q^{i(\mu-i)} & \text{if } 0 < i < \mu \\ 1 & \text{if } i = 0 \text{ or } i = \mu \end{cases}$$

and as a direct corollary of these two inequalities we obtain

$$\begin{bmatrix} \mu \\ i \end{bmatrix}_{1/q^2} q^{i(\mu-i)} \leq \begin{bmatrix} \mu \\ i \end{bmatrix}_q. \tag{6}$$

Now the inequality $\left(\prod_{j=0}^{a-1}(q^c - q^j)\right)^b > \left(\prod_{j=0}^{b-1}(q^c - q^j)\right)^a$ for $a, b, c \in \mathbb{N}$ with $0 \leq a < b < c$ yields

$$\textstyle\prod_{j=0}^{i-1}(q^{\mathfrak{M}} - q^j) > \left(\prod_{j=0}^{\mu-1}(q^{\mathfrak{M}} - q^j)\right)^{i/\mu} = q^{i\,\mathfrak{M}}\left({\gamma_{q,m,\eta}}^{-1}\right)^{i/\mu} \tag{7}$$

where we introduce the notation ${\gamma_{q,m,\eta}}^{-1} := \prod_{j=\mathfrak{M}-\mu+1}^{\mathfrak{M}}(1-(1/q)^j)$. Combining (6) and (7) lead to a new lower bound on the number of matrices of rank t.

Proposition 2. *For $m, \eta, i \in \mathbb{N}$ with $i \leq \mu$, we have the lower bound*

$$\left({\gamma_{q,m,\eta}}^{-1/\mu}\right)^i \begin{bmatrix} \mu \\ i \end{bmatrix}_{1/q^2} q^{i(m+\eta-i)} \leq \mathrm{NM}_q(m,\eta,i).$$

Next, we can obtain an upper bound for the number of matrices by introducing the function $\kappa_{q,m,\eta}(t) := \left(\frac{(1-q^{-m})(1-q^{-\eta})}{(1-q^{-1})}\right)^t$ and writing

$$\mathrm{NM}_q(m,\eta,t) = \left(\textstyle\prod_{i=1}^{t} \frac{(1-q^{-m+i-1})(1-q^{-\eta+i-1})}{(1-q^{-i})}\right) q^{t(m+\eta-t)}.$$

Proposition 3. *For $m, \eta, t \in \mathbb{N}$ we have the following upper bound*

$$\mathrm{NM}_q(m, \eta, t) \;\leq\; \kappa_{q,m,\eta}(t) q^{t(m+\eta-t)}.$$

In [17] an upper bound is derived using $\mathrm{NM}_q(m, \eta, t) \leq \gamma_q q^{t(m+\eta-t)}$. By doing similar steps with $\kappa_{q,m,\eta}(t)$ instead of γ_q we obtain Theorem 5.

Theorem 5. *Given positive integers m, η, ℓ, t and a prime power q, it holds*

$$\left|\mathcal{S}_t^{m,\eta,\ell,q}\right| \leq \kappa_{q,m,\eta}(t) \binom{\ell+1-1}{\ell-1} q^{t(m+\eta-\frac{t}{\ell})}.$$

Finally we state a strong form of log-concavity for $(\mathrm{NM}_q(m, \eta, i))_{i=0}^{\mu}$ that we apply later to Theorem 8.

Theorem 6. *For $0 < i < \mu$ we have*

$$\frac{\mathrm{NM}_q(m, \eta, i)^2}{\mathrm{NM}_q(m, \eta, i-1)\,\mathrm{NM}_q(m, \eta, i+1)} = \frac{(q^m - q^{i-1})}{(q^m - q^i)} \frac{(q^\eta - q^{i-1})}{(q^\eta - q^i)} \frac{q^i(q^{i+1}-1)}{q^{i-1}(q^i-1)} \geq q^2.$$

Moreover, since convolution preserves log-concavity, it holds that for all ℓ that the sequence $\left(\left|\mathcal{S}_i^{m,\eta,\ell,q}\right|\right)_{i=0}^{\mu\ell}$ is log-concave.

4.1 Integral Upper Bound

Let $f(x)$ and $g(x)$ be two real-valued functions defined on the natural numbers (or on a larger domain). We define the **discrete convolution** by $[f * g](t) := \sum_{i=0}^{t} f(i)g(t-i)$, for $t \in \mathbb{N}$. The ℓ-fold discrete convolution $[f * f * \cdots * f]$ (well-defined by associativity of $*$) is denoted as $f^{*\ell}$. Let $C(t)$ be a real-valued function depending on parameters m, η, q and satisfying

$$\left|\mathcal{S}_t^{m,\eta,1,q}\right| \leq C(t) q^{t(m+\eta-t)} \quad \text{and} \quad C(t_1)C(t_2) = C(t_3)C(t_4)$$

whenever $t_1 + t_2 = t_3 + t_4$. By Proposition 3, examples of such functions are γ_q and $\kappa_{q,m,\eta}(t)$. The reason for looking at these functions is because they work well with discrete convolutions, i.e., $[C(x)f(x) * C(x)g(x)]\,(t) = C(0)C(t)[f * g](t)$. Therefore, we can upper bound the sphere sizes as follows

$$\left|\mathcal{S}_t^{m,\eta,\ell,q}\right| \leq \left(C(x) q^{x(m+\eta-x)}\right)^{*\ell}(t) = C(0)^{\ell-1} C(t) \left(q^{x(m+\eta-x)}\right)^{*\ell}(t).$$

Proposition 4 provides a formula to compute convolutions.

Proposition 4. *Consider $f_\ell(x) := q^{x(m+\eta-x/\ell)}$ for $x \in \mathbb{R}$ and $\ell \in \mathbb{N}$. Functions of this form satisfy the following relation on their discrete convolutions*

$$[f_{\ell_1} * f_{\ell_2}](t) \leq \left(1 + \sqrt{\frac{\ell_1 \ell_2 \pi}{(\ell_1 + \ell_2)\ln q}}\right) f_{\ell_1+\ell_2}(t).$$

The bound is obtained by bounding summations by integrals and by noticing $[f_{\ell_1} * f_{\ell_2}](t) = f_{\ell_1+\ell_2}(t) \sum_{i=0}^{t} q^{-\left(\frac{1}{\ell_1}+\frac{1}{\ell_2}\right)\left(i-\frac{\ell_1}{\ell_1+\ell_2}t\right)^2}$. Setting $\ell_1 = 1$ and applying Proposition 4 inductively for $\ell_2 = 1, \ldots, \ell - 1$ we obtain upper bounds on the sphere sizes.

Theorem 7. *Let m, η, ℓ, q, t be positive integers. Choosing $C(t)$ equal to γ_q or $\kappa_{q,m,\eta}(t)$, we observe the following bounds, respectively*

$$\left|\mathcal{S}_t^{m,\eta,\ell,q}\right| \leq \gamma_q^{\ell} \prod_{k=1}^{\ell-1} \left(1 + \sqrt{\tfrac{k\pi}{(k+1)\ln q}}\right) q^{t(m+\eta-t/\ell)}$$
$$\left|\mathcal{S}_t^{m,\eta,\ell,q}\right| \leq \kappa_{q,m,\eta}(t) \prod_{k=1}^{\ell-1} \left(1 + \sqrt{\tfrac{k\pi}{(k+1)\ln q}}\right) q^{t(m+\eta-t/\ell)}$$

where the further simplifications $\sqrt{\frac{k\pi}{(k+1)\ln q}} \leq \sqrt{\frac{(\ell-1)\pi}{\ell \ln q}} < \sqrt{\frac{\pi}{\ln q}}$ can be made.

4.2 Lower Bound via Ordinary Generating Functions

An alternative approach is not to bound the number of matrices first, but to bound the generating function $\mathcal{S}^{m,\eta,1,q}(z)$ coefficient-wise with another polynomial $\mathcal{F}(z)$ whose ℓ-th power can be computed more easily. The polynomial $\mathcal{F}$ that we use to obtain a lower bound can be factored nicely into linear parts by the q-binomial theorem.

Proposition 5. *Let $m, \eta \in \mathbb{N}$. Then,*

$$\mathcal{F}(z) := \sum_{i=0}^{\mu} q^{i(m+\eta-i)} \begin{bmatrix} \mu \\ i \end{bmatrix}_{1/q^2} z^i = \prod_{i=1}^{\mu} (1 + q^{m+\eta-2i+1} z).$$

This polynomial satisfies the following chain of coefficient-wise inequalities

$$\textstyle\sum_{i=0}^{\mu} \gamma_q^{-1} q^{i(m+\eta-i)} z^i \;\preccurlyeq_c\; \gamma_q^{-1} \mathcal{F}(z) \;\preccurlyeq_c\; \mathcal{F}(\gamma_{q,m,\eta}{}^{-1/\mu} z) \;\preccurlyeq_c\; \mathcal{S}^{m,\eta,1,q}(z).$$

The first inequality follows from $\begin{bmatrix} \mu \\ i \end{bmatrix}_{1/q^2} \geq 1$, the second from $\gamma_q^{-1} \leq \gamma_{q,m,\eta}{}^{-1} \leq \gamma_{q,m,\eta}{}^{-i/\mu} \leq 1$ for $0 \leq i \leq \mu$ and the third from Proposition 2. Since this coefficient-wise inequality is preserved under convolution, we obtain

$$\left(\textstyle\sum_{i=0}^{\mu} \gamma_q^{-1} q^{i(m+\eta-i)} z^i\right)^{\ell} \;\preccurlyeq_c\; \mathcal{F}(\gamma_{q,m,\eta}{}^{-1/\mu} z)^{\ell} \;\preccurlyeq_c\; \mathcal{S}^{m,\eta,\ell,q}(z). \tag{8}$$

If we look now at $\mathcal{F}(z)^{\ell} = \prod_{i=1}^{\mu} \left(\sum_{j=0}^{\ell} \binom{\ell}{j} q^{j(m+\eta-2i+1)} z^j\right)$, we can lower bound $[z^t]\mathcal{F}(z)^{\ell}$ as follows: let $t = t_*\ell + r$ with $t_* \in \mathbb{N}$ and $0 \leq r < \ell$. Then using, depending on i, the inequality

$$\left(\textstyle\sum_{j=0}^{\ell} \binom{\ell}{j} q^{j(m+\eta-2i+1)} z^j\right) \succcurlyeq_c \begin{cases} q^{\ell(m+\eta-2i+1)} z^{\ell} & \text{for } \; 1 \leq i \leq t_* \\ \binom{\ell}{r} q^{r(m+\eta-2i+1)} z^r & \text{for } \; i = t_* + 1 \\ 1 & \text{for } \; t_* + 2 \leq i \leq \mu \end{cases} \tag{9}$$

we obtain

$$\sum_{t=0}^{\mu\ell} \binom{\ell}{r} q^{t(m+\eta-\frac{t}{\ell})+\frac{r^2}{\ell}-r} z^t \preccurlyeq_c \mathcal{F}(z)^\ell.$$

Finally, substituting ${\gamma_{q,m,\eta}}^{-1/\mu} z$ for z in this inequality and applying equation (8) we get the following result.

Theorem 8. *Let $t = t_*\ell + r$ with $t_* \in \mathbb{N}$ and $0 \leq r < \ell$. Then*

$$\left({\gamma_{q,m,\eta}}^{-1}\right)^{t/\mu} \binom{\ell}{r} q^{t(m+\eta-\frac{t}{\ell})+\frac{r^2}{\ell}-r} \leq \left|\mathcal{S}_t^{m,\eta,\ell,q}\right|.$$

Notice that remarkably, aside for the coefficient in front, we have obtained the same lower bound as [16, Lemma 2] via a completely different method. However, by choosing different inequalities in (9) there is still room for future optimization. Since $\left(\left|\mathcal{S}_i^{m,\eta,\ell,q}\right|\right)_{i=0}^{\mu\ell}$ is log-concave, we can take the smallest concave sequence that is coefficient-wise greater or equal to the sequence $\left(\log_q\left(\left({\gamma_{q,m,\eta}}^{-1}\right)^{i/\mu}\binom{\ell}{r} q^{i(m+\eta-\frac{i}{\ell})+\frac{r^2}{\ell}-r}\right)\right)_{i=0}^{\mu\ell}$ (i.e. its convex hull) for a slightly improved lower bound on $\log_q\left|\mathcal{S}_t^{m,\eta,\ell,q}\right|$.

5 Comparison of Bounds

In this section, we compare the new bounds presented in this paper with the existing bounds related to the sphere size in the sum-rank metric. In Figure 1 the relationship between the growth rate $\frac{1}{\ell}\log_q\left|\mathcal{S}_t^{m,\eta,\ell,q}\right|$ of the sphere size and the normalized radius ρ is shown. We observe that the upper bound using Theorem 2 and the lower bound using Theorem 4 are the tightest bounds and very close to the exact values. The computation of these bounds necessitates the evaluation of the entropy H_ρ. Computing H_ρ is straightforward for a specified β, whereas determining β for a given ρ cannot be achieved in a closed-form manner, as outlined in (2). For scenarios where prioritizing closed-form expressions dependent on ρ is essential, the derived alternative bounds may better suit the intended use-cases. In Figure 1, the upper bounds from Theorem 7 using $\kappa_{q,m,\eta}$, Theorem 7 using γ_q and Theorem 5 are consolidated into a single piece-wise function by selecting the minimum value among these bounds. The transition points are indicated by circles. We observe that for the new closed-form upper and lower bounds we improve significantly in comparison to the already existing closed-form bounds given in [17, Theorem 5] and [16, Lemma 2]. Furthermore, the new bounds are potentially useful tools for obtaining improved closed-form Gilbert-Varshamov or sphere-packing bounds, as introduced in [1] and [16].

In Figure 2 we show the tightness of the improved bounds for different numbers of blocks. We choose the same values for the parameters q, m, t and n as for the bounds given in [17]. Notably, the bounds proposed in [17] exhibit considerable looseness in scenarios where ℓ becomes substantially large (i.e., when the sum-rank metric converges to the Hamming metric). While superior bounds are already established for the Hamming metric (i.e., $\ell = n$), our analysis illustrates substantial enhancements for $\ell < 60$ compared to existing bounds.

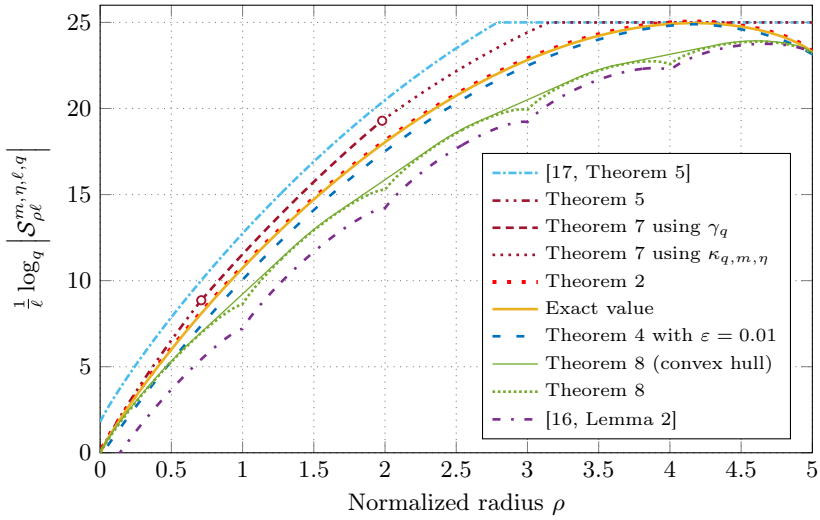


Fig. 1. Comparison of upper and lower bounds for the sphere $\mathcal{S}_{\rho\ell}^{m,\eta,\ell,q}$ as function of ρ with parameters $q = 2$, $m = 5$, $\eta = 5$, $\ell = 100$.

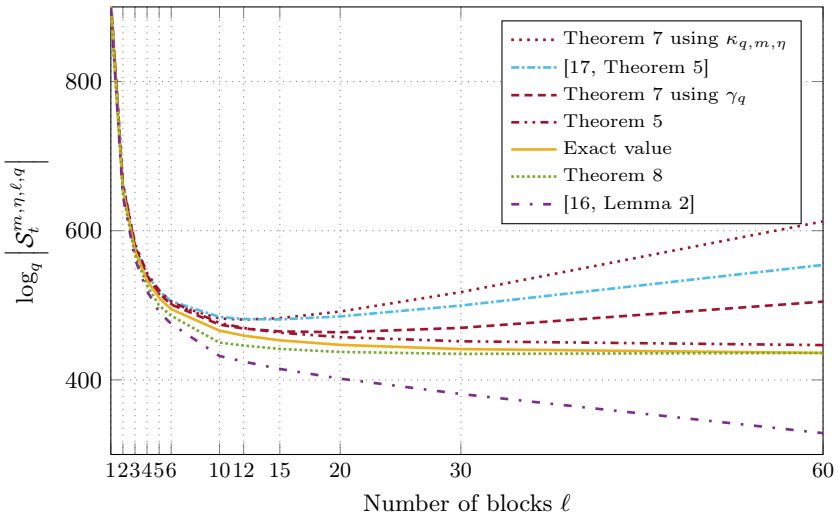


Fig. 2. Comparison of upper and lower bounds for the sphere $\mathcal{S}_t^{m,\eta,\ell,q}$ as function of ℓ with parameters $q = 2$, $m = 40$, $t = 10$ and keeping $n = \eta\ell = 60$ constant.

Acknowledgments. H. Sauerbier Couvée is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 801434) and the Bavarian State Ministry of Science and Arts via the project EQAP. T. Jerkovits and J. Bariffi acknowledge the financial support by the Federal Ministry of Education and Research of Germany in the program of "Souverän. Digital. Vernetzt." Joint project 6G-RIC, project identification number: 16KISK022. The authors thank H. Bartz, A. Baumeister, G. Liva, and A. Wachter-Zeh for their insights and discussions.

References

1. Byrne, E., Gluesing-Luerssen, H., Ravagnani, A.: Fundamental properties of sum-rank-metric codes. IEEE Transactions on Information Theory **67**(10), 6456–6475 (2021). https://doi.org/10.1109/TIT.2021.3074190
2. Flajolet, P., Sedgewick, R.: Analytic combinatorics. Cambridge University Press (2009)
3. Gardy, D., Solé, P.: Saddle point techniques in asymptotic coding theory. In: Workshop on Algebraic Coding. pp. 75–81. Springer (1991)
4. Gilbert, E.N.: A comparison of signalling alphabets. The Bell System Technical Journal **31**(3), 504–522 (1952)
5. Greferath, M., O'Sullivan, M.E.: On bounds for codes over frobenius rings under homogeneous weights. Discrete Mathematics **289**(1), 11–24 (2004). https://doi.org/https://doi.org/10.1016/j.disc.2004.10.002
6. Gruica, A., Horlemann, A.L., Ravagnani, A., Willenborg, N.: Densities of codes of various linearity degrees in translation-invariant metric spaces. Designs, Codes and Cryptography (May 2023), https://doi.org/10.1007/s10623-023-01236-2
7. Hamming, R.W.: Error detecting and error correcting codes. The Bell system technical journal **29**(2), 147–160 (1950)
8. Horlemann-Trautmann, A.L., Weger, V.: Information set decoding in the Lee metric with applications to cryptography. Advances in Mathematics of Communications **15**(4) (2021). https://doi.org/10.3934/amc.2020089
9. Ihringer, F., Justus Liebig University Giessen: Finite geometry intersecting algebraic combinatorics : an investigation of intersection problems related to Erdös-Ko-Rado theorems on Galois geometries with help from algebraic combinatorics. Ph.D. thesis (2015). https://doi.org/10.22029/JLUPUB-9701
10. Lee, C.: Some properties of nonbinary error-correcting codes. IRE Transactions on Information Theory **4**(2), 77–82 (1958)
11. Loeliger, H.A.: An upper bound on the volume of discrete spheres. IEEE Transaction of Information Theory **40**(6), 2071–2073 (1994)
12. Lu, H.F., Kumar, P.V.: A unified construction of space-time codes with optimal rate-diversity tradeoff. IEEE Transactions on Information Theory **51**(5) (2005). https://doi.org/10.1109/TIT.2005.846403
13. Martínez-Peñas, U., Kschischang, F.R.: Universal and dynamic locally repairable codes with maximal recoverability via sum-rank codes. IEEE Transactions on Information Theory **65**(12), 7790–7805 (2019). https://doi.org/10.1109/TIT.2019.2924888
14. Migler, T., Morrison, K.E., Ogle, M.: Weight and rank of matrices over finite fields (2004)
15. da Nóbrega, R.W., Filho, B.F.U.: Multishot codes for network coding using rank-metric codes. 2010 Third IEEE International Workshop on Wireless Network Coding pp. 1–6 (2010), https://api.semanticscholar.org/CorpusID:10178357
16. Ott, C., Puchinger, S., Bossert, M.: Bounds and genericity of sum-rank-metric codes. In: 2021 17th International Symposium Problems of Redundancy in Information and Control Systems, REDUNDANCY 2021 (2021). https://doi.org/10.1109/REDUNDANCY52534.2021.9606442
17. Puchinger, S., Renner, J., Rosenkilde, J.: Generic decoding in the sum-rank metric. IEEE Transactions on Information Theory **68**(8), 5075–5097 (2022)
18. Sauerbier Couvée, H., Jerkovits, T., Bariffi, J.: Bounds on sphere sizes in the sum-rank metric and coordinate-additive metrics (2024), https://arxiv.org/abs/2404.10666

19. Shehadeh, M., Kschischang, F.R.: Space-time codes from sum-rank codes (2021)
20. Tchebichef, P.: Sur les valeurs limites des intégrales. Journal de mathématiques pures et appliquées **19**, 157–160 (1874)

On the Recursive Behaviour of the Number of Irreducible Polynomials with Certain Properties over Finite Fields

Max Schulz

University of Rostock, Germany
max.schulz@uni-rostock.de

Abstract. Let $\mathbb{F}_q$ be the field with q elements and of characteristic p. For $a \in \mathbb{F}_p$ consider the set

$$S_a(n) = \{f \in \mathbb{F}_q[x] \mid \deg(f) = n,\ f \text{ irreducible, monic and } \mathrm{Tr}(f) = a\}.$$

In a recent paper, Robert Granger proved for $q = 2$ and $n \geq 1$

$$|S_1(n)| - |S_0(n)| = \begin{cases} 0, & \text{if } 2 \nmid n \\ |S_1(n/2)|, & \text{if } 2 \mid n. \end{cases}$$

We will prove a generalization of this result for all finite fields. This is possible due to an observation about the size of certain subsets of monic irreducible polynomials arising in the context of a group action of subgroups of $\mathrm{PGL}_2(\mathbb{F}_q)$ on monic polynomials. Additionally, it enables us to apply these methods to prove two further results that are very similar in nature.

Keywords: Irreducible Polynomials, Group Action, Projective General Linear Group

Introduction

Let $\mathbb{F}_q$ be the finite field with q elements, p the prime dividing q, $\mathcal{I}_q$ the set of monic irreducible polynomials in $\mathbb{F}_q[x]$ and $\mathcal{I}_q^n$ the set of monic irreducible polynomials of degree n. Moreover, $\mathrm{Tr}_{q/p}$ is the absolute trace. Since $\mathrm{Tr}_{q^n/p}(\alpha) = \mathrm{Tr}_{q^n/p}(\alpha^q)$ for a root α of $f \in \mathcal{I}_q^n$ we can define $\mathrm{Tr}(f) := \mathrm{Tr}_{q^n/p}(\alpha)$. For an $a \in \mathbb{F}_p$ consider the set

$$S_a(n) := \{f \in \mathcal{I}_q^n \mid \mathrm{Tr}(f) = a\}.$$

Let $f \in \mathcal{I}_q^n$ be of the form $f = x^n + \sum_{i=0}^{n-1} a_i x^i$ then the trace is given by

$$\mathrm{Tr}(f) = -\mathrm{Tr}_{q/p}(a_{n-1}).$$

In [6] it is proved that for $q = 2$ and $n \geq 1$

$$|S_1(n)| - |S_0(n)| = \begin{cases} 0, & \text{if } n \equiv 1 \pmod 2 \\ |S_1(n/2)|, & \text{otherwise.} \end{cases}$$

We are going to prove the following extension of this result:

Theorem 1. *For all $n \geq 1$ and all finite fields $\mathbb{F}_q$ we have*

$$\sum_{a \in \mathbb{F}_p^*} |S_a(n)| - (p-1)|S_0(n)| = \begin{cases} 0, & \text{if } p \nmid n \\ \sum\limits_{a \in \mathbb{F}_p^*} |S_a(n/p)|, & \text{otherwise.} \end{cases}$$

Remark 2. Note that the balanced case, that is,

$$\sum_{a \in \mathbb{F}_p^*} |S_a(n)| - (p-1)|S_0(n)| = 0,$$

where $p \nmid n$, is not hard to see. Let $f \in S_0(n)$, so $\mathrm{Tr}(f) = 0$ which means that $\mathrm{Tr}_{q^n/p}(\alpha) = 0$ for α a root of f. Let $a \in \mathbb{F}_p^*$ and consider an element $b_a \in \mathbb{F}_q^*$ such that $\mathrm{Tr}_{q^n/p}(b_a) = a$. Such an element exists since $\mathrm{Tr}_{q^n/p} = \mathrm{Tr}_{q/p} \circ \mathrm{Tr}_{q^n/q}$ and for all $b \in \mathbb{F}_q^*$

$$\mathrm{Tr}_{q^n/q}(b) = n \cdot b \neq 0$$

if $p \nmid n$, hence $\mathrm{Tr}_{q^n/p}$ as a map from $\mathbb{F}_q$ to $\mathbb{F}_p$ is surjective as $\mathrm{Tr}_{q/p} : \mathbb{F}_q \to \mathbb{F}_p$ and $\mathrm{Tr}_{q^n/q}|_{\mathbb{F}_q} : \mathbb{F}_q \to \mathbb{F}_q$ are surjective. The polynomial $f(x - b_a)$ has trace a, so the map $f(x) \mapsto f(x - b_a)$ is a bijection between $S_0(n)$ and $S_a(n)$, thus $|S_0(n)| = |S_a(n)|$ for all $a \in \mathbb{F}_p$ and the balanced case follows. A similar idea does not work for the case that $p \mid n$ since then $\mathrm{Tr}_{q^n/q}|_{\mathbb{F}_q}$ is not surjective anymore.

Another example that exhibits a similar pattern is the following: Let q be odd and $u, v \in \mathbb{F}_q$ with $u \neq v$. Define the following two sets for $n \geq 2$

$$C_{u,v}(n) := \left\{ f \in \mathcal{I}_q^n \;\middle|\; \left(\frac{f(u) \cdot f(v)}{q}\right) = -1 \right\}$$
$$D_{u,v}(n) := \mathcal{I}_q^n \setminus C_{u,v}(n) = \left\{ f \in \mathcal{I}_q^n \;\middle|\; \left(\frac{f(u) \cdot f(v)}{q}\right) = 1 \right\}.$$

Here $\left(\frac{\cdot}{q}\right) : \mathbb{F}_q^* \to \{1, -1\} \leq \mathbb{F}_q^*$ denotes the Legendre-Symbol

$$\left(\frac{a}{q}\right) = a^{(q-1)/2} = \begin{cases} 1, & a \text{ is a square in } \mathbb{F}_q^* \\ -1, & \text{otherwise.} \end{cases}$$

We prove the following theorem:

Theorem 3. *Let $q \equiv 1 \pmod 2$. For all $u, v \in \mathbb{F}_q$ with $u \neq v$ and $n \geq 2$ we have*

$$|C_{u,v}(n)| - |D_{u,v}(n)| = \begin{cases} 0, & 2 \nmid n \\ |C_{u,v}(n/2)|, & 2 \mid n. \end{cases}$$

In [6] a group action of subgroups of $\mathrm{PGL}_2(\mathbb{F}_q)$ on irreducible polynomials over $\mathbb{F}_q$ played a crucial role in some of the proofs, so we thought that ideas out of our recent paper [13] could be utilized to prove similar results. Our proof of Theorem 1 relies on a general underlying principle which can be used to obtain Theorem 3 as well. We give a quick overview:

For an element $A \in \mathrm{GL}_2(\mathbb{F}_q)$ we write $[A] \in \mathrm{PGL}_2(\mathbb{F}_q)$ as its coset in $\mathrm{PGL}_2(\mathbb{F}_q)$ and if A is of the form

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then set

$$[A] \circ x = \frac{ax+b}{cx+d}$$

as the corresponding linear rational function. For a subgroup $G \leq \mathrm{PGL}_2(\mathbb{F}_q)$ consider the set of G-invariant rational functions

$$\mathbb{F}_q(x)^G := \{Q(x) \in \mathbb{F}_q(x) \mid Q([A] \circ x) = Q(x) \text{ for all } [A] \in G\}.$$

This is a subfield of $\mathbb{F}_q(x)$ with $[\mathbb{F}_q(x) : \mathbb{F}_q(x)^G] = |G|$. Moreover, by Lüroth's Theorem, there is a rational function $Q(x) = g(x)/h(x) \in \mathbb{F}_q(x)$ of degree $\deg(Q) = \max\{\deg(g), \deg(h)\} = |G|$ such that $\mathbb{F}_q(x)^G = \mathbb{F}_q(Q(x))$. Note that we always assume that the numerator and denominator of a rational function have no common factors. Every such generator Q of $\mathbb{F}_q(x)^G$ can be normalized so that $Q = g/h$ with $\deg(g) = |G|$ and $0 \leq \deg(h) < \deg(g)$, we call these rational functions *quotient maps* for G and in what follows we write Q_G for an arbitrary quotient map for G. In [13], we studied the factorization of rational transformations with quotient maps. A rational transformation of a polynomial F with a rational function $Q = g/h$ is defined as

$$F^Q(x) := h(x)^{\deg(F)} \cdot F\left(\frac{g(x)}{h(x)}\right) \tag{1}$$

so it is the numerator polynomial of the rational function $F(Q(x))$. To avoid ambiguity we set the numerator polynomial g of Q to be monic. Define the following two sets

$$C(Q_G, n) := \left\{ f \in \mathcal{I}_q^n \mid f^{Q_G} \in \mathcal{I}_q^{|G|n} \right\}$$
$$D(Q_G, n) := \mathcal{I}_q^n \setminus C(Q_G, n).$$

So $C(Q_G, n)$ is the set of irreducible polynomials f of degree n that yield irreducible polynomials of degree $|G| \cdot n$ after transformation with quotient map Q_G. The following theorem will be the backbone of our proofs of Theorem 1 and 3:

Theorem 4. *Let $G \leq \mathrm{PGL}_2(\mathbb{F}_q)$ be a cyclic subgroup of prime order s and $Q_G \in \mathbb{F}_q(x)$ a quotient map for G. For all $n > d(G)$ (the number $d(G)$ will be defined before Remark 13) we have*

$$|C(Q_G, n)| - (s-1)|D(Q_G, n)| = \begin{cases} 0, & \text{if } s \nmid n \\ |C(Q_G, \frac{n}{s})|, & \text{if } s \mid n. \end{cases}$$

Both results are immediate consequences of this theorem by choosing the right cyclic subgroups G and quotient maps Q_G. The set $C(Q_G, n)$ can be occasionally described in terms of arithmetic properties that the coefficients of irreducible polynomials in $C(Q_G, n)$ need to satisfy if the quotient map Q_G was chosen carefully.

In this extended abstract we focus on how to apply Theorem 4 but will omit proving it. If the reader is interested in the proof we refer them to our arxiv-paper [12], which is the in-depth version of this abstract.

1 Invariant Polynomials and Rational Transformations

Every $[A] \in \mathrm{PGL}_2(\mathbb{F}_q)$ induces a bijective map on $\overline{\mathbb{F}}_q \cup \{\infty\}$ via

$$[A] \circ v = \frac{av + b}{cv + d},$$

i.e. just plugging v into the linear rational function belonging to $[A]$. This induces a left group action of $\mathrm{PGL}_2(\mathbb{F}_q)$ on $\overline{\mathbb{F}}_q \cup \{\infty\}$. An intimately related group action on polynomials is given by

Definition 5. *Define* $* : \mathrm{PGL}_2(\mathbb{F}_q) \times \mathbb{F}_q[x] \to \mathbb{F}_q[x]$ *with*

$$[A] * f(x) = \lambda_{A,f}(cx + d)^{\deg(f)} f\left(\frac{ax + b}{cx + d}\right).$$

The factor $\lambda_{A,f} \in \mathbb{F}_q^*$ *makes the output-polynomial monic.*

In other words, $[A] * f$ is the normalized $([A] \circ x)$-transformation of f. This transformation and its variations are well-studied objects over finite fields, see for example [4], [10], [11], [14] and it has some theoretic applications, see for example [6], [7] and [9].

Let $G \leq \mathrm{PGL}_2(\mathbb{F}_q)$ be a subgroup and $G \circ \infty := \{[A] \circ \infty \mid [A] \in G\}$ be the G-orbit of ∞. Define

$$\mathcal{NR}_q^G := \{f \in \mathbb{F}_q[x] \mid f \text{ monic and } f(\alpha) \neq 0 \text{ for all } \alpha \in G \circ \infty\}$$

where $f(\infty) = \infty$ if $\deg(f) \geq 1$ and $a(\infty) = a$ for all $a \in \mathbb{F}_q$.

Lemma 6 ([13, Lemma 7]). *Let* $G \leq \mathrm{PGL}_2(\mathbb{F}_q)$. *For all* $f, g \in \mathcal{NR}_q^G$ *and* $[A], [B] \in G$ *the following hold:*

1. $\deg([A] * f) = \deg(f)$
2. $[AB] * f = [B] * ([A] * f)$ *and* $[I_2] * f = f$, *so* $*$ *is a right group action of* G *on* $\mathcal{NR}_q^G$
3. $[A] * (fg) = ([A] * f)([A] * g)$

4. f irreducible if and only if $[A] * f$ irreducible.

The first and fourth item show that G also acts on $\mathcal{I}_q^n$ for $n \geq 2$ (since $G \circ \infty \subseteq \mathbb{F}_q \cup \{\infty\}$) and on

$$\mathcal{I}_q^G := \mathcal{NR}_q^G \cap \mathcal{I}_q.$$

We denote a G-orbit in $\mathcal{NR}_q^G$ as $G * r := \{[A] * r \mid [A] \in G\}$.

Definition 7. *A polynomial $f \in \mathcal{NR}_q^G$ is called G-orbit polynomial if there is an irreducible polynomial $r \in \mathcal{I}_q^G$ such that*

$$f = \prod_{t \in G*r} t =: \prod G * r.$$

A G-invariant polynomial is a polynomial $f \in \mathcal{NR}_q^G$ such that $[A] * f = f$ for all $[A] \in G$. Every G-invariant polynomial can be written as the product of G-orbit polynomials (which are G-invariant by Lemma 6 3.), so G-orbit polynomials can be seen as the atoms of G-invariant polynomials.

Next we want to recollect some facts about rational transformations. For $Q = g/h$ with $\gcd(g, h) = 1$ and $F \in \mathbb{F}_q[x]$ we write $F^Q \in \mathbb{F}_q[x]$ as the Q-transform of F with Q as in (1). If F^Q is irreducible then F has to be irreducible if $F(Q(\infty)) \neq 0$. The following lemma gives a necessary and sufficient condition for the irreducibility of F^Q:

Lemma 8 ([3, Lemma 1]). *Let $Q(x) = g(x)/h(x) \in \mathbb{F}_q(x)$ and $F \in \mathbb{F}_q[x]$ such that $F(Q(\infty)) \neq 0$. Then $F^Q \in \mathbb{F}_q[x]$ is irreducible if and only if $F \in \mathbb{F}_q[x]$ is irreducible and $g(x) - \alpha h(x)$ is irreducible over $\mathbb{F}_q(\alpha)[x]$, where α is a root of F.*

Now consider a quotient map $Q_G \in \mathbb{F}_q(x)$ of $G \leq \mathrm{PGL}_2(\mathbb{F}_q)$. We have

Lemma 9 ([13, Lemma 13 and Lemma 14]). *Let $F \in \mathbb{F}_q[x]$ be a monic polynomial, then $F^{Q_G} \in \mathcal{NR}_q^G$ and F^{Q_G} is G-invariant. Moreover, F^{Q_G} is of degree $\deg(F^{Q_G}) = |G| \cdot \deg(F)$.*

Theorem 10 ([13, Main Theorem, Theorem 22 and Corollary 23]). *Let $F \in \mathbb{F}_q[x]$ be monic and irreducible, $G \leq \mathrm{PGL}_2(\mathbb{F}_q)$ a subgroup and $Q_G = g/h \in \mathbb{F}_q(x)$ a quotient map for G. Then there is an irreducible monic polynomial $r \in \mathbb{F}_q[x]$ with $\deg(F) | \deg(r)$ and an integer $k > 0$ such that*

$$F^{Q_G}(x) = \left(\prod G * r\right)^k.$$

Additionally F^{Q_G} is an orbit polynomial, i.e. $k = 1$, if $|G \circ v| = |G|$ for a root $v \in \overline{\mathbb{F}}_q$ of F^{Q_G}. In the case that F^{Q_G} is an orbit polynomial the degree of every irreducible factor of F^{Q_G} can be calculated via

$$\deg(r) = \frac{|G|}{|G * r|} \cdot \deg(F).$$

The polynomials $F \in \mathcal{I}_q$ for which $F^{Q_G} = (\prod G * r)^k$ with $k > 1$ are of degree $\deg(F) \leq 2$. To show that we use the fact that F^{Q_G} is an orbit polynomial if every (or equivalently just one) root $v \in \overline{\mathbb{F}}_q$ of F^{Q_G} is contained in a regular G-orbit, i.e. $|G \circ v| = |G|$ (Theorem 10) and irreducible polynomials F not satisfying that condition are of degree less than or equal to 2, as the following lemma shows:

Lemma 11 ([1, Lemma 2.1]). *Let $G \leq \mathrm{PGL}_2(\mathbb{F}_q)$ and set*

$$P_G := \left\{ v \in \overline{\mathbb{F}}_q \cup \{\infty\} \mid |G \circ v| < |G| \right\}.$$

We have $P_G \subseteq \mathbb{F}_{q^2} \cup \{\infty\}$ and $|P_G| \leq 2(|G| - 1)$. Moreover, $[\mathbb{F}_q(v) : \mathbb{F}_q] \leq 2$ for all $v \in P_G \setminus \{\infty\}$.

Let $F \in \mathcal{I}_q$. Note that if F^{Q_G} has roots in non-regular G-orbits, then it only has irreducible factors of degree less than 3 by the lemma above. Moreover, we know that if r is an irreducible factor of F^{Q_G}, then $\deg(F) \mid \deg(r)$, so $\deg(F) \leq 2$, which is exactly what we wanted to show. Furthermore, there are only finitely many irreducible monic polynomials $F \in \mathcal{I}_q$ such that F^{Q_G} is not a G-orbit polynomial but a proper power thereof, as the number of non-regular G-orbits in $\overline{\mathbb{F}}_q \cup \{\infty\}$ is finite.

The next corollary is one of our main tools we make use of in this paper. Define $\mathcal{I}_q^G/G$ as the set of G-orbits in $\mathcal{I}_q^G$, that is,

$$\mathcal{I}_q^G/G := \{G * r \mid r \in \mathcal{I}_q^G\}.$$

Corollary 12 ([13, Corollary 25]). *The map $\delta_{Q_G} : \mathcal{I}_q \to \mathcal{I}_q^G/G$ with $F \mapsto G * r$ such that $F^{Q_G} = \prod (G * r)^k$ is a bijection.*

An irreducible monic G-invariant polynomial f is a G-orbit polynomial, thus f can be written as $f = F^{Q_G}$ for F an irreducible monic polynomial if a root of f is contained in a regular G-orbit by Theorem 10 and Corollary 12.

2 Application of Theorem 4

Set $d(Q_G) \in \mathbb{N}_0$ as the biggest number such that there exists an irreducible polynomial $F \in \mathcal{I}_q$ of degree $d(Q_G)$ with $F^{Q_G} = \prod (G * r)^k$ and $k > 1$. If no such polynomial exists then $d(Q_G) := 0$. Recall that $d(Q_G) \leq 2$.

Remark 13. If $Q_G, Q'_G \in \mathbb{F}_q(x)$ are quotient maps for G, then $d(Q_G) = d(Q'_G)$.

Proof. Let $F \in \mathcal{I}_q$ be of degree $d(Q_G)$ such that $F^{Q_G} = (\prod G * r)^k$ and $k > 1$. It can be shown that there are $a \in \mathbb{F}_q^*$ and $b \in \mathbb{F}_q$ such that $Q_G(x) = aQ'_G(x) + b$, for reference see [1, Proposition 3.4]. Since we want the numerator polynomial of quotient maps to be monic we write for $Q'_G = g'/h'$:

$$Q_G(x) = aQ'_G(x) + b = \frac{g'(x)}{a^{-1}h'(x)} + b.$$

Thus we have

$$F^{Q_G}(x) = (a^{-1}h'(x))^{\deg(F)} \cdot F(Q_G(x)) = h'(x)^{\deg(F)} \cdot \left((a^{-1})^{\deg(F)} F(aQ'_G(x) + b)\right)$$
$$= h'(x))^{\deg(F)} \cdot (\underbrace{[A] * F}_{=:H(x)})^{Q'_G},$$

where

$$A = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}.$$

Therefore $H(x) = [A] * F(x)$ is an irreducible polynomial of degree $\deg(H) = \deg(F) = d(Q_G)$. Additionally, $H^{Q'_G} = F^{Q_G} = \prod(G * r)^k$ with $k > 1$, so $d(Q'_G) \geq d(Q_G)$. Because of symmetry we get $d(Q'_G) = d(Q_G)$. □

This is why we can write $d(G)$ instead of $d(Q_G)$.

We will often determine $d(G)$ by using the second part of Theorem 10 in a contrapositive way, that means:

$$F^{Q_G} = \prod(G * r)^k \text{ with } k > 1 \Rightarrow |G \circ v| < |G| \text{ for a root } v \in \overline{\mathbb{F}}_q \text{ of } F^{Q_G}.$$

Since there are only finitely many non-regular G-orbits we only have to check finitely many irreducible polynomials F of degree 1 or 2. In this paper we only have to check at most 2 polynomials.

2.1 Proof of Theorem 1

For proving Theorem 1 we need to consider the cyclic subgroup

$$G := \left\langle \left[\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right] \right\rangle.$$

It has order $p = \text{char}(\mathbb{F}_q)$ and a quotient map is $Q_G(x) = x^p - x$. Note that the Q_G-transformation of polynomials $F \in \mathbb{F}_q[x]$ with Q_G is just the composition of F with Q_G, that is, $F^{Q_G}(x) = F(Q_G(x))$. The condition for $F(Q_G(x)) = F(x^p - x)$ to be irreducible is well-known and originally due to Varshamov, see for example [2, Lemma 1.1] and the references therein. For $F \in \mathcal{I}_q^n$ and $\alpha \in \overline{\mathbb{F}}_q$ a root of F we have

$$F(Q_G(x)) \in \mathbb{F}_q[x] \text{ is irreducible} \Leftrightarrow \text{Tr}_{q^n/p}(\alpha) \neq 0.$$

As mentioned in the introduction we can write the condition as follows:

$$F(Q_G(x)) \in \mathbb{F}_q[x] \text{ is irreducible } \Leftrightarrow \text{Tr}(F) \neq 0.$$

Hence

$$\bigcup_{a \in \mathbb{F}_p^*} S_a(n) = C(Q_G, n)$$

and $D(Q_G, n) = S_0(n)$. The number $d(G) = 0$ since the only non-regular G-orbit in $\overline{\mathbb{F}}_q \cup \{\infty\}$ is $\{\infty\}$. Applying Theorem 4 gives

$$\begin{aligned}\left|\bigcup_{a\in\mathbb{F}_p^*} S_a(n)\right| - (p-1)|S_0(n)| &= |C(Q_G,n)| - (p-1)|D(Q_G,n)| \\ &= \begin{cases} 0, & p \nmid n \\ |C(Q_G, n/p)|, & p \mid n \end{cases} \\ &= \begin{cases} 0, & p \nmid n \\ |\bigcup_{a\in\mathbb{F}_p^*} S_a(n/p)|, & p \mid n \end{cases}\end{aligned}$$

for all $n > d(G) = 0$. This proves Theorem 1.

2.2 Proof of Theorem 3

Assume $q \equiv 1 \pmod 2$ and let $u, v \in \mathbb{F}_q$ such that $u \neq v$. Consider the matrix

$$A_{u,v} := \begin{pmatrix} \frac{1}{2}(u+v) & -uv \\ 1 & -\frac{1}{2}(u+v) \end{pmatrix}.$$

Since $[A_{u,v}]^2 = [I_2]$ the cyclic group $G_{u,v} := \langle [A_{u,v}] \rangle \leq \mathrm{PGL}_2(\mathbb{F}_q)$ contains only 2 elements. As a quotient map for $G_{u,v}$ we choose

$$Q_{G_{u,v}}(x) = \frac{1}{2}(x + [A_{u,v}] \circ x) = \frac{x^2 - uv}{2x - (u+v)}.$$

Note that $Q_{G_{u,v}}(u) = u$ and $Q_{G_{u,v}}(v) = v$ since $[A_{u,v}] \circ u = u$ and $[A_{u,v}] \circ v = v$. Moreover $\{u\}$ and $\{v\}$ are the only non-regular $G_{u,v}$-orbits in $\overline{\mathbb{F}}_q \cup \{\infty\}$, hence $d(G_{u,v})$ is the highest degree of the two polynomials $F_1, F_2 \in \mathcal{I}_q$ (if they exist) such that

$$\begin{aligned} F_1^{Q_{G_{u,v}}}(x) &= (x-u)^2 \\ F_2^{Q_{G_{u,v}}}(x) &= (x-v)^2 \end{aligned}$$

by Theorem 10. We chose $Q_{G_{u,v}}$ so that $(x-u)^{Q_{G_{u,v}}} = (x-u)^2$ and $(x-v)^{Q_{G_{u,v}}} = (x-v)^2$, thus $d(G_{u,v}) = 1$.

Let $F \in \mathcal{I}_q$, then $F^{Q_{G_{u,v}}}$ is irreducible if and only if

$$\begin{aligned} P(x) &:= (x^2 - uv) - \gamma(2x - (u+v)) \\ &= x^2 - 2\gamma x + (\gamma(u+v) - uv) \in \mathbb{F}_q(\gamma)[x] \end{aligned}$$

is irreducible for $\gamma \in \overline{\mathbb{F}}_q$ a root of F by Lemma 8. A quadratic polynomial over $\mathbb{F}_q(\gamma)$ is irreducible if and only if it has no roots in $\mathbb{F}_q(\gamma)$. For P this is equivalent

to $4\gamma^2 - 4(\gamma(u+v) - uv)$ being a non-square in $\mathbb{F}_q(\gamma) = \mathbb{F}_{q^{\deg(F)}}$. Hence

$$\begin{aligned}
-1 &= \left(\frac{4\gamma^2 - 4(\gamma(u+v) - uv)}{q^{\deg(F)}}\right) = \left(\frac{\gamma^2 - (u+v)\gamma + uv}{q^{\deg(F)}}\right) \\
&= \left(\frac{(\gamma-u)(\gamma-v)}{q^{\deg(F)}}\right) = ((\gamma-u)(\gamma-v))^{\frac{q^{\deg(F)}-1}{2}} \\
&= ((u-\gamma)(v-\gamma))^{\frac{q^{\deg(F)}-1}{q-1}\cdot\frac{q-1}{2}} \\
&= \left(\prod_{i=0}^{\deg(F)-1}(u-\gamma^{q^i})\right)^{\frac{q-1}{2}} \cdot \left(\prod_{i=0}^{\deg(F)-1}(v-\gamma^{q^i})\right)^{\frac{q-1}{2}} \\
&= F(u)^{\frac{q-1}{2}} \cdot F(v)^{\frac{q-1}{2}} = \left(\frac{F(u)\cdot F(v)}{q}\right).
\end{aligned}$$

This calculation is very similar to a calculation trick that Meyn used in [8, Proof of Theorem 8]. We showed that $C(Q_{G_{u,v}}, n) = \{f \in \mathcal{I}_q^n \mid \left(\frac{f(u)f(v)}{q}\right) = -1\}$, the rest of the proof of Theorem 3 follows from Theorem 4.

2.3 An Example similar to Theorem 3

Let $\mathbb{F}_q$ be an arbitrary finite field, s a prime dividing $q-1$ and $c \in \mathbb{F}_q$. Define

$$\begin{aligned}
T_c(n) &:= \left\{ f \in \mathcal{I}_q^n \mid f(c)^{\frac{q-1}{s}} \neq (-1)^{\frac{(q-1)n}{s}} \right\} \\
U_c(n) &:= \mathcal{I}_q^n \setminus T(n) = \left\{ f \in \mathcal{I}_q^n \mid f(c)^{\frac{q-1}{s}} = (-1)^{\frac{(q-1)n}{s}} \right\}.
\end{aligned}$$

We are going to prove the following theorem

Theorem 14. *Let $\mathbb{F}_q$ be an arbitrary finite field and s a prime dividing $q-1$. Moreover let $c \in \mathbb{F}_q$, then we have for all $n \geq 2$ that*

$$|T_c(n)| - (s-1)|U_c(n)| = \begin{cases} 0, & \text{if } s \nmid n \\ |T_c(n/s)|, & \text{if } s \mid n. \end{cases}$$

Before we start with the proof we need to formulate a lemma first, which is folklore.

Lemma 15. *Let $\mathbb{F}_q$ be an arbitrary finite field, $c \in \mathbb{F}_q^*$ and s a prime dividing $q-1$. The polynomial $x^s - c \in \mathbb{F}_q[x]$ is irreducible if and only if $c^{\frac{q-1}{s}} \neq 1$.*

Proof (Theorem 14). We consider the following matrix

$$A := \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}.$$

where $a \in \mathbb{F}_q^*$ has order $s > 1$ which is prime and $b \in \mathbb{F}_q$ is arbitrary. The group $G := \langle [A] \rangle$ has order s. The fixed points of $[A]$ are ∞ and $c := \frac{-b}{a-1}$ and these are

again the only non-regular orbits in $\overline{\mathbb{F}}_q \cup \{\infty\}$. Note that we can obtain every $c \in \mathbb{F}_q$ for fixed $a \in \mathbb{F}_q^* \setminus \{1\}$ by choosing $b = (1-a) \cdot c$. A quotient map for G is

$$Q_G(x) = (x-c)^s + c.$$

and $(x-c)^{Q_G} = (x-c)^s$, so $d(G) = 1$.

Let $F \in \mathcal{I}_q^n$, then $F^{Q_G}(x) = F(Q_G(x))$ is irreducible if and only if

$$P(x) = Q_G(x) - \gamma = (x-c)^s + c - \gamma \in \mathbb{F}_{q^n}[x]$$

is irreducible by Lemma 8, where γ is a root of F. The polynomial $P(x) \in \mathbb{F}_{q^n}[x]$ is irreducible if and only if $P(x+c) = x^s - (\gamma - c) \in \mathbb{F}_{q^n}[x]$ is irreducible. By Lemma 15 this is the case exactly when $(\gamma-c)^{(q^n-1)/s} \neq 1$. Now we calculate:

$$\begin{aligned} 1 \neq (\gamma-c)^{\frac{q^n-1}{s}} &= (-1)^{\frac{q^n-1}{q-1}\cdot\frac{q-1}{s}} \cdot (c-\gamma)^{\frac{q^n-1}{q-1}\cdot\frac{q-1}{s}} \\ &= \left(\prod_{i=0}^{n-1}(-1)^{q^i}\right)^{\frac{q-1}{s}} \cdot \left(\prod_{i=0}^{n-1}(c-\gamma^{q^i})\right)^{\frac{q-1}{s}} = (-1)^{\frac{(q-1)n}{s}} \cdot f(c)^{\frac{q-1}{s}}. \end{aligned}$$

Hence $C(Q_G, n) = T_c(n)$, the rest follows from Theorem 4 again.

Remark 16. 1. If we take $s = 2$ the condition in $T_c(n)$ is

$$(-1)^{\frac{(q-1)n}{2}} \neq f(c)^{\frac{q-1}{2}} = \left(\frac{f(c)}{q}\right).$$

This looks quite similar to the defining condition of $C_{u,v}(n)$ in Theorem 3.

2. In Theorem 14 we used the criterion of Lemma 15 for the irreducibility of binomials of the form $x^s - c$ where s is a prime dividing $q-1$. If the reader is interested in a recent paper that explains the factorization of polynomials $x^n - c \in \mathbb{F}_q[x]$ for arbitrary n we refer them to [5].

References

[1] Antonia W. Bluher. "Explicit Artin maps into PGL_2". In: *Expositiones Mathematicae* 40 (2022), pp. 45–93.

[2] Xiwang Cao and Lei Hu. "On the reducibility of some composite polynomials over finite fields". In: *Designs, Codes and Cryptography* 64 (Sept. 2012), pp. 229–239.

[3] Stephen D. Cohen. "On irreducible polynomials of certain types in finite fields". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 66 (1969), pp. 335–344.

[4] Theodoulos Garefalakis. "On the action of $GL_2(\mathbb{F}_q)$ on irreducible polynomials over $\mathbb{F}_q$". In: *Journal of Pure and Applied Algebra* 215 (Aug. 2011), pp. 1835–1843.

[5] Anna-Maurin Graner. *Closed formulas for the factorization of $X^n - 1$, the n-th cyclotomic polynomial, $X^n - a$ and $f(X^n)$ over a finite field for arbitrary positive integers n.* 2024. arXiv: `2306.11183 [math.NT]`.

[6] Robert Granger. "Three proofs of an observation on irreducible polynomials over GF(2)". In: *Finite Fields and Their Applications* 88 (2023).
[7] Giorgos Kapetanakis. "Prescribing coefficients of invariant irreducible polynomials". In: *Journal of Number Theory* 180 (2017), pp. 615–628.
[8] Helmut Meyn. "On the Construction of Irreducible Self-Reciprocal Polynomials Over Finite Fields". In: *AAECC 1* (1990), pp. 43–53.
[9] Daniel Panario, Lucas Reis, and Qiang Wang. "Construction of irreducible polynomials through rational transformations". In: *Journal of Pure and Applied Algebra* 224 (2020).
[10] Lucas Reis. "Möbius-like maps on irreducible polynomials and rational transformations". In: *Journal of Pure and Applied Algebra* 224 (May 2019), pp. 169–180.
[11] Lucas Reis. "On the existence and number of invariant polynomials". In: *Finite Fields and Their Applications* 61 (2020).
[12] Max Schulz. *On the Recursive Behaviour of the Number of Irreducible Polynomials with Certain Properties over Finite Fields*. 2023. arXiv: 2310.01872 [math.NT].
[13] Max Schulz. *Rational Transformations and Invariant Polynomials*. 2023. arXiv: 2306.13502 [math.NT].
[14] Henning Stichtenoth and Alev Topuzoğlu. "Factorization of a class of polynomials over finite fields". In: *Finite Fields and Their Applications* 18 (2012), pp. 108–122.

Stabilizers of graphs of linear functions and rank-metric codes

Valentino Smaldore[1], Corrado Zanella[1], and Ferdinando Zullo[2]

[1] Dipartimento di Tecnica e Gestione dei Sistemi Industriali
Università degli Studi di Padova
Stradella S. Nicola, 3
36100 Vicenza VI - Italy valentino.smaldore@unipd.it, corrado.zanella@unipd.it
[2] Dipartimento di Matematica e Fisica,
Università degli Studi della Campania "Luigi Vanvitelli",
Viale Lincoln 5,
81100 Caserta CE - Italy ferdinando.zullo@unicampania.it

Abstract. In this extended abstract we will study the action of $\mathbb{F}_{q^n}^{2\times 2}$ on the graph of an $\mathbb{F}_q$-linear function of $\mathbb{F}_{q^n}$ into itself. In particular we will see that, under certain combinatorial assumptions, its stabilizer (together with the sum and product of matrices) is a field. Moreover, we will establish a connection between such a stabilizer and the right idealizer of the rank-metric code defined by the linear function and give some structural results in the case in which the polynomials are partially scattered.

Keywords: rank-metric codes, partially scattered polynomials, graphs of functions

1 Rank-metric codes

Rank-metric codes have been originally introduced by Delsarte [8] in 1978. They have been intensively investigated in recent years because of their applications; we refer to [18] for a survey on this topic. The set $\mathbb{F}_q^{m\times n}$ of $m \times n$ matrices over $\mathbb{F}_q$ can be endowed with the *rank metric* defined by

$$d(A, B) = \mathrm{rk}\,(A - B).$$

A subset $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ equipped with the rank metric is called a *rank-metric code*. The *minimum distance* of $\mathcal{C}$ is defined as $d = \min\{d(A, B)\colon A, B \in \mathcal{C},\ A \neq B\}$. We will denote the parameters of a rank-metric code $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ with minimum distance d by $(m, n, q; d)$. Delsarte showed in [8] that the parameters of these codes must fulfill a Singleton-like bound.

Theorem 1 [8] *If $\mathcal{C}$ is a rank-metric code with parameters $(m, n, q; d)$, then*

$$|\,\mathcal{C}\,| \leq q^{\max\{m,n\}(\min\{m,n\}-d+1)}.$$

When equality holds, we say that $\mathcal{C}$ is a *maximum rank distance* (*MRD* for short) code.

We will be mainly interested in $\mathbb{F}_q$-*linear* rank-metric codes, that is $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^{m\times n}$. Two $\mathbb{F}_q$-linear rank-metric codes $\mathcal{C}$ and $\mathcal{C}'$ in $\mathbb{F}_q^{m\times n}$ are *equivalent* if and only if there exist $X \in \mathrm{GL}(m,q)$, $Y \in \mathrm{GL}(n,q)$, and a field automorphism σ of $\mathbb{F}_q$ such that $\mathcal{C}' = \{XC^\sigma Y \colon C \in \mathcal{C}\}$. The *left* and *right idealizers* of a rank-metric code $\mathcal{C} \subseteq \mathbb{F}_q^{m\times n}$ are defined as

$$L(\mathcal{C}) = \{Y \in \mathbb{F}_q^{m\times m} \colon YC \in \mathcal{C} \text{ for all } C \in \mathcal{C}\},$$

$$R(\mathcal{C}) = \{Z \in \mathbb{F}_q^{n\times n} \colon CZ \in \mathcal{C} \text{ for all } C \in \mathcal{C}\},$$

respectively. They are powerful tools to study the equivalence issue among rank-metric codes. These notions have been introduced by Liebhold and Nebe in [10, Definition 3.1]; they are invariant under equivalences of rank-metric codes. Further invariants have been introduced in [9, 15]. In [14], idealizers have been studied in details (under the name of *middle* and *right nuclei*) and the following result has been proved.

Theorem 2 [14] *Let $\mathcal{C}$ and $\mathcal{C}'$ be $\mathbb{F}_q$-linear rank-metric codes of $\mathbb{F}_q^{m\times n}$.*

- *If $\mathcal{C}$ and $\mathcal{C}'$ are equivalent, then their left and right idealizers are isomorphic as $\mathbb{F}_q$-algebras (*[14, Proposition 4.1]*).*
- *Let $\mathcal{C}$ be an MRD code with minimum distance $d > 1$. If $m \leq n$, then $L(\mathcal{C})$ is a finite field with $|L(\mathcal{C})| \leq q^m$. If $m \geq n$, then $R(\mathcal{C})$ is a finite field with $|R(\mathcal{C})| \leq q^n$. In particular, when $m = n$, $L(\mathcal{C})$ and $R(\mathcal{C})$ are both finite fields (*[14, Theorem 5.4 and Corollary 5.6]*).*

We may see the nonzero elements of an $\mathbb{F}_q$-linear rank-metric code $\mathcal{C}$ with parameters $(m, n, q; d)$ as:

- matrices of $\mathbb{F}_q^{m\times n}$ having rank at least d and with at least one matrix of rank exactly d;
- $\mathbb{F}_q$-linear maps $V \to W$ where $V = V(n,q)$ and $W = V(m,q)$, having usual map rank at least d and with at least one map of rank exactly d;
- when $m = n$, elements of the $\mathbb{F}_q$-algebra $\mathcal{L}_{n,q}$ of q-polynomials over $\mathbb{F}_{q^n}$ modulo $x^{q^n} - x$, having rank at least d and with at least one polynomial of rank exactly d, where the rank is just the rank of the associated matrix.

2 Linearized polynomials and linear sets

A *q-polynomial* (or *linearized polynomial*) over the finite field $\mathbb{F}_{q^n}$ has form $f = \sum_{i=0}^{k} a_i x^{q^i} \in \mathbb{F}_{q^n}[x]$; if $a_k \neq 0$ then the *q-degree* of f is k. The set of linearized polynomials over $\mathbb{F}_{q^n}$ will be denoted as $L_{n,q}$. Such set, equipped with the operations of sum, multiplication by elements of $\mathbb{F}_q$ and the composition, results to be an $\mathbb{F}_q$-algebra. The quotient algebra $\mathcal{L}_{n,q} = L_{n,q}/(x^{q^n} - x)$ has

the property that its elements are in one-to-one correspondence with the $\mathbb{F}_q$-linear endomorphisms of $\mathbb{F}_{q^n}$. Note that we can identify the elements of $\mathcal{L}_{n,q}$ with the q-polynomials having q-degree smaller than n. Following [12], let f be a q-linearized polynomial over $\mathbb{F}_{q^n}$, t a divisor of n such that $1 < t < n$, so that $n = tt'$. We say that f is L-q^t-*partially scattered* if for any $y, z \in \mathbb{F}_{q^n}^*$,

$$\frac{f(y)}{y} = \frac{f(z)}{z} \Longrightarrow \frac{y}{z} \in \mathbb{F}_{q^t}, \tag{1}$$

and that f is R-q^t-*partially scattered* if for any $y, z \in \mathbb{F}_{q^n}^*$,

$$\frac{f(y)}{y} = \frac{f(z)}{z} \text{ and } \frac{y}{z} \in \mathbb{F}_{q^t} \Longrightarrow \frac{y}{z} \in \mathbb{F}_q. \tag{2}$$

A polynomial f which is both L-q^t-partially scattered and R-q^t-partially scattered is called *scattered* (see [17]).

2.1 Graphs of functions

Let $f \in \mathbb{F}_q[x]$, the *graph* of f is defined as the following set of affine points

$$\mathcal{G}_f = \{(y, f(y)) \colon y \in \mathbb{F}_q\} \subseteq \mathrm{AG}(2, q).$$

We can see the projective plane $\mathrm{PG}(2, q)$ as the union of $\mathrm{AG}(2, q)$ and the line at infinity ℓ_∞. In coordinates, the points of $\mathrm{PG}(2, q)$ are of the form $\langle (x_0, x_1, x_2) \rangle_{\mathbb{F}_q}$ for some $(x_0, x_1, x_2) \in \mathbb{F}_q^3 \setminus \{(0, 0, 0)\}$ and we may assume that ℓ_∞ is the set of points defined by the vectors with last component equal to zero and so the points in $\mathrm{AG}(2, q)$ are those of the form $\langle (a, b, 1) \rangle_{\mathbb{F}_q}$ for some $a, b \in \mathbb{F}_q$, which in $\mathrm{AG}(2, q)$ is defined by the pair (a, b).

The set of *directions* of $f \in \mathbb{F}_q[x]$ is defined as

$$\mathcal{D}_f = \{PQ \cap \ell_\infty \colon P, Q \in \mathcal{G}_f, \ \ P \neq Q\},$$

where PQ denotes the line through the points P and Q. Note that

$$\mathcal{D}_f = \{\langle (1, m, 0) \rangle_{\mathbb{F}_q} \colon m \in D_f\},$$

where D_f is the set of slopes of the lines used in $\mathcal{D}_f$, that is

$$D_f = \left\{ \frac{f(y) - f(z)}{y - z} \colon y, z \in \mathbb{F}_q, \ \ y \neq z \right\}.$$

Combinatorial conditions on $\mathcal{G}_f$ and/or $\mathcal{D}_f$ can give algebraic properties on f; see for instance the well-celebrated results in [1, 2] where some conditions on the intersections between $\mathcal{G}_f$ and the affine lines together with bounds on the number of directions yield some linearity conditions on f. Since f is $\mathbb{F}_q$-linear, the affine lines meet $\mathcal{G}_f$ in either zero points or in a power of q points. Here we investigate the natural action of the group $(\mathbb{F}_{q^n}^{2\times 2}, +)$ on $\mathcal{G}_f$ by considering the

set $\mathbb{S}_f = \{A \in \mathbb{F}_{q^n}^{2\times 2} \colon A\mathcal{G}_f \subseteq \mathcal{G}_f\}$, where $A\mathcal{G}_f = \left\{A\begin{pmatrix} y \\ f(y)\end{pmatrix} : y \in \mathbb{F}_{q^n}\right\}$. $\mathbb{S}_f$, together with $+$ and $\cdot$ the usual sum and product of matrices in $\mathbb{F}_{q^n}^{2\times 2}$ and $\star$ the multiplication by a scalar in $\mathbb{F}_q$, forms an $\mathbb{F}_q$-algebra. The proof relies on the following property.

Proposition 3 [19, Proposition 2.2] *If $A, B \in \mathbb{S}_f$, then $A+B,\ AB \in \mathbb{S}_f$.*

Let f and g be two linearized polynomials over $\mathbb{F}_{q^n}$ and consider the two related graphs $\mathcal{G}_f$ and $\mathcal{G}_g$ in $\mathrm{AG}(2,q^n)$. We will prove that when f has *low weight*, that is if for every affine line ℓ, $|\ell \cap \mathcal{G}_f| < q^{n/2}$, then $(\mathbb{S}_f, +, \cdot)$ is a field. We say that f and g are *equivalent* if there exists $\varphi \in \Gamma\mathrm{L}(2,q^n)$ such that $\varphi(\mathcal{G}_f) = \mathcal{G}_g$, that is, there exist $A \in \mathrm{GL}(2,q^n)$ and $\sigma \in \mathrm{Aut}(\mathbb{F}_{q^n})$ with the property that for each $x \in \mathbb{F}_{q^n}$ there exists $y \in \mathbb{F}_{q^n}$ satisfying

$$A\begin{pmatrix} x^\sigma \\ f(x)^\sigma\end{pmatrix} = \begin{pmatrix} y \\ g(y)\end{pmatrix},$$

see [3, Section 1] and [5, Section 1].

This definition of equivalence preserves the property of being R-q^t- and L-q^t-partially scattered:

Proposition 4 [4, Proposition 7.1] *Let f and g be two equivalent q-polynomials in $\mathcal{L}_{n,q}$. If f is* R-q^t*-partially scattered (resp.* L-q^t*-partially scattered), then g is* R-q^t*-partially scattered (resp.* L-q^t*-partially scattered).*

2.2 Linear sets

Let V be an r-dimensional $\mathbb{F}_{q^n}$-vector space and let $\Lambda = \mathrm{PG}(V, \mathbb{F}_{q^n}) = \mathrm{PG}(r-1, q^n)$. Let U be an $\mathbb{F}_q$-subspace of V such that $\dim_{\mathbb{F}_q}(U) = k$, then the set

$$L_U = \{\langle u\rangle_{\mathbb{F}_{q^n}} : u \in U \setminus \{0\}\} \subseteq \mathrm{PG}(r-1, q^n)$$

is said to be an $\mathbb{F}_q$-*linear set* of rank k.

Definition 5 *Consider a polynomial $f \in \mathcal{L}_{n,q}$.*

- *The linear set associated to f is*

$$L_f = L_{\mathcal{G}_f} = \{\langle (y, f(y))\rangle_{q^n} \mid y \in \mathbb{F}_{q^n}^*\}.$$

- *The* weight *of a point $P = \langle v\rangle_{q^n} \in \mathrm{PG}(1,q^n)$ in L_f is*

$$w_{L_f}(P) = \dim_q(\mathcal{G}_f \cap \langle v\rangle_{q^n}).$$

- *L_f is called* scattered *if all points of L_f have weight one.*
- *f is called* low weight *if all points of L_f have weight less than or equal to $\frac{n}{2}$.*

Note that the polynomial $f \in \mathcal{L}_{n,q}$ is scattered if and only if L_f is.

3 On the stabilizer of low weight polynomials

We study the action of the group $(\mathbb{F}_{q^n}^{2\times 2}, +)$ on the graph of a low weight linearized polynomial f, by showing that the stabilizer $\mathbb{S}_f$ of its graph is a field.

A motivation to study the structure of $\mathbb{S}_f$ regards the equivalence issue.

Lemma 6 [4, Lemma 7.2] *Let f and g be two q-polynomials over $\mathbb{F}_{q^n}$. If f and g are equivalent then $\mathbb{S}_f$ and $\mathbb{S}_g$ are isomorphic.*

In [19] we proved that when f is a low weight polynomial, then $(\mathbb{S}_f, +, \cdot)$ is a field.

Theorem 7 [19, Theorem 2.3] *Let f be a q-polynomial in $\mathcal{L}_{n,q}$. If f is a low weight polynomial, then $(\mathbb{S}_f, +, \cdot)$ is a field.*

Proof. (Sketch of Proof) By Proposition 3, it is enough to prove that for any rank-one 2×2 matrix M with elements in $\mathbb{F}_{q^n}$, $M\mathcal{G}_f$ is not contained in $\mathcal{G}_f$. Consider $Z \neq O$ such that $MZ = O$ and let C be a nonzero column of M. Define $\mu : \mathcal{G}_f \to \mathbb{F}_{q^n}^2$, $(y, f(y)) \mapsto M(y, f(y))^T$.
So, $\ker\mu \subseteq \langle Z\rangle_{q^n} \cap \mathcal{G}_f \Rightarrow dim_q(\ker\mu) < \frac{n}{2}$, then $\dim_q(Im\mu) > \frac{n}{2}$. Assume $M\mathcal{G}_f \subseteq \mathcal{G}_f$, then $Im\mu \subseteq \langle C\rangle_{q^n} \cap \mathcal{G}_f$, $\dim_q(Im\mu) < \frac{n}{2}$, contradiction.

Theorem 7 allows us to find a large class of polynomials for which $\mathbb{S}_f$ is a field.

Proposition 8 [19, Proposition 2.4] *Let f be a q-polynomial in $\mathcal{L}_{n,q}$. If f has q-degree k with $1 < k < n/2$ then it is a low weight polynomial. In particular $\mathbb{S}_f$ is a field.*

Proposition 9 [19, Proposition 3.4] *Let t be a nontrivial divisor of n.*

(*i*) *If f is a* R-q^t*-partially scattered polynomial in $\mathcal{L}_{n,q}$, then $w_{L_f}(P) \leq n/2$ for any point $P \in \mathrm{PG}(1, q^n)$.*
(*ii*) *If f is a* L-q^t*-partially scattered polynomial in $\mathcal{L}_{n,q}$, then $w_{L_f}(P) \leq t$ for any point $P \in \mathrm{PG}(1, q^n)$.*

There are indeed examples of partially scattered polynomials which are low weight polynomials.

4 Examples

In this section we give some examples of stabilizers of linearized polynomials. Firstly we list low weight polynomials.

The first family of low weight polynomials is given by the scattered polynomials, as in this case the maximum size of intersection between the related graph and the affine lines is q. We list here the known examples of polynomials for which $\mathbb{S}_f$ has been already determined, including also some non-scattered ones.

- $f = x^{q^s} \in \mathcal{L}_{n,q}$ with $\gcd(s,n) = 1$, then $|\mathbb{S}_f| = q^n$, see [6, Section 6];
- $f = \delta x^{q^s} + x^{q^{n(s-1)}} \in \mathcal{L}_{n,q}$ with $\gcd(s,n) = 1$, $\delta \neq 0$ and $n \geq 4$, then $|\mathbb{S}_f| = q^2$ if n is even and $|\mathbb{S}_f| = q$ if n is odd, see [6, Section 6] (we call these polynomials *LP polynomials* even in case they are not scattered);
- $f = \delta x^{q^s} + x^{q^{s+n/2}} \in \mathcal{L}_{n,q}$ with $\delta \neq 0$, n even and $\gcd(s,n) = 1$, then $|\mathbb{S}_f| = q^{n/2}$, see [6, Corollary 5.2];
- $f = x^q + x^{q^3} + \delta x^{q^5} \in \mathcal{L}_{6,q}$ with q odd and $\delta^2 + \delta = 1$, then $|\mathbb{S}_f| = q^2$, see [7, Proposition 5.2].
- $f = x^{q^s} + x^{q^{s(t-1)}} + \eta^{1+q^s} x^{q^{s(t+1)}} + \eta^{1-q^{s(2t-1)}} x^{q^{s(2t-1)}} \in \mathcal{L}_{n,q}$ with q odd prime power, $t, s, n \in \mathbb{N}$ with $n = 2t$, $t \geq 5$, $\gcd(s,n) = 1$ and $N_{q^n/q^t}(\eta) = -1$, then $|\mathbb{S}_f| = q^2$, see [13, Proposition 3.4].
- $f = x^{q^{s(t-1)}} + x^{q^{s(2t-1)}} + m(x^{q^s} - x^{q^{s(t+1)}}) \in \mathcal{L}_{n,q}$ with q odd prime power, $t, s, n \in \mathbb{N}$ with $n = 2t$, $t \geq 5$, $\gcd(s,n) = 1$, $m \in \mathbb{F}_q^t$, then $|\mathbb{S}_f| = q$ if t is odd and $m \neq 1$, $|\mathbb{S}_f| = q^2$ otherwise, see [20, Proposition 3.1].

Partially scattered polynomials are not far from being low weight polynomials.

Proposition 10 *Let t be a nontrivial divisor of n.*

(i) If f is a R-q^t-partially scattered polynomial in $\mathcal{L}_{n,q}$, then $w_{L_f}(P) \leq n/2$ for any point $P \in \mathrm{PG}(1, q^n)$.
(ii) If f is a L-q^t-partially scattered polynomial in $\mathcal{L}_{n,q}$, then $w_{L_f}(P) \leq t$ for any point $P \in \mathrm{PG}(1, q^n)$.

4.1 Non-low weight partially scattered polynomials

The next two results are listed in order to characterize the L-q^t-partially scattered polynomials f, and to give examples of R-q^t-partially scattered polynomials f whose related algebras $\mathbb{S}_f$ are not fields.

Theorem 11 *Let t be a proper divisor of n. Let $f \in \mathbb{F}_{q^n}[x]$ be an L-q^t-partially scattered polynomial in $\mathcal{L}_{n,q}$. Then $\mathbb{S}_f$ is not a field if and only if f is equivalent to $\ell^{q^t} - \ell$ for some $\ell \in \mathcal{L}_{t,q}$, and $n = 2t$.*

Example 12 *Let $p = \sum_{k=0}^{n-1} \left(\sum_{\ell=0}^{t-1} (u_\ell + u_\ell^{q^s} \xi) \lambda_\ell^{*q^k} \right) x^{q^k}$, where $\{u_0, \ldots, u_{t-1}\}$ is an $\mathbb{F}_q$-basis of $\mathbb{F}_{q^t}$ and $(\lambda_0^*, \ldots, \lambda_{n-1}^*)$ is the dual basis of $(u_0 + \mu u_0^{q^s} \xi, \ldots, u_{t-1} + \mu u_{t-1}^{q^s} \xi, u_0 + u_0^{q^s} \xi, \ldots, u_{t-1} + u_{t-1}^{q^s} \xi)$. Then p is an R-q^t-partially scattered polynomial and the stabilizer of $\mathcal{G}_p$ is not a field.*

5 Applications to two-dimensional linear rank-metric codes

When $m = n$, $\mathbb{F}_q$-linear rank-metric codes can be studied in terms of linearized polynomials. Indeed, an $\mathbb{F}_q$-linear rank-metric code $\mathcal{C}$ is an $\mathbb{F}_q$-subspace of $\mathcal{L}_{n,q}$

endowed with the rank metric and all the notions already given for rank-metric codes can be read in this context. For any linearized polynomial $f \in \mathcal{L}_{n,q}$, we can consider the following $\mathbb{F}_{q^n}$-linear rank-metric code

$$\mathcal{C}_f = \langle x, f \rangle_{\mathbb{F}_{q^n}}.$$

Sheekey in [17] pointed out that $\mathcal{C}_f$ is an MRD code if and only if f is a scattered polynomial. When $\mathcal{C}_f$ is an MRD code, we already know that its right idealizer is a field, cf. Theorem 2. In the next result we will see that we can relax the condition of being MRD codes when considering two-dimensional $\mathbb{F}_{q^n}$-linear rank-metric codes. To this aim we first prove a relation between $\mathbb{S}_f$ and $R(\mathcal{C}_f)$, extending [11, Lemma 4.1] where the result was proved under the assumption that f is scattered.

Theorem 13 [19, Theorem 4.1] *Let $f \in \mathcal{L}_{n,q}$ and denote by $\mathcal{C}_f$ the associated rank-metric code in $\mathcal{L}_{n,q}$. Suppose that $f \notin \langle x \rangle_{\mathbb{F}_{q^n}}$. Then the $\mathbb{F}_q$-algebras $\mathbb{S}_f$ and $R(\mathcal{C}_f)$ are isomorphic.*

Proof. (Sketch of Proof) The proof relies on showing that the following is an isomorphism

$$\psi \colon \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto ax + bf.$$

As a consequence we obtain the following.

Corollary 14 *Let f be a linearized polynomial in $\mathcal{L}_{n,q}$. If $d(\mathcal{C}_f) > n/2$, then $R(\mathcal{C}_f)$ is a field.*

Proof. By [16, Theorem 2] it follows that $w_{L_f}(P) < n/2$ for any point P, that is f has low weight, and the assertion then follows from Theorems 7 and 13.

Remark 15 *The above corollary is part of* [14, Lemma 4.4].

5.1 The right idealizer of the MRD codes associated with partially scattered polynomials

In Theorem 3.3 of [12], the authors showed that if $n = tt'$ and $f \in \mathcal{L}_{n,q}$ is an R-q^t-partially scattered polynomial then

$$\tilde{\mathcal{C}}_f = \{F_{|\mathbb{F}_{q^t}} : \mathbb{F}_{q^t} \to \mathbb{F}_{q^n} : F \in \mathcal{C}_f\}$$

is an MRD code with parameters $(n, t, q; t-1)$ The left idealizer $L(\tilde{\mathcal{C}}_f)$ contains a copy of $\mathbb{F}_{q^n}$ and, since $t \leq n$, $R(\tilde{\mathcal{C}}_f)$ is a field with $|R(\tilde{\mathcal{C}}_f)| \leq q^t$ by Theorem 2.

In the next results we find a relation between the right idealizer of $\mathcal{C}_f$ and those of $\tilde{\mathcal{C}}_f$.

Let $\mathcal{L}_{t,n,q} = \{g \in \mathcal{L}_{n,q} \mid g(\mathbb{F}_{q^t}) = \mathbb{F}_{q^t}\}$ be the $\mathbb{F}_q$-vector space of the $\mathbb{F}_q$-endomorphisms of $\mathbb{F}_{q^n}$ which fix setwise $\mathbb{F}_{q^t}$. Then consider the equivalence

relation $\approx$, such that $g \approx g'$ if and only if $g_{|\mathbb{F}_{q^t}} = g'_{|\mathbb{F}_{q^t}}$. Then the projection map

$$\tilde{\pi} : \mathcal{L}_{n,q} \longrightarrow \mathcal{L}_{n,q}/\approx,$$

maps $\mathcal{L}_{t,n,q}$ onto a vector space isomorphic to $\mathcal{L}_{t,q}$, and let

$$\Phi : \tilde{\pi}(g) \in \mathcal{L}_{t,n,q}/\approx \longrightarrow g_{|\mathbb{F}_{q^t}} \in \mathcal{L}_{t,q}.$$

Theorem 16 [19, Theorem 4.8] *Let $f \in \mathcal{L}_{n,q}$ with $f \notin \langle x \rangle_{\mathbb{F}_{q^n}}$ and such that f is* R-q^t*-partially scattered. Consider*

$$\mathcal{C}_f = \langle x, f \rangle_{\mathbb{F}_{q^n}} \subseteq \mathcal{L}_{n,q}$$

and

$$\tilde{\mathcal{C}}_f = \{F_{|\mathbb{F}_{q^t}} : \mathbb{F}_{q^t} \to \mathbb{F}_{q^n} : F \in \mathcal{C}_f\} \subseteq \mathrm{Hom}_{\mathbb{F}_q}(\mathbb{F}_{q^t}, \mathbb{F}_{q^n}).$$

Then

$$(\Phi \circ \tilde{\pi})(R(\mathcal{C}_f) \cap \mathcal{L}_{t,n,q}) \subseteq R(\tilde{\mathcal{C}}_f).$$

Corollary 17 [19, Corollary 4.9] *Let $f \in \mathcal{L}_{n,q}$ with $f \notin \langle x \rangle_{\mathbb{F}_{q^n}}$ and such that f is* R-q^t*-partially scattered. Then*

$$|R(\tilde{\mathcal{C}}_f)| \geq |\mathcal{L}_{t,n,q} \cap R(\mathcal{C}_f)|. \tag{3}$$

Bibliography

[1] S. Ball: The number of directions determined by a function over a finite field, *J. Combin. Theory Ser. A* **104** (2003) 341–350.

[2] S. Ball, A. Blokhuis, A.E. Brouwer, L. Storme and T. Szőnyi: On the number of slopes of the graph of a function defined over a finite field, *J. Combin. Theory Ser. A* **86** (1999) 187–196.

[3] D. Bartoli, G. Micheli, G. Zini and F. Zullo: r-fat linearized polynomials, *J. Combin. Theory Ser. A* **189** (2022): 105609.

[4] D. Bartoli, G. Zini and F. Zullo: Investigating the exceptionality of scattered polynomials, *Finite Fields Appl.* **77** (2022), 101956.

[5] B. Csajbók, G. Marino and O. Polverino: A Carlitz type result for linearized polynomials, *Ars Math. Contemp.* **16(2)** (2019), 585–608.

[6] B. Csajbók, G. Marino, O. Polverino and C. Zanella: A new family of MRD-codes, *Linear Algebra Appl.* **548** (2018), 203–220.

[7] B. Csajbók, G. Marino and F. Zullo: New maximum scattered linear sets of the projective line, *Finite Fields Appl.* **54** (2018), 133–150.

[8] P. Delsarte: Bilinear forms over a finite field, with applications to coding theory, *J. Combin. Theory Ser. A* **25** (1978), 226–241.

[9] L. Giuzzi and F. Zullo: Identifiers for MRD-codes, *Linear Algebra Appl.* **575** (2019), 66–86.

[10] D. Liebhold and G. Nebe: Automorphism groups of Gabidulin-like codes, *Arch. Math.* **107(4)** (2016), 355–366.

[11] G. Longobardi, G. Marino, R. Trombetti and Y. Zhou: A large family of maximum scattered linear sets of $\mathrm{PG}(1, q^n)$ and their associated MRD codes, *Combinatorica* **43** (2023), 681-716.

[12] G. Longobardi, C. Zanella: Partially scattered linearized polynomials and rank metric codes, *Finite Fields Appl.* **76** (2021), 101914.

[13] G. Longobardi, C. Zanella: A standard form for scattered linearized polynomials and properties of the related translation planes, https://arxiv.org/abs/2205.15429.

[14] G. Lunardon, R. Trombetti and Y. Zhou: On kernels and nuclei of rank metric codes, *J. Algebraic Combin.* **46** (2017), 313–340.

[15] A. Neri, S. Puchinger, A. Horlemann-Trautmann: Equivalence and Characterizations of Linear Rank-Metric Codes Based on Invariants, *Linear Algebra Appl.* **603** (2020), 418-469.

[16] T.H. Randrianarisoa: A geometric approach to rank metric codes and a classification of constant weight codes, *Des. Codes Cryptogr.* **88** (2020), 1331–1348.

[17] J. Sheekey: A new family of linear maximum rank distance codes, *Adv. Math. Commun.* **10(3)** (2016), 475–488.

[18] J. Sheekey: MRD codes: constructions and connections, *Combinatorics and finite fields: Difference sets, polynomials, pseudorandomness and applications*, Radon Series on Computational and Applied Mathematics **23**, K.-U. Schmidt and A. Winterhof (eds.), De Gruyter (2019).

[19] V. Smaldore, C. Zanella, F. Zullo: On the stabilizer of the graph of linear functions over finite fields, https://arxiv.org/abs/2401.06085.

[20] V. Smaldore, C. Zanella, F. Zullo: New scattered linearized quadrinomials, https://arxiv.org/abs/2402.14742.

Further Investigation on Differential Properties of the Generalized Ness-Helleseth Mapping

Yongbo Xia*, Furong Bao*, Shaoping Chen†, Chunlei Li ‡, and Tor Helleseth ‡

Abstract

Let n be an odd positive integer, p be a prime with $p \equiv 3 \pmod 4$, $d_1 = \frac{p^n-1}{2} - 1$ and $d_2 = p^n - 2$. The mapping from $\mathbb{F}_{p^n}$ to itself defined by $f_u(x) = ux^{d_1} + x^{d_2}$ is called the generalized Ness-Helleseth mapping, where $u \in \mathbb{F}_{p^n}$. It was initially studied by Ness and Helleseth in the ternary case. In this paper, for $p^n \equiv 3 \pmod 4$ and $p^n \geq 7$, we provide the necessary and sufficient condition for $f_u(x)$ to be an APN function. In addition, for each u satisfying $\chi(u+1) = \chi(u-1)$, the differential spectrum of $f_u(x)$ is investigated, and it is expressed in terms of some quadratic character sums involving cubic polynomials, where $\chi(\cdot)$ denotes the quadratic charactor of $\mathbb{F}_{p^n}$.

1 Introduction

Let $\mathbb{F}_{p^n}$ be the finite field with p^n elements and $\mathbb{F}_{p^n}^* = \mathbb{F}_{p^n} \setminus \{0\}$, where p is a prime and n is a positive integer. Let $F(x)$ be a mapping from $\mathbb{F}_{p^n}$ to itself. The *derivative function* of $F(x)$ at an element $a \in \mathbb{F}_{p^n}$, denoted by $\mathbb{D}_a F$, is given by

$$\mathbb{D}_a F(x) = F(x+a) - F(x).$$

For any a, $b \in \mathbb{F}_{p^n}$, let $\delta_F(a,b) = |\{x \in \mathbb{F}_{p^n} \mid \mathbb{D}_a F(x) = b\}|$, where $|S|$ denotes the cardinality of a set S. The *differential uniformity* of $F(x)$, denoted by $\delta(F)$, is defined as

$$\delta(F) = \max\{\delta_F(a,b) \mid a \in \mathbb{F}_{p^n}^*,\ b \in \mathbb{F}_{p^n}\}.$$

*Y. Xia and F. Bao are with the Department of Mathematics and Statistics, South-Central University for Nationalities, Wuhan 430074, China (xia@mail.scuec.edu.cn; baokekebpsp@163.com). Y. Xia is also with the Hubei Key Laboratory of Intelligent Wireless Communications, South-Central University for Nationalities, Wuhan 430074, China.

†S. Chen is with the Hubei Key Laboratory of Intelligent Wireless Communications, South-Central University for Nationalities, Wuhan 430074, China (spchen@scuec.edu.cn).

‡C. Li and T. Helleseth are with the Department of Informatics, University of Bergen, N-5020 Bergen, Norway (e-mail: chunlei.li@uib.no; tor.helleseth@uib.no).

A function $F(x)$ is said to be *differentially δ-uniform* if $\delta(F) = \delta$. Differential uniformity is an important concept in cryptography introduced by Nyberg [9], which can be used to quantify the security of the block cipher with respect to the differential attack if $F(x)$ used in an S-box. The lower the differential uniformity of $F(x)$ is, the stronger it is to resist the differential attack. When $\delta(F) = 1$, $F(x)$ is said to be *a perfect nonlinear (PN) function*. When $\delta(F) = 2$, $F(x)$ is said to be *an almost perfect nonlinear (APN) function*. PN and APN functions are important in cryptography [9] and also useful in coding theory [3, 4], mathematics [5, 2] and combinatorics [6]. Recent research on PN and APN functions can be found in [1] and the references therein.

Besides the differential uniformity, there is another concept that is used to measure the differential property of a nonlinear function more precisely. This concept is called the *differential spectrum* and is defined as follows.

Definition 1 *Let $F(x)$ be a mapping over $\mathbb{F}_{p^n}$ with differential uniformity δ, and define*

$$\omega_i = |\{(a, b) \in \mathbb{F}_{p^n}^* \times \mathbb{F}_{p^n} \mid \delta_F(a, b) = i\}|, \ 0 \le i \le \delta.$$

The differential spectrum of $F(x)$ is defined to be an ordered sequence

$$\mathbb{S} = [\omega_0, \omega_1, \ldots, \omega_\delta].$$

According to the Definition 1, the differential spectrum of a nonlinear mapping $F(x)$ with $\delta(F) = \delta$ satisfies the following two identities:

$$\sum_{i=0}^{\delta} \omega_i = (p^n - 1)p^n \text{ and } \sum_{i=0}^{\delta} (i \times \omega_i) = (p^n - 1)p^n, \tag{1}$$

which play an important role in determining the differential spectrum. It is an absorbing topic to determine the differential spectra of the nonlinear mappings with low differential uniformity. However, this problem is relatively challenging. Up to now, only for some power mappings and a few polynomials, the differential spectra were calculated. Such results can be found in [11, 12, 13] and the references therein.

Let n be an odd positive integer, p be an odd prime satisfying $p \equiv 3 \pmod 4$, $d_1 = \frac{p^n-1}{2} - 1$ and $d_2 = p^n - 2$. Let

$$f_u(x) = ux^{d_1} + x^{d_2}, \tag{2}$$

where $u \in \mathbb{F}_{p^n}$. For $p = 3$, the mapping $f_u(x)$ was initially studied by Ness and Helleseth in [8], and was further investigated in [11], where it was called *the ternary Ness-Helleseth mapping* and its differential uniformity was completely determined. For general p, we call this mapping *the generalized Ness-Helleseth mapping*. When $p \geq 7$, the differential properties of $f_u(x)$ were partially studied by Zeng et al. in [14] and by Zha in his PhD thesis [15]. Let $\chi(\cdot)$ denote the quadratic character of $\mathbb{F}_{p^n}$. Their results were summarized as follows.

Theorem 1 *[14] Let n be an odd integer, $p \equiv 3 \pmod 4$ and $p \geq 7$. Let u be an element in $\mathbb{F}_{p^n}$ such that $\chi(u+1) = \chi(u-1) = -\chi(5u+3)$ or $\chi(u+1) = \chi(u-1) = -\chi(5u-3)$. Then, the generalized Ness-Helleseth mapping $f_u(x)$ defined in (2) is an APN mapping, where $\chi(\cdot)$ denotes the quadratic character of $\mathbb{F}_{p^n}$.*

Theorem 2 *[15] Let n be an odd integer, $p \equiv 3 \pmod 4$ and $p \geq 7$. Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2). Then, the differential uniformity of $f_u(x)$ is equal to 3 when u satisfies $\chi(u+1) = \chi(u-1) \neq -\chi(5u+3)$ and $\chi(u+1) = \chi(u-1) \neq -\chi(5u-3)$, and is at most 4 when u satisfies $\chi(u+1) \neq \chi(u-1)$.*

Since the generalized Ness-Helleseth mapping in the ternary case was well studied in [8] and [11], we only focus on the case $p \geq 7$ in this paper. According to Theorems 1 and 2, in order to completely characterize the differential uniformity of $f_u(x)$, we still need to compute the exact differential uniformity of $f_u(x)$ for u satisfying $\chi(u+1) \neq \chi(u-1)$. After this goal is achieved, we can characterize the necessary and sufficient condition for $f_u(x)$ to be APN. Moreover, we will compute the differential spectrum of $f_u(x)$ for some u in this paper.

For convenience, we introduce the following sets

$$\begin{aligned}
\mathcal{U}_1 = \{u \in \mathbb{F}_{p^n} \mid \chi(u+1) = \chi(u-1) = -\chi(5u+3) \text{ or } \\
\chi(u+1) = \chi(u-1) = -\chi(5u-3)\}, \\
\mathcal{U}_2 = \{u \in \mathbb{F}_{p^n} \mid \chi(u+1) = \chi(u-1) \neq -\chi(5u+3) \text{ and } \\
\chi(u+1) = \chi(u-1) \neq -\chi(5u-3)\}, \\
\mathcal{U}_3 = \{u \in \mathbb{F}_{p^n} \mid \chi(u+1) \neq \chi(u-1)\}.
\end{aligned} \tag{3}$$

Note that $\mathcal{U}_i$, $i = 1, 2, 3$, are pairwise disjoint and $\cup_{i=1}^{3} \mathcal{U}_i = \mathbb{F}_{p^n}$. We also require the following three quadratic character sums to express our main results in the sequel

$$\Gamma_{p,n} = \sum_{x\in\mathbb{F}_{p^n}} \chi\left(x(x+1)(x+4)\right),$$

$$\Gamma_1(u) = \sum_{x\in\mathbb{F}_{p^n}} \chi\left((u+1)^2x^3 + (u^2-2u-2)x^2 + (1-u^2)x\right), \tag{4}$$

$$\Gamma_2(u) = -\sum_{x\in\mathbb{F}_{p^n}} \chi\left((u+1)x^3 - 4(u+2)(u+1)x^2 + 4(u+2)^2(u+1)x - 16u^2(u+1)^2\right).$$

Let $u \in \mathcal{U}_3$ and define

$$\begin{cases} g_1(x) = -(u+1)x, \\ g_2(x) = x^2 - 4(u+1)x, \\ g_3(x) = x^2 + 4(u-1)x, \\ g_4(x) = x^2 - 4x + 4u^2, \\ g_5(x) = (2+2\sqrt{1-u^2})x - 4u^2. \end{cases} \tag{5}$$

When $u \in \mathcal{U}_3$, each of the polynomials $g_2(x)$, $g_3(x)$ and $g_4(x)$ can be factored into a product of two linear polynomials over $\mathbb{F}_{p^n}$. Let I be an arbitrary subset of $\{1, 2, \cdots, 5\}$. In order to make sure that every product $\prod_{i\in I} g_i(x)$ cannot be written in the form of $c(g(x))^2$ for some polynomial $g(x)$ and some constant c, it is required that $u \in \mathcal{U}_3 \setminus \mathcal{U}_0$, where $\mathcal{U}_0 = \{0, \pm 1, \pm\frac{4}{5}\}$ and it is a subset of $\mathcal{U}_3$.

2 The differential uniformity of $f_u(x)$ for $u \in \{0, \pm 1\}$

Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2) and $p \geq 7$. According to Theorems 1 and 2, for $u \in \mathcal{U}_1$, $f_u(x)$ is an APN mapping, and when $u \in \mathcal{U}_2$, the differential uniformity of $f_u(x)$ is equal to 3. In the sequel, we will focus on the situation $u \in \mathcal{U}_3$. Note that $\{0, \pm 1\} \subset \mathcal{U}_3$. We deal with the special case $u \in \{0, \pm 1\}$ separately.

When $u = 0$, we have $f_0(x) = x^{p^n-2}$. The differential uniformity of $f_0(x)$ was studied by Dobbertin et al. in [7], which is 2 when $p^n \equiv 2 \pmod 3$ and 4 when $p^n \equiv 1 \pmod 3$. Later, the differential spectrum of $f_0(x)$ was determined by Zhang in his Master thesis [16]. Note that $f_u(-x) = -f_{-u}(x)$, and thus, $f_u(x)$ and $f_{-u}(x)$ have the same differential properties. For $u \in \{\pm 1\}$, we have the following result.

Proposition 1 *Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2) and $p \geq 7$. Then $f_1(x)$ and $f_{-1}(x)$ have the same differential uniformity $\frac{p^n+1}{4}$. Moreover the differential spectrum of $f_{\pm 1}(x)$ is*

$$[\,\omega_0 = \frac{(p^n-1)(p^n+1-\Gamma_{p,n})}{8},\ \omega_1 = \frac{(p^n-1)(2p^n-2+\Gamma_{p,n})}{4},$$
$$\omega_2 = \frac{(p^n-1)(p^n+1-\Gamma_{p,n})}{8},\ \omega_3 = \cdots = \omega_{\frac{p^n-3}{4}} = 0,\ \omega_{\frac{p^n+1}{4}} = (p^n-1)\,],$$

where $\Gamma^{(1)}_{p,n}$ is defined in (4) and $|\Gamma^{(1)}_{p,n}| \leq 2p^{n/2}$ by the Weil bound.

3 The conditions for the differential equation of $f_u(x)$ to have four solutions

For given $(a,b) \in \mathbb{F}_{p^n}^* \times \mathbb{F}_{p^n}$, the differential equation of $f_u(x)$ is given by

$$\mathbb{D}_a f_u(x) = f_u(x+a) - f_u(x) = b. \tag{6}$$

Let $N(a,b)$ denote the number of solutions of (6) in $\mathbb{F}_{3^n}$. According to Theorem 2, when $u \in \mathcal{U}_3 \setminus \{0, \pm 1\}$, we have $N(a,b) \leq 4$. Moreover, we can prove the following result.

Proposition 2 *Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2). When $u \in \mathcal{U}_3 \setminus \{0, \pm 1\}$, the differential equation $\mathbb{D}_a f_u(x) = b$ has four solutions if and only if there exist pairs $(a,b) \in \mathbb{F}_{p^n}^* \times \mathbb{F}_{p^n}$ satisfying the following conditions*

$$\begin{cases} \chi\left(\frac{a(u+1)}{b}\right) = -1, \\ \chi\left(a^2b^2 - 4(u+1)ab\right) = 1, \\ \chi\left(a^2b^2 + 4(u-1)ab\right) = 1, \\ \chi\left(4u^2 + a^2b^2 - 4ab\right) = 1, \\ \chi\left(-4u^2 + 2ab + 2ab\sqrt{1-u^2}\right) = 1. \end{cases} \tag{7}$$

4 The differential uniformity of $f_u(x)$ for $u \in \mathcal{U}_3 \setminus \{0, \pm 1\}$

4.1 The differential uniformity of $f_u(x)$ for $u \in \mathcal{U}_3 \setminus \{0, \pm 1, \pm\frac{4}{5}\}$

Let $\mathcal{A}$ be the set of the 3-tuples (p, n, u) given in Table 1, where the elements u belong to the corresponding set $\mathcal{U}_3 \setminus \{0, \pm 1, \pm\frac{4}{5}\}$. Based on Proposition 2, we can use quadratic character sums to count the number of $(a, b) \in \mathbb{F}_{p^n}^* \times \mathbb{F}_{p^n}$ satisfying the conditions (7) when $u \in \mathcal{U}_3 \setminus \{0, \pm 1, \pm\frac{4}{5}\}$. Then, by evaluating the character sums, we can show the following result.

Proposition 3 *Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2). When $p \geq 7$ and $u \in \mathcal{U}_3 \setminus \{0, \pm 1, \pm\frac{4}{5}\}$, the differential uniformity $\delta(f_u)$ of $f_u(x)$ is given as follows:*

$$\delta(f_u) = \begin{cases} 3, & \text{if } (p, n, u) \in \mathcal{A}, \\ 4, & \text{otherwise}, \end{cases} \tag{8}$$

where $\mathcal{A}$ is the set of the 3-tuples (p, n, u) given in Table 1.

Table 1: The differential uniformity of $f_u(x)$ for given (p, n, u)

(p, n, u)	$\delta(f_u)$	(p, n, u)	$\delta(f_u)$
$(11, 1, \pm 5)$	3	$(19, 1, \pm 2)$	3
$(23, 1, \pm 4)$	3	$(31, 1, \pm 10)$	3
$(31, 1, \pm 13)$	3	$(47, 1, \pm 11)$	3
$(59, 1, \pm 15)$	3	$(71, 1, \pm 13)$	3
$(83, 1, \pm 4)$	3	$(83, 1, \pm 38)$	3
$(151, 1, \pm 22)$	3		

4.2 The differential uniformity of $f_u(x)$ for $u = \pm\frac{4}{5}$

Note that $f_u(x)$ and $f_{-u}(x)$ have the same differential properties. Hence, for $u = \pm\frac{4}{5}$, we only need to investigate the differential uniformity of $f_{\frac{4}{5}}(x)$. With the notation introduced above, we can show that when $u = \frac{4}{5}$, the differential equation $\mathbb{D}_a f_u(x) = b$ has at most three solutions, and $N(a, b) = 3$ if and only if one of the following two conditions holds:

(i)

$$\begin{cases} \chi\left(-(u+1)ab\right) = 1, \\ \chi\left(a^2b^2 - 4(u+1)ab\right) = 1, \\ \chi\left(a^2b^2 - 4ab + 4u^2\right) = 1, \\ \chi\left(\varphi(u)ab - 4u^2\right) = 1, \end{cases} \tag{9}$$

where $u = \frac{4}{5}$ and $\varphi(u) = 2 + 2\sqrt{1-u^2}$;

(ii)

$$\chi(a^2b^2 - 4ab + 4u^2) = 1 \text{ and } \chi(\varphi(u)ab - 4u^2) = 1, \tag{10}$$

where $u = \frac{4}{5}$ and $ab = 1 + u$ or $1 - u$.

Note that there is no pair $(a, b) \in \mathbb{F}_{p^n}^* \times \mathbb{F}_{p^n}$ satisfying (9) and (10) simultaneously since $\chi(-(u+1)^2) = -1$ and $\chi(-(1-u^2)) = -1$. Moreover, when $u = \frac{4}{5}$, we can show that there always exists $(a, b) \in \mathbb{F}_{p^n}^* \times \mathbb{F}_{p^n}$ satisfying the conditions (9) or (10). This leads to the following result.

Proposition 4 *Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2) and $p \geq 7$. When $u = \pm\frac{4}{5}$, the differential uniformity of $f_u(x)$ is equal to 3.*

5 The differential uniformity of $f_u(x)$

Let n be an odd integer, $p \geq 7$ be an odd prime with $p \equiv 3 \pmod 4$ and $f_u(x)$ be the function defined in (2). According to Proposition 1, 3 and 4, we can obtain the following result.

Theorem 3 *Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2) and $p \geq 7$. Then, the differential uniformity of $f_u(x)$ for $u \in \mathcal{U}_3$ is given as follows*

$$\delta(f_u(x)) = \begin{cases} 2, & \text{if } u = 0 \text{ and } p^n \equiv 2 \pmod 3, \\ 4, & \text{if } u = 0 \text{ and } p^n \equiv 1 \pmod 3, \\ \frac{p^n+1}{4}, & \text{if } u = \pm 1, \\ 3, & \text{if } u = \pm\frac{4}{5}, \\ 3, & \text{if } (p, n, u) \in \mathcal{A}, \\ 4, & \text{otherwise}, \end{cases}$$

where $\mathcal{A}$ is the set of the 3-tuples (p, n, u) given in Table 1.

Theorem 3 together with Theorems 1 and 2 gives the differential uniformity of the generalized Ness-Helleseth mapping for each $u \in \mathbb{F}_{p^n}$. Based on these results, we get the necessary and sufficient condition for $f_u(x)$ to be APN as follows.

Corollary 1 *Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2) and $p \geq 7$. Then, $f_u(x)$ is APN if and only if n, p and u satisfy one of the following conditions:*
(i) $p^n \equiv 2 \pmod 3$ and $u = 0$;
(ii) $p = 7$, $n = 1$ and $u = \pm 1$;
(iii) $u \in \mathcal{U}_1$, where $\mathcal{U}_1$ is defined in (3).

6 The differential spectrum of $f_u(x)$ for $u \in \mathcal{U}_1 \cup \mathcal{U}_2$

6.1 The differential spectrum of $f_u(x)$ for $u \in \mathcal{U}_1$

Let $u \in \mathbb{F}_{p^n}$ satisfy $\chi(u+1) = \chi(u-1)$, i.e., $u \in \mathcal{U}_1 \cup \mathcal{U}_2$. Then, the differential equation (6) of $f_u(x)$ has exactly two solutions if and only if one of the following two conditions (i) and (ii) holds:

(i) $\chi\left(\frac{a(u+1)}{b}\right) = -1$, $\chi\left(a^2b^2 - 4(u+1)ab\right) = 1$ and $\chi\left(a^2b^2 - 4ab + 4u^2\right) = 1$.

(ii) $\chi\left(\frac{a(u+1)}{b}\right) = 1$, $\chi\left(a^2b^2 + 4(u-1)ab\right) = 1$ and $\chi\left(a^2b^2 - 4ab + 4u^2\right) = 1$.

Let $z = ab$, and let $N_1(u)$ and $N_2(u)$ denote the number of $z \in \mathbb{F}_{p^n}^*$ satisfying the conditions (i) and (ii), respectively. Then, we can show that

$$8(N_1(u) + N_2(u)) = 2p^n - 14 + \Gamma_1(u) + \Gamma_2(u) + \Gamma_1(-u) + \Gamma_2(-u), \tag{11}$$

where $\Gamma_1(u)$ and $\Gamma_2(u)$ are defined in (4).

Recall that when $u \in \mathcal{U}_1$, the generalized Ness-Helleseth mapping $f_u(x)$ defined in (2) is an APN mapping. Therefore, according to the identities in (1), we can express the differential spectrum of $f_u(x)$ for $u \in \mathcal{U}_1$ as follows.

Proposition 5 *Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2) and $p \geq 7$. When $u \in \mathcal{U}_1$, the differential spectrum of $f_u(x)$ can be expressed as*

$$[\,\omega_0 = (p^n-1)(N_1(u)+N_2(u)), \omega_1 = (p^n-1)(p^n-2(N_1(u)+N_2(u))), \omega_2 = (p^n-1)(N_1(u)+N_2(u))\,],$$

where $N_1(u) + N_2(u)$ is given in (11).

6.2 The differential spectrum of $f_u(x)$ for $u \in \mathcal{U}_2$

Recall that the mapping $f_u(x)$ is 3-uniform when $u \in \mathcal{U}_2$. Assume that in this case the differential spectrum of $f_u(x)$ is $[\omega_0, \omega_1, \omega_2, \omega_3]$. From the proof of Theorem 2 in [15], we know that in this case the differential equation $\mathbb{D}_a f_u(x) = b$ of $f_u(x)$ has three solutions if and only if $ab = 1 \pm u$. Thus, we have

$$\omega_3 = 2 \cdot (p^n - 1).$$

If $ab = 1 + u$, then the pairs (a, b) satisfy the condition (ii) in Subsection 6.1, and when $ab = 1 - u$, the pairs (a, b) satisfy the condition (i) in Subsection 6.1. Therefore, we have

$$\omega_2 = (p^n - 1)(N_1(u) + N_2(u)) - 2 \cdot (p^n - 1).$$

According to identities in (1), we get the following result.

Proposition 6 *Let $f_u(x)$ be the generalized Ness-Helleseth mapping defined in (2) and $p \geq 7$. For $u \in \mathcal{U}_2$, the differential spectrum of $f_u(x)$ can be expressed as*

$$\begin{bmatrix} \omega_0 = (p^n - 1)(N_1(u) + N_2(u) + 2), & \omega_1 = (p^n - 1)\left(p^n - 2 - 2(N_1(u) + N_2(u))\right), \\ \omega_2 = (p^n - 1)(N_1(u) + N_2(u) - 2), & \omega_3 = 2(p^n - 1) \end{bmatrix},$$

where $N_1(u) + N_2(u)$ is given in (11).

References

[1] C. Beierle and G. Leander, "New instances of quadratic APN functions," *IEEE Trans. Inf. Theory*, vol. 68, no. 1, pp. 670-678, 2022.

[2] J. Bierbrauer, "New semifields, PN and APN functions," *Des. Codes Cryptogr.*, vol. 54, no. 3, pp. 189-200, 2010.

[3] C. Carlet, P. Charpin, and V. Zinoviev, "Codes, bent functions and permutations suitable for DES-like cryptosystems," *Des. Codes Cryptogr.*, vol. 15, no. 2, pp. 125-156, 1998.

[4] C. Carlet, C. Ding, and J. Yuan, "Linear codes from perfect nonlinear mappings and their secret schemes," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2089-2102, 2005.

[5] R. S. Coulter and R. W. Matthews, "Planar functions and planes of Lenz-Barlotti class II," *Des. Codes Cryptogr.*, vol. 10, no. 2, pp. 167-184, 1997.

[6] C. Ding and J. Yuan, "A family of skew Hadamard difference sets," *J. Comb. Theory, Ser. A*, vol. 113, no. 7, pp. 1526-1535, 2006.

[7] H. Dobbertin, D. Mills, E. N. Müller, A. Pott, and W. Willems, "APN functions in odd characteristic," *Discret. Math.*, vol. 267, no. 1-3, pp. 95-112, 2003.

[8] G. J. Ness and T. Helleseth, "A new family of ternary almost perfect nonlinear mappings," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2581-2586, 2007.

[9] K. Nyberg, "Differentially uniform mappings for cryptography," in *Advances in cryptology-EUROCRYPT'93 (Lecture Notes in Computer Science)*, vol. 765, T. Helleseth Eds. Berlin, Germany: Springer-Verlag, 1994, pp. 55-64.

[10] Y. Wu, N. Li, and X. Zeng, "Linear codes from perfect nonlinear functions over finite fields," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 3-11, 2020.ao

[11] Y. Xia, F. Bao, S. Chen, C. Li and T. Helleseth, "More Differential Properties about the Ness-Helleseth Nonlinear Mapping," submitted to *IEEE Trans. Inf. Theory*, July 2023.

[12] H. Yan, Y. Xia, C. Li, T. Helleseth, M. Xiong, and J. Luo, "The differential spectrum of the power mapping x^{p^n-3}," *IEEE Trans. Inf. Theory*, vol. 68, no. 8, pp. 5535 - 5547, 2022.

[13] H. Yan, S. Mesnager, and X. Tan. "The complete differential spectrum of a class of power permutations over odd characteristic finite fields," *IEEE Trans. Inf. Theory*, vol. 69, no. 11, pp. 7426 - 7438, 2023.

[14] X. Zeng, L. Hu, Y. Yang, and W Jiang, "On the inequivalence of Ness-Helleseth APN functions," *IACR Cryptol. ePrint Arch*, https://eprint.iacr.org/2007/379.

[15] Z. Zha, "Research on low differential unifomity functions," PhD thesis, Hunan University, 2008.

[16] X. Zhang, "The differential spectra and nonlinearity of some power mappings with low uniformity," Master thesis, South-Central University for Nationalities, 2020.