Workshop on Coding and Cryptography
Perugia, Italy, the 06/20/2024.

Introducing locality in a generalization of AG-codes

Bastien Pacifico
ECo, LIRMM, Montpellier.

Linear codes

A linear code $\mathcal{C} \subset \mathbb{F}_q^n$ is a linear subspace.

We denote by $[n, k, d]$ a code if

- n is its length,
- k is its dimension,
- d is its minimum distance.

Theorem (Singleton Bound)

$$d \leqslant n - k + 1.$$

Such a code can be defined by the image of an injective map $\mathcal{C} : \mathbb{F}_q^k \longrightarrow \mathbb{F}_q^n$.

Locally Recoverable Codes (LRCs)

Definition

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 $c = (c_1, \ldots, c_n) \in \mathcal{C}$ can be recovered using a subset of at most r other symbols. The smallest such r is called the locality of the code.

$$(c_1, \cdots, c_{i-1}, \cancel{c_i}, c_{i+1}, \ldots, c_{n-1}, c_n)$$

Theorem

Let $\mathcal{C}$ be a $q-$ary linear code with parameters $[n, k, d]$ with locality r. The minimum distance d of $\mathcal{C}$ verifies

$$d \leqslant n - k - \left\lceil \frac{k}{r} \right\rceil + 2. \quad \textit{(Singleton Bound for LRCs)}$$

The rate of such a code verifies

$$\frac{k}{n} \leqslant \frac{r}{r+1}$$

Ex : Reed-Solomon codes

A Reed-Solomon code $RS(n,k)$ of length n and dimension k is defined by the image of an application

$$RS(n,k) : \begin{array}{ccc} \mathbb{F}_q[x]_{<k} & \longrightarrow & \mathbb{F}_q^n \\ f & \longmapsto & (f(\alpha_1), \ldots, f(\alpha_n)) \,, \end{array}$$

where $\alpha_1, \ldots, \alpha_n$ are distinct elements of $\mathbb{F}_q$.

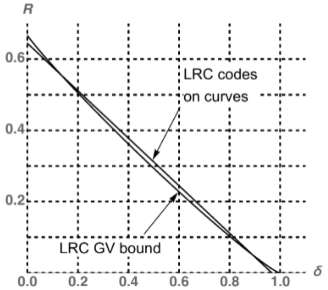
The minimum distance of $RS(n,k)$ veririfes

$$d = n - k + 1.$$

These codes have locality k and (also) reach the Singleton-type bound for LRCs.

Some known constructions

- Tamo-Barg ($n \leqslant q$),
- LRCs from algebraic surfaces ($n \leqslant 4q$),
- Tamo-Barg-Vladuts ($n \mapsto \infty$),
- Concatenated codes ($n \mapsto \infty$).



The bound (46) together with the GV-type bound, $r = 2, q_0 = 32$. [1]

[1] Tamo, Barg and Vladut, *Locally recoverable codes on algebraic curves*, 2015

Some known constructions

- Tamo-Barg ($n \leqslant q$),
- LRCs from algebraic surfaces ($n \leqslant 4q$),
- Tamo-Barg-Vladuts ($n \mapsto \infty$),
- Concatenated codes ($n \mapsto \infty$).

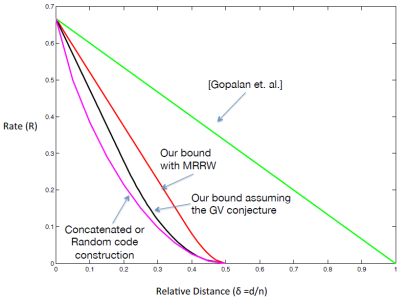


Fig. 2. A depiction of our achievability result (10) through the trade-off between the *rate*, k/n, and *relative distance*, d/n, for binary codes (q=2) for large values of n, with locality $r = 2$. We compare this achievable rate with our upper bounds: assuming respectively MRRW bound, Eq. (8), and the Gilbert-Varshamov (GV) bound, Eq. (9), as the asymptotically optimal rate for binary error-correcting codes as $n \to \infty$. If the GV bound were true rate-distance trade-off, then our achievability scheme is quite good for large distances. [2]

[2] Cadambe and Mazumbar, *Bounds on the Size of Locally Recoverable Codes*, 2015

Concatenated codes

Definition

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

$$\mathcal{C}_{\text{out}}(m) = (c_1, , \dots, c_n),$$

where $m \in \mathbb{F}_{q^{k'}}^k$ and $c_1, \dots, c_n \in \mathbb{F}_{q^{k'}}$. Then the concatenated code $\mathcal{C}_{\text{conc}}$ of $\mathcal{C}_{\text{out}}$ and $\mathcal{C}_{\text{in}}$ is defined by

$$\mathcal{C}_{\text{conc}}(m) = (\mathcal{C}_{\text{in}}(c_1) \mid \cdots \mid \mathcal{C}_{\text{in}}(c_n)).$$

Proposition

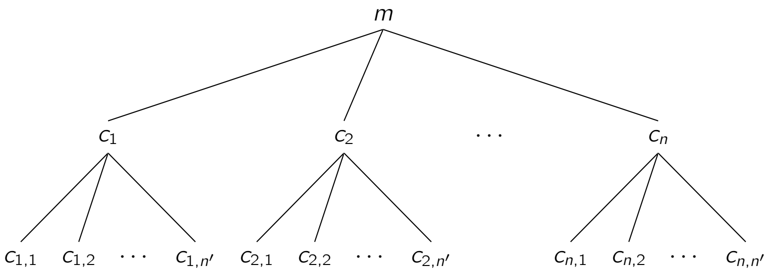
The code $\mathcal{C}_{\text{conc}}$ is a $[nn', kk', \geqslant dd']$ linear code over $\mathbb{F}_q$.

Moreover, the locality of a concatenated code is given by the one of the inner code.

Concatenated LRCs

Let

- $\mathcal{C}_{\text{out}}$ be a $q^{k'} - ary$ linear code of parameters $[n, k, d]$,
- $\mathcal{C}_{\text{in}}$ be a $q - ary$ linear code of parameters $[n', k', d']$ with locality r,
- $\mathcal{C}_{\text{out}}(m) = (c_1, \dots, c_n)$.
- $\mathcal{C}_{\text{in}}(c_i) = (c_{i,1}, \dots, c_{i,n'})$.



The code $\mathcal{C}_{\text{conc}}$ is a $[nn', kk', \geqslant dd']$ $q-$ary linear code with locality r.

A concatenated LRC using RS codes

Let

- $q \geqslant 3$,
- $\mathcal{C}_{\text{out}} := \text{RS}(3,2)$ be a $q^2 - ary$ linear code of parameters $[3,2,2]$,
- $\mathcal{C}_{\text{in}} := \text{RS}(3,2)$ be a $q - ary$ linear code of parameters $[3,2,2]$,
- $a_1, a_2, a_3 \in \mathbb{F}_{q^2}$.

$$f \in \mathbb{F}_{q^2}[x]_{<2}$$

$$f(a_1) \mapsto f_1 \in \mathbb{F}_q[x]_{<2} \qquad f(a_2) \mapsto f_2 \in \mathbb{F}_q[x]_{<2} \qquad f(a_3) \mapsto f_3 \in \mathbb{F}_q[x]_{<2}$$

$$f_1(0) \quad f_1(1) \quad f_1(2) \quad f_2(0) \quad f_2(1) \quad f_2(2) \quad f_3(0) \quad f_3(1) \quad f_3(2)$$

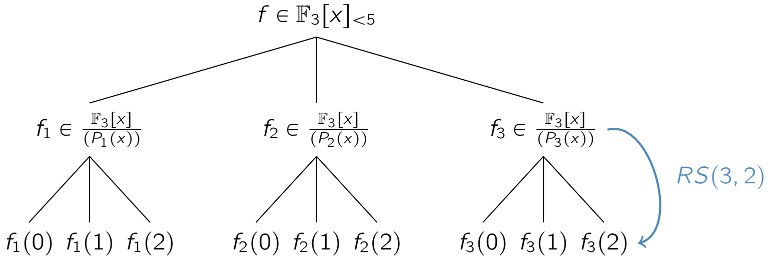
The code $\mathcal{C}_{\text{conc}}$ is a $[9, 4, \geqslant 4]$ $q-$ary linear code with locality 2.

New construction

An optimal example

Let $P_1(x) = x^2 + 2x + 2$, $P_2(x) = x^2 + 1$, and $P_3(x) = x^2 + x + 2$,

and the Reed-Solomon code $RS(3,2)$: $\begin{array}{ccc} \mathbb{F}_3[x]_{<2} & \longrightarrow & \mathbb{F}_3^3 \\ f & \longmapsto & (f(0), f(1), f(2)) \, . \end{array}$



This code is a $[9,5,3]$ linear code with locality 2, reaching the Singleton Bound for LRC .

Actually, some optimal examples

The latter example generalizes to any prime power $q \geqslant 3$.

Proposition

Let $q \geqslant 3$ 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.

Remark : the dimension is not a multiple of the locality.

Are these concatenated codes ?

New example	**Concatenated LRC example**
Polynomials of degree k over $\mathbb{F}_q$	Polynomials of degree k_0 over $\mathbb{F}_{q^r}$
Evaluation modulo s degree r polynomials	Evaluation at s elements of $\mathbb{F}_{q^r}$
Polynomials of degree r over $\mathbb{F}_q$	Polynomials of degree r over $\mathbb{F}_{q^r}$
Evaluation at $r+1$ elements of $\mathbb{F}_q$	Evaluation at $r+1$ elements of $\mathbb{F}_q$
$[s(r+1), k]$ linear code over $\mathbb{F}_q$	$[s(r+1), k_0 r]$ linear code over $\mathbb{F}_q$:

Are these concatenated codes ?

New example	**Concatenated LRC example**
Polynomials of degree k over $\mathbb{F}_q$	Polynomials of degree k_0 over $\mathbb{F}_{q^r}$
Evaluation modulo s degree r polynomials	Evaluation at s elements of $\mathbb{F}_{q^r}$
Polynomials of degree r over $\mathbb{F}_q$	Polynomials of degree r over $\mathbb{F}_{q^r}$
Evaluation at $r+1$ elements of $\mathbb{F}_q$	Evaluation at $r+1$ elements of $\mathbb{F}_q$
$[s(r+1), k]$ linear code over $\mathbb{F}_q$	$[s(r+1), k_0 r]$ linear code over $\mathbb{F}_q$:

What are these codes ?

Are these concatenated codes ?

New example	**Concatenated LRC example**
Polynomials of degree k over $\mathbb{F}_q$	Polynomials of degree k_0 over $\mathbb{F}_{q^r}$
Evaluation modulo s degree r polynomials	Evaluation at s elements of $\mathbb{F}_{q^r}$
Polynomials of degree r over $\mathbb{F}_q$	Polynomials of degree r over $\mathbb{F}_{q^r}$
Evaluation at $r+1$ elements of $\mathbb{F}_q$	Evaluation at $r+1$ elements of $\mathbb{F}_q$
$[s(r+1), k]$ linear code over $\mathbb{F}_q$	$[s(r+1), k_0 r]$ linear code over $\mathbb{F}_q$:

What are these codes ? Generalized AG-Codes !

Algebraic-Geometric (AG) codes

Let $F/\mathbb{F}_q$ be a function field of genus g.
Let $\mathcal{D}$ and G be divisors of F, with $\mathcal{D} = P_1 + \cdots + P_n$, where $P_1, \ldots, P_n$ are distinct rational places (points) of F.
Suppose that $Supp(G) \bigcap \{P_1, \ldots, P_n\} = \varnothing$.
An AG code $\mathcal{C}(\mathcal{D}, G)$ is defined by the image of an application

$$\mathcal{C}(\mathcal{D}, G) : \begin{array}{ccc} \mathcal{L}(G) & \longrightarrow & \mathbb{F}_q^n \\ f & \longmapsto & (f(P_1), \ldots, f(P_n)) \,. \end{array}$$

If $2g - 2 < \deg G < n$, the code $\mathcal{C}(\mathcal{D}, G)$ has dimension

$$k = \deg(G) - g + 1$$

and minimum distance

$$d \geqslant n - \deg(G).$$

Generalized AG codes (GAG)[3]

Let $F/\mathbb{F}_q$ be an algebraic function field defined over $\mathbb{F}_q$ of genus g, and

- $P_1, \ldots, P_s$ 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 $1 \leqslant i \leqslant s$:

- $k_i = \deg(P_i)$ the degree of P_i,
- C_i is a $[n_i, k_i, d_i]_q$ linear code,
- π_i is a fixed $\mathbb{F}_q-$linear isomorphism mapping $\mathbb{F}_{q^{k_i}}$ to C_i.

Consider the application

$$\alpha \ : \ \begin{array}{ccc} \mathcal{L}(G) & \longrightarrow & \mathbb{F}_q^n \\ f & \longmapsto & (\pi_1(f(P_1)), \ldots, \pi_s(f(P_s)))\,. \end{array} .$$

Definition

The image of α is called a generalized algebraic-geometric code, denoted by $C(P_1, \ldots, P_s : G : C_1, \ldots, C_s)$.

[3]Xing, Niederreiter and Lam, *A Generalization of Algebraic-Geometric Codes*, 1999.

Proposition

Observation : if $k_1 = \ldots = k_s =: k$, the code defined above has locality k. More formally,

Proposition

Let $\mathcal{C} = C(P_1, \ldots, P_s : G : C_1, \ldots, C_s)$ be a generalized AG-code as in the previous slide. If there exists $r \in \mathbb{N}$ such that for all $1 \leqslant i \leqslant s$, we have

- $1 < k_i \leqslant r$,
- $n_i > \deg(P_i)$, *and*
- C_i *has locality* k_i,

then $\mathcal{C}$ has locality r.

An optimal example (bis)

Let $\mathbb{F}_3(x)$ be the rational function field. Set $G = 4P_\infty$.
Let $P_1(x) = x^2 + 2x + 2$, $P_2(x) = x^2 + 1$, and $P_3(x) = x^2 + x + 2$,
and the Reed-Solomon code $RS(3,2)$: $\begin{array}{lcl} \mathbb{F}_3[x]_{<2} & \longrightarrow & \mathbb{F}_3^3 \\ f & \longmapsto & (f(0), f(1), f(2)) . \end{array}$

$$f \in \mathcal{L}(4P\infty) \simeq \mathbb{F}_3[x]_{<5}$$

$$F_{P_1} \ni f(P_1) \mapsto f_1 \in \frac{\mathbb{F}_3[x]}{(P_1(x))} \qquad f_2 \in \frac{\mathbb{F}_3[x]}{(P_2(x))} \qquad f_3 \in \frac{\mathbb{F}_3[x]}{(P_3(x))}$$

$RS(3,2)$

$$f_1(0) \ f_1(1) \ f_1(2) \qquad f_2(0) \ f_2(1) \ f_2(2) \qquad f_3(0) \ f_3(1) \ f_3(2)$$

The code $C(P_1, P_2, P_3 : 4P_\infty : RS(3,2), RS(3,2), RS(3,2))$ is a $[9,5,3]$ linear code with locality 2, reaching the Singleton Bound for LRC .

Practical proposition

Proposition

Let $\mathcal{C} = C(P_1, \ldots, P_s : G : C_1, \ldots, C_s)$ *be a generalized AG-code as defined previously. Suppose that*

- $\deg P_1 = \cdots = \deg P_s = r$, *and*
- $\mathcal{C}' = C_1 = \cdots = C_s$ *is a* $[n', r, d']$ *linear code with locality* r.

If $2g - 1 \leqslant \deg(G) < rs$, *then* $\mathcal{C}$ *is a*

$$[sn', deg(G) - g + 1, \geqslant d'\left(s - \left\lfloor \frac{\deg G}{r} \right\rfloor\right)]$$

linear code over $\mathbb{F}_q$ *with locality* r.

More examples : set-up

We (randomly) constructed several codes over $\mathbb{F}_3$ using evaluation at places of degree 2, then encoding the evaluations with $RS(3,2)$ as previously.

We use the following curves.

- The rational function field $\mathbb{F}_3(x)$, of genus 0, that contains 3 places of degree 2. Then one can construct codes of length at most 9.
- The elliptic curve defined by the equation $y^2 = x^3 + x$ of genus 1, that contains 6 places of degree 2. Then one can construct codes of length at most 18.
- The Klein quartic defined by the equation $x^4 + y^4 + 1 = 0$ of genus 3, that contains 12 places of degree 2. Then one can construct codes of length at most 36.

This gives $[3s, k, \geqslant 2\left(s - \left\lfloor \frac{k+g-1}{2} \right\rfloor\right)]$ linear code with locality 2, where s is the number of places of degree 2 used in the construction.

More examples : results

		$\mathbb{F}_3(x)$		$y^2 = x^3 + x$		$x^4 + y^4 + 1$	
n	k	d	defect	d	defect	d	defect
9	3	4	2	4	2	4	2
	4	4	1	4	1	4	1
	5	3	0	3	0	3	0
12	4	-	-	5	3	6	2
	5	-	-	4	2	4	2
	6	-	-	3	2	4	1
15	5	-	-	6	3	6	3
	6	-	-	4	4	5	3
	7	-	-	4	2	4	2
	8	-	-	3	2	4	1
18	6	-	-	6	5	6	5
	7	-	-	6	3	6	3
	8	-	-	4	4	4	4
	9	-	-	4	2	4	2
	10	-	-	2	3	3	2
21	7	-	-	-	-	8	4
	8	-	-	-	-	6	5
	9	-	-	-	-	5	4
	10	-	-	-	-	4	4
	11	-	-	-	-	4	2
	12	-	-	-	-	4	1
24	8	-	-	-	-	8	6
	9	-	-	-	-	7	5
	10	-	-	-	-	6	5
	11	-	-	-	-	6	3
	12	-	-	-	-	4	4
	13	-	-	-	-	4	2
	14	-	-	-	-	3	2
	15	-	-	-	-	3	1
27	9	-	-	-	-	8	7
	10	-	-	-	-	8	6
	11	-	-	-	-	7	5

		$x^4 + y^4 + 1$	
n	k	d	defect
27	12	6	5
	13	6	3
	14	4	4
	15	4	2
	16	3	2
30	10	10	7
	11	8	7
	12	7	7
	13	7	5
	14	6	5
	15	6	3
	16	4	4
	17	4	2
	18	3	2
33	11	10	8
	12	10	7
	13	8	7
	14	8	6
	15	6	6
	16	6	5
	17	5	4
	18	4	4
	19	4	2
36	12	10	10
	13	10	8
	14	8	9
	15	8	7
	16	6	8
	17	6	6
	18	5	6
	19	4	5
	20	4	4

Construction

Proposition

Let

- $F/\mathbb{F}_q$ *be a function field of genus* g *containing* s *places of degree* r*, denoted by* $P_1, \ldots, P_s$,
- $\mathcal{C}_{\text{par}}$ *the* $q-$*ary single parity check code of length* $r+1$ *and dimension* r *and minimum distance* 2,
- G *be a divisor of* F *of degree* $k+g-1$*, where* $g-1 < k < rs-g+1$,

Then, the code $C(P_1, \ldots, P_s : G : \mathcal{C}_{\text{par}}, \ldots, \mathcal{C}_{\text{par}})$ *is a* $[n, k, \geqslant d]$ *linear code over* $\mathbb{F}_q$ *with locality* r*, such that*

$$n = (r+1)s,$$

$$d \geqslant 2\left(s - \left\lfloor \frac{k+g-1}{r} \right\rfloor\right).$$

It follows that the rate of this code verifies

$$\frac{k}{n} \geqslant \frac{r}{r+1} - \frac{r}{2}\delta - \frac{g-1}{n},$$

where $\delta = \frac{d}{n}$.

Drinfeld-Vladut Bound

In this context, we need functions fields with a lot of places (of degree r) relatively to their genus. The best we can expect is given by the following.

Definition (Drinfeld-Vladut Bound of order r)

Let $F/\mathbb{F}_q$ be a function field over $\mathbb{F}_q$ and $B_r(F/\mathbb{F}_q)$ denotes 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,

$$\limsup_{g \longrightarrow +\infty} \frac{B_r(q, g)}{g} \leqslant \frac{1}{r}(q^{\frac{r}{2}} - 1).$$

- $r = 1$: (classical) Drinfeld-Vladut Bound.
- **Example :** Garcia-Stichtenoth recursively defined tower of function fields.[4]

[4] Garcia and Stichtenoth, *A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound*, 1995.

Asymptotic study

Ballet and Rolland[5] studied the descent of the tower of Garcia-Stichtenoth to the field of constant $\mathbb{F}_q$.

The authors also proved that these towers reach the Drinfeld-Vladut bound at order 2.

This allows us to prove the existence of infinite families of linear code with locality 2.

Proposition

Let $q > 3$ be a prime power. Then, Construction 1 provides an infinite family of linear code with locality 2 verifying

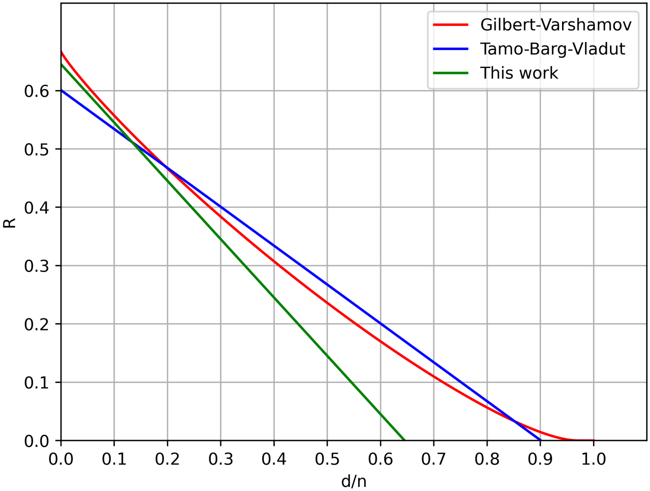
$$\frac{k}{n} \geqslant \frac{2}{3}\left(1 - \frac{q}{q^2 - q - 2}\right) - \delta.$$

[5] Ballet and Rolland, *Families of curves over any finite field attaining the generalized Drinfeld-Vladut bound*, 2011.

Comparison with concatenated codes

GAG-construction	**Concatenated construction**
$F/\mathbb{F}_q$ $P_1, \dots, P_s$ places of degree r	$F/\mathbb{F}_{q^r}$ $P_1, \dots, P_s$ rational places
$g-1 < k < rs - g + 1$ $\deg G = k + g - 1$	$g-1 < k_0 < rs - g + 1$ $\deg G = k_0 + g - 1$
$\left[s(r+1), k, 2\left(s - \left\lfloor \frac{k+g-1}{r} \right\rfloor\right)\right]$	$\left[s(r+1), k = k_0 r, 2\left(s - \frac{k}{r} - g + 1\right)\right]$
$\frac{k}{n} \geqslant \frac{r}{r+1} - \frac{r}{2}\delta - \frac{g-1}{n}$	$\frac{k}{n} \geqslant \frac{r}{r+1} - \frac{r}{2}\delta - \frac{r(g-1)}{n}$
if $q > 3$ and $r = 2$: $\frac{k}{n} \geqslant \frac{2}{3}\left(1 - \frac{q}{q^2-q-2}\right) - \delta$	if $q \geqslant 3$ and $2 \mid r$: $\frac{k}{n} \geqslant \frac{r}{r+1}\left(1 - \frac{r+1}{2}\delta - \frac{1}{q^{\frac{r}{2}}-1}\right)$

Comparison with known-results



Possible further developments

- Improvements (places of other degrees, multiplicities, use other "subcodes",..)
- Hierarchical LRCs ?
- Question : can we use this construction to obtain code of any dimension $k \in \mathbb{N}$?

Possible further developments

- Improvements (places of other degrees, multiplicities, use other "subcodes",..)
- Hierarchical LRCs ?
- Question : can we use this construction to obtain code of any dimension $k \in \mathbb{N}$?

Thanks for your attention!