PIR Codes, Unequal-Data-Demand Codes, and the Griesmer Bound

Henk D.L. Hollmann, Martin Puškin and Ago-Erik Riet

University of Tartu

2024

Notation

In this presetation,

- $\mathbb{F}_q$ denotes the q-element finite field;
- $[n] = \{1, 2, \dots, n\}$;
- given a vector $\mathbf{s}$, we denote by s_j its j-th component.

Error-correction codes

A $[n, k]_q$ error-correction code C is a k-dimensional subspace of $\mathbb{F}_q^n$.

Error-correction codes

A $[n,k]_q$ error-correction code C is a k-dimensional subspace of $\mathbb{F}_q^n$. An *encoder* of the code C is a linear bijection $\epsilon\colon \mathbb{F}_q^k \to C$ which maps the *message word* $m \in \mathbb{F}_q^k$ to a corresponding *code word* $mG \in C$, where $G \in \mathrm{Mat}_{k,n}(\mathbb{F}_q)$. We call G the *generator matrix* for C.

Error-correction codes

A $[n,k]_q$ error-correction code C is a k-dimensional subspace of $\mathbb{F}_q^n$. An *encoder* of the code C is a linear bijection $\epsilon\colon \mathbb{F}_q^k \to C$ which maps the *message word* $m \in \mathbb{F}_q^k$ to a corresponding *code word* $mG \in C$, where $G \in \mathrm{Mat}_{k,n}(\mathbb{F}_q)$. We call G the *generator matrix* for C.

An important parameter of C is its *Hamming distance* $d = \min\{w(u) \mid u \in C\backslash\{\mathbf{0}\}\}$ where $w(u)$ equals the number of non-zero components of u. If d is known, we also call C an $[n,k,d]_q$ error-correction code.

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.

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.

PIR *codes* are a way to make this process more efficient by only storing a part of the data in each server while still allowing for the scheme to work.

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.

PIR *codes* are a way to make this process more efficient by only storing a part of the data in each server while still allowing for the scheme to work.

So PIR codes are used to reduce the storage overhead in the classic PIR scheme.

PIR Codes

Definition (PIR Codes)

Given a (one-to-one) encoder map $\epsilon\colon \mathbb{F}_q^k \to \mathbb{F}_q^n$, 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 $\mathbf{c}_I = (c_{i_1}, c_{i_2}, \ldots, c_{i_s})$ of a codeword $\mathbf{c} = \epsilon(a)$ uniquely determines the j-th data symbol a_j. The encoder map ϵ is a t-PIR code if there exists for every $j \in [k]$ a collection of t *disjoint* recovery sets for the j-th data symbol.

PIR Codes

Definition (PIR Codes)

Given a (one-to-one) encoder map $\epsilon\colon \mathbb{F}_q^k \to \mathbb{F}_q^n$, 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 $\mathbf{c}_I = (c_{i_1}, c_{i_2}, \ldots, c_{i_s})$ of a codeword $\mathbf{c} = \epsilon(a)$ uniquely determines the j-th data symbol a_j. The encoder map ϵ is a t-PIR code if there exists for every $j \in [k]$ a collection of t *disjoint* recovery sets for the j-th data symbol.

We say that a $k \times n$ matrix $\mathbf{G}$ with entries from $\mathbb{F}_q$ is a (linear) t-PIR code if the corresponding encoder $\epsilon\colon \mathbf{a}^\top \to \mathbf{a}^\top\mathbf{G}$ is t-PIR. In that case we say that $\mathbf{G}$ generates a t-PIR code, or that $\mathbf{G}$ is t-PIR.

PIR Codes

Definition (PIR Codes)

Given a (one-to-one) encoder map $\epsilon\colon \mathbb{F}_q^k \to \mathbb{F}_q^n$, a set of positions $I = (i_1, \dots, i_s) \subseteq [n]$ is called a recovery set for the j-th data symbol if the restriction $\mathbf{c}_I = (c_{i_1}, c_{i_2}, \dots, c_{i_s})$ of a codeword $\mathbf{c} = \epsilon(a)$ uniquely determines the j-th data symbol a_j. The encoder map ϵ is a t-PIR code if there exists for every $j \in [k]$ a collection of t *disjoint* recovery sets for the j-th data symbol.

We say that a $k \times n$ matrix $\mathbf{G}$ with entries from $\mathbb{F}_q$ is a (linear) t-PIR code if the corresponding encoder $\epsilon\colon \mathbf{a}^\top \to \mathbf{a}^\top\mathbf{G}$ is t-PIR. In that case we say that $\mathbf{G}$ generates a t-PIR code, or that $\mathbf{G}$ is t-PIR.

Note that being t-PIR is a property of the *encoder* of the code.

PIR Codes

Example

Let $q = 2$, and let C be the binary linear code with (linear) encoder $\epsilon\colon \mathbf{a}^\top \to \mathbf{a}^\top\mathbf{G}$, where

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Then the first data symbol has recovery sets $\{1\}$, $\{2, 3\}$, $\{4\}$ and the second data symbol has recovery sets $\{2\}$ and $\{1, 3\}$. As it's easy to see that the second data symbol cannot have three recovery sets, $\mathbf{G}$ is 2-PIR.

UEP Codes

The higher the distance of the error-correction code, the more protected the message.

UEP Codes

The higher the distance of the error-correction code, the more protected the message.

An *unequal error protection (UEP) code* is an error-correction code where some bits of the message word may be more protected than others and can sometimes be recovered independently.

UEP Codes

The higher the distance of the error-correction code, the more protected the message.

An *unequal error protection (UEP) code* is an error-correction code where some bits of the message word may be more protected than others and can sometimes be recovered independently.

Example

We can define the encoder ϵ to map (a, b) to (a, a, a, b). Now, clearly the first coordinate is more protected than the second.

UEP Codes

In 1978, Dunning and Robins introduced the concept of a *separation vector* as a generalization of distance in order to characterize the error protection capability of UEP codes.

UEP Codes

In 1978, Dunning and Robins introduced the concept of a *separation vector* as a generalization of distance in order to characterize the error protection capability of UEP codes.

Definition

For an encoder $\epsilon\colon \mathbb{F}_q^k \to \mathbb{F}_q^n$, define the *separation vector* $\mathbf{s}(\epsilon) \in \mathbb{Z}_+^k$ by defining for each $j \in [k]$

$$s_j(\epsilon) = \min\{d(\epsilon(\mathbf{a}), \epsilon(\mathbf{a}')) \mid \mathbf{a}, \mathbf{a}' \in \mathbb{F}_q^k, a_j \neq a'_j\}.$$

UEP Codes

In 1978, Dunning and Robins introduced the concept of a *separation vector* as a generalization of distance in order to characterize the error protection capability of UEP codes.

Definition

For an encoder $\epsilon\colon \mathbb{F}_q^k \to \mathbb{F}_q^n$, define the *separation vector* $\mathbf{s}(\epsilon) \in \mathbb{Z}_+^k$ by defining for each $j \in [k]$

$$s_j(\epsilon) = \min\{d(\epsilon(\mathbf{a}), \epsilon(\mathbf{a}')) \mid \mathbf{a}, \mathbf{a}' \in \mathbb{F}_q^k, a_j \neq a_j'\}.$$

Given a separation vector $\mathbf{s}(\epsilon)$, we can decode the i-th data symbol correctly by decoding to the nearest codeword if at most $\lfloor (s_i(\epsilon) - 1)/2 \rfloor$ errors have occurred.

UEP Codes

In 1978, Dunning and Robins introduced the concept of a *separation vector* as a generalization of distance in order to characterize the error protection capability of UEP codes.

Definition

For an encoder $\epsilon\colon \mathbb{F}_q^k \to \mathbb{F}_q^n$, define the *separation vector* $\mathbf{s}(\epsilon) \in \mathbb{Z}_+^k$ by defining for each $j \in [k]$

$$s_j(\epsilon) = \min\{d(\epsilon(\mathbf{a}), \epsilon(\mathbf{a}')) \mid \mathbf{a}, \mathbf{a}' \in \mathbb{F}_q^k, a_j \neq a_j'\}.$$

Given a separation vector $\mathbf{s}(\epsilon)$, we can decode the i-th data symbol correctly by decoding to the nearest codeword if at most $\lfloor (s_i(\epsilon) - 1)/2 \rfloor$ errors have occurred.
We denote by $\mathbf{s}(\mathbf{G})$ the separation vector of a linear code encoded with the generating matrix $\mathbf{G}$.

UEP Codes

Example

We can concatenate an $[n_1, q^{k_1}, d_1]_q$ code and an $[n_2, q^{k_2}, d_2]_q$ code C_2 to form a UEP code with codewords $(\mathbf{c}_1, \mathbf{c}_2)$, $c_i \in C_i$ and a separation vector $\mathbf{s}(\epsilon)$ for which

$$s_i(\epsilon) \geq \begin{cases} d_1 & \text{if } i \text{ is among the first } n_1 \text{ positions,} \\ d_2 & \text{if } i \text{ is among the last } n_2 \text{ positions.} \end{cases}$$

UEP Codes

Example

For a code with the separation vector $(3,2)$, the trivial construction has length 5 as it needs two repetition codes with the encoder $\epsilon(a,b) = aaabb$ $(a, b \in \mathbb{F}_q)$.
Now consider the linear UEP code generated by the matrix $\mathbf{G}$:

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

This has the separation vector $(3,2)$ but its length is only 4.

UDD 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.

UDD Codes

Definition

Let $T = (t_1, \ldots, t_k)$ where $t_1, \ldots, t_k \in \mathbb{Z}$ with $t_1 \geq \ldots \geq t_k \geq 0$. An UDD T-PIR code of length n is an encoder $\epsilon \colon \mathbb{F}_Q^k \to \mathbb{F}_q^n$ where the j-th data symbol has at least t_j mutually disjoint recovery sets for all $j \in [k]$.

UDD Codes

Definition

Let $T = (t_1, \dots, t_k)$ where $t_1, \dots, t_k \in \mathbb{Z}$ with $t_1 \geq \dots \geq t_k \geq 0$. An UDD T-PIR code of length n is an encoder $\epsilon \colon \mathbb{F}_Q^k \to \mathbb{F}_q^n$ where the j-th data symbol has at least t_j mutually disjoint recovery sets for all $j \in [k]$.

We say that a $k \times n$ matrix $\mathbf{G}$ with entries from $\mathbb{F}_q$ is a linear T-PIR code if the corresponding encoder $\epsilon \colon \mathbf{a}^\top \to \mathbf{a}^\top \mathbf{G}$ is T-PIR. In that case we say that $\mathbf{G}$ generates a T-PIR code.

UDD Codes

Definition

Let $T = (t_1, \ldots, t_k)$ where $t_1, \ldots, t_k \in \mathbb{Z}$ with $t_1 \geq \ldots \geq t_k \geq 0$. An UDD T-PIR code of length n is an encoder $\epsilon\colon \mathbb{F}_Q^k \to \mathbb{F}_q^n$ where the j-th data symbol has at least t_j mutually disjoint recovery sets for all $j \in [k]$.

We say that a $k \times n$ matrix $\mathbf{G}$ with entries from $\mathbb{F}_q$ is a linear T-PIR code if the corresponding encoder $\epsilon\colon \mathbf{a}^\top \to \mathbf{a}^\top \mathbf{G}$ is T-PIR. In that case we say that $\mathbf{G}$ generates a T-PIR code.

We can once again get a trivial construction by concatenating t_j-PIR codes.

UDD Codes

Definition

Let $T = (t_1, \ldots, t_k)$ where $t_1, \ldots, t_k \in \mathbb{Z}$ with $t_1 \geq \ldots \geq t_k \geq 0$. An UDD T-PIR code of length n is an encoder $\epsilon \colon \mathbb{F}_Q^k \to \mathbb{F}_q^n$ where the j-th data symbol has at least t_j mutually disjoint recovery sets for all $j \in [k]$.

We say that a $k \times n$ matrix $\mathbf{G}$ with entries from $\mathbb{F}_q$ is a linear T-PIR code if the corresponding encoder $\epsilon \colon \mathbf{a}^\top \to \mathbf{a}^\top \mathbf{G}$ is T-PIR. In that case we say that $\mathbf{G}$ generates a T-PIR code.

We can once again get a trivial construction by concatenating t_j-PIR codes. But the same matrix $\mathbf{G}$ as before provides an example where this is not optimal.

Griesmer Bound

It is well known that the associated code of a t-PIR code has distance $d \geq t$. A stronger result is

Theorem

Let C be a $[n, q^k, d]_q$ code with encoder $\epsilon\colon \mathbb{F}_q^k \to \mathbb{F}_q^n$ and the separation vector $\mathrm{s}(\epsilon)$. If ϵ is an UDD T-PIR code, where $T = (t_1, \dots, t_k)$ with $t_1 \geq \dots \geq t_k \geq 0$, then $s_j(\epsilon) \geq t_j$ for all $j \in [k]$.

Griesmer Bound

It is well known that the associated code of a t-PIR code has distance $d \geq t$. A stronger result is

Theorem

Let C be a $[n, q^k, d]_q$ code with encoder $\epsilon\colon \mathbb{F}_q^k \to \mathbb{F}_q^n$ and the separation vector $\mathbf{s}(\epsilon)$. If ϵ is an UDD T-PIR code, where $T = (t_1, \ldots, t_k)$ with $t_1 \geq \ldots \geq t_k \geq 0$, then $s_j(\epsilon) \geq t_j$ for all $j \in [k]$.

The Griesmer bound for linear UEP codes now directly yields the following for UDD codes.

Theorem (Griesmer Bound for UDD codes)

Suppose that the $k \times n$ matrix $\mathbf{G}$ over $\mathbb{F}_q$ generates a linear UDD T-PIR code, where $T = (t_1, \ldots, t_k)$ with $t_1 \geq \ldots \geq t_k \geq 0$. Then

$$n \geq \sum_{j=1}^{k} \left\lceil \frac{t_j}{q^{j-1}} \right\rceil .$$

An ILP Problem Related to PIR Codes

It would be nice to have an argument that would prove all these Griesmer-type bounds *simultaneously*, in a *uniform* way. We will set up an integer linear programming problem to achieve this.

An ILP Problem Related to PIR Codes

The hyperplanes in $\mathbb{F}_q^k$ and the collection of vectors $\mathcal{P}_k$ of the form $\mathbf{h} = (0, \dots, 0, 1, \dots)$ are in a one-to-one correspondence. The hyperplane corresponding to the vector $\mathbf{h} \in \mathcal{P}_k$ is $\mathbf{h}^\perp := \{\mathbf{a} \in \mathbb{F}_q^k \colon \langle \mathbf{a}, \mathbf{h} \rangle = 0\}$.

An ILP Problem Related to PIR Codes

The hyperplanes in $\mathbb{F}_q^k$ and the collection of vectors $\mathcal{P}_k$ of the form $\mathbf{h} = (0, \dots, 0, 1, \dots)$ are in a one-to-one correspondence. The hyperplane corresponding to the vector $\mathbf{h} \in \mathcal{P}_k$ is $\mathbf{h}^\perp := \{\mathbf{a} \in \mathbb{F}_q^k \colon \langle \mathbf{a}, \mathbf{h} \rangle = 0\}$.

For $\mathbf{h} \in \mathcal{P}_k$, define

$$\nu(\mathbf{h}) = \min\{j \in [k] \colon h_j \neq 0\}.$$

An immediate consequence of the definition is that $\mathbf{h}_{\nu(\mathbf{h})} = 1$.

An ILP Problem Related to PIR Codes

The hyperplanes in $\mathbb{F}_q^k$ and the collection of vectors $\mathcal{P}_k$ of the form $\mathbf{h} = (0, \dots, 0, 1, \dots)$ are in a one-to-one correspondence. The hyperplane corresponding to the vector $\mathbf{h} \in \mathcal{P}_k$ is $\mathbf{h}^\perp := \{\mathbf{a} \in \mathbb{F}_q^k \colon \langle \mathbf{a}, \mathbf{h} \rangle = 0\}$.

For $\mathbf{h} \in \mathcal{P}_k$, define

$$\nu(\mathbf{h}) = \min\{j \in [k] \colon h_j \neq 0\}.$$

An immediate consequence of the definition is that $\mathbf{h}_{\nu(\mathbf{h})} = 1$.

Theorem

Let $\mathbf{G}$ be a $k \times n$ matrix over $\mathbb{F}_q$ that generates a UDD T-PIR code, where $T = (t_1, \dots, t_k)$ with $t_1 \geq \dots \geq t_k \geq 0$. Suppose $\mathbf{G}$ has $n_{\mathbf{i}}$ columns equal to $\mathbf{i}$, $\mathbf{i} \in \mathbb{F}_q^k$. Then for all $\mathbf{h} \in \mathcal{P}_k$, we have

$$\sum_{\langle \mathbf{i}, \mathbf{h} \rangle \neq 0} n_{\mathbf{i}} \geq t_{\nu(\mathbf{h})}.$$

An ILP Problem Related to PIR Codes

So for $T = (t_1, \ldots, t_k) \in \mathbb{Z}^k$ with $t_1 \geq \ldots \geq t_k \geq 0$, define $\nu(T)$ to be the solution the following ILP problem:

$$ILP(T)\colon \begin{cases} n_{\mathbf{i}} \in \mathbb{Z},\ n_{\mathbf{i}} \geq 0 & (\mathbf{i} \in \mathbb{F}_q^k \backslash \{\mathbf{0}\}) \\ \displaystyle\sum_{\{\mathbf{i}\colon \langle \mathbf{i}, \mathbf{h} \rangle \neq 0\}} n_{\mathbf{i}} \geq t_{\nu(\mathbf{h})} & (\mathbf{h} \in \mathcal{P}_k) \\ \text{minimize } n = \displaystyle\sum_{\mathbf{i} \in \mathbb{F}_q^k \backslash \{\mathbf{0}\}} n_{\mathbf{i}}. \end{cases}$$

An ILP Problem Related to PIR Codes

So for $T = (t_1, \ldots, t_k) \in \mathbb{Z}^k$ with $t_1 \geq \ldots \geq t_k \geq 0$, define $\nu(T)$ to be the solution the following ILP problem:

$$ILP(T): \begin{cases} n_{\mathbf{i}} \in \mathbb{Z},\ n_{\mathbf{i}} \geq 0 & (\mathbf{i} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}) \\ \displaystyle\sum_{\{\mathbf{i}:\ \langle \mathbf{i}, \mathbf{h} \rangle \neq 0\}} n_{\mathbf{i}} \geq t_{\nu(\mathbf{h})} & (\mathbf{h} \in \mathcal{P}_k) \\ \text{minimize } n = \displaystyle\sum_{\mathbf{i} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}} n_{\mathbf{i}}. & \end{cases}$$

According to the previous theorem, if $\mathbf{G}$ generates a UDD T-PIR code with $T = (t_1, \ldots, t_k)$ and $t_1 \geq \ldots \geq t_k \geq 0$, then $n \geq n - n_{\mathbf{0}} \geq \nu(T)$. For an optimal solution, we of course take $n_{\mathbf{0}} = 0$.

An ILP Problem Related to PIR Codes

Example

Let $q = 2$ and $k = 2$ and let $T = (t_1, t_2) \in \mathbb{Z}^2$ with $t_1 \geq t_2 \geq 0$. Associate the numbers 1, 2, 3 with the vectors $(1, 0)$, $(0, 1)$, and $(1, 1)$, respectively. The ILP is the problem to minimize $n = n_1 + n_2 + n_3$, where $n_i \geq 0$ is an integer ($i \in [3]$) under the conditions

$$n_1 + n_3 \geq t_1,$$
$$n_2 + n_3 \geq t_2,$$
$$n_1 + n_2 \geq t_1.$$

Here the inequalities correspond to the hyperplanes $(1, 0)^\top$, $(0, 1)^\top$, and $(1, 1)^\top$, respectively. It is not difficult to see that the minimum value for n under these conditions equals $t_1 + \left\lceil \frac{t_2}{2} \right\rceil$.

A Lower Bound for the ILP Problem

Theorem

Let $\nu(T)$ *be the optimal solution to our ILP problem, where* $\mathbf{G}$ *is a* $k \times n$ *matrix over* $\mathbb{F}_q$ *and* $T = (t_1, \ldots, t_k)$ *with* $t_1 \geq \ldots \geq t_k \geq 0$*. Then*

$$\nu(T) \geq \sum_{j=1}^{k} \left\lceil \frac{t_j}{q^{j-1}} \right\rceil .$$

Proof idea. Induction on the dimension k.

The Griesmer Bound for Linear Error Correction Codes from the ILP Problem

The Griesmer bound for linear codes can also be proved by our ILP argument.

The Griesmer Bound for Linear Error Correction Codes from the ILP Problem

The Griesmer bound for linear codes can also be proved by our ILP argument.

Assume that $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_n]$ 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.

The Griesmer Bound for Linear Error Correction Codes from the ILP Problem

The Griesmer bound for linear codes can also be proved by our ILP argument.

Assume that $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_n]$ 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 $\mathbf{G}$ has $n_{\mathbf{i}}$ columns equal to $\mathbf{i}$ ($\mathbf{i} \in \mathbb{F}_q^k$).

The Griesmer Bound for Linear Error Correction Codes from the ILP Problem

The Griesmer bound for linear codes can also be proved by our ILP argument.

Assume that $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_n]$ 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 $\mathbf{G}$ has $n_{\mathbf{i}}$ columns equal to $\mathbf{i}$ ($\mathbf{i} \in \mathbb{F}_q^k$).

Let $\mathbf{h}^\perp$ be a hyperplane ($\mathbf{h} \in \mathbb{F}_q^k \backslash \{\mathbf{0}\}$). Consider $\mathbf{c}^\top := \mathbf{h}^\top \mathbf{G}$. Then $c_j = 0$ iff $\mathbf{h}^\top \mathbf{g}_j = 0$, so $w(\mathbf{c}) = \sum_{\langle \mathbf{h}, \mathbf{i} \rangle \neq 0} n_{\mathbf{i}}$.

The Griesmer Bound for Linear Error Correction Codes from the ILP Problem

The Griesmer bound for linear codes can also be proved by our ILP argument.

Assume that $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_n]$ 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 $\mathbf{G}$ has $n_{\mathbf{i}}$ columns equal to $\mathbf{i}$ ($\mathbf{i} \in \mathbb{F}_q^k$).

Let $\mathbf{h}^{\perp}$ be a hyperplane ($\mathbf{h} \in \mathbb{F}_q^k \backslash \{\mathbf{0}\}$). Consider $\mathbf{c}^{\top} := \mathbf{h}^{\top}\mathbf{G}$. Then $c_j = 0$ iff $\mathbf{h}^{\top}\mathbf{g}_j = 0$, so $w(\mathbf{c}) = \sum_{\langle \mathbf{h}, \mathbf{i} \rangle \neq 0} n_{\mathbf{i}}$.

It follows that $\sum_{\langle \mathbf{h}, \mathbf{i} \rangle \neq 0} n_{\mathbf{i}} \geq d$ for every $\mathbf{h} \in \mathbb{F}_q^k \backslash \{\mathbf{0}\}$.

The Griesmer Bound for linear UEP Codes from the ILP Problem

Suppose that the linear UEP code is generated by a $k \times n$ matrix $\mathbf{G}$ over $\mathbb{F}_q$. Then the separation vector $(s_1, \dots, s_k)$ of the code is given by

$$s_j = s_j(\mathbf{G}) = \min\{w(\mathbf{h}^\top \mathbf{G}) \colon h_j \neq 0\},$$

$j \in [k]$.

The Griesmer Bound for linear UEP Codes from the ILP Problem

Suppose that the linear UEP code is generated by a $k \times n$ matrix $\mathbf{G}$ over $\mathbb{F}_q$. Then the separation vector $(s_1, \ldots, s_k)$ of the code is given by

$$s_j = s_j(\mathbf{G}) = \min\{w(\mathbf{h}^\top \mathbf{G}) \colon h_j \neq 0\},$$

$j \in [k]$.

Suppose that the rows of $\mathbf{G}$ are ordered in such a way that $s_1 \geq \ldots \geq s_k$ and that $\mathbf{G}$ has $n_{\mathbf{i}}$ columns equal to $\mathbf{i}$ ($\mathbf{i} \in \mathbb{F}_q^k$).

The Griesmer Bound for linear UEP Codes from the ILP Problem

Suppose that the linear UEP code is generated by a $k \times n$ matrix $\mathbf{G}$ over $\mathbb{F}_q$. Then the separation vector $(s_1, \dots, s_k)$ of the code is given by

$$s_j = s_j(\mathbf{G}) = \min\{w(\mathbf{h}^\top \mathbf{G}) \colon h_j \neq 0\},$$

$j \in [k]$.

Suppose that the rows of $\mathbf{G}$ are ordered in such a way that $s_1 \geq \dots \geq s_k$ and that $\mathbf{G}$ has $n_\mathbf{i}$ columns equal to $\mathbf{i}$ ($\mathbf{i} \in \mathbb{F}_q^k$).

Then, if $\mathbf{h} \in \mathcal{P}_k$ with $\nu(\mathbf{h}) = j$, we have

$$\sum_{\langle \mathbf{h}, \mathbf{i} \rangle \neq 0} n_\mathbf{i} = |\{l \in [n] \colon \mathbf{h}^\top \mathbf{g}_l \neq 0\}| = w(\mathbf{h}^\top \mathbf{G}) \geq s_j = s_{\nu(\mathbf{h})}.$$

The Griesmer Bound for linear UEP Codes from the ILP Problem

Conversely, let $(n_{\mathbf{i}})_{\mathbf{i}\in\mathbb{F}_q\setminus\{\mathbf{0}\}}$ be a feasible solution to our ILP and $n = \sum n_{\mathbf{i}}$.

The Griesmer Bound for linear UEP Codes from the ILP Problem

Conversely, let $(n_{\mathbf{i}})_{\mathbf{i}\in\mathbb{F}_q\setminus\{\mathbf{0}\}}$ be a feasible solution to our ILP and $n = \sum n_{\mathbf{i}}$.

Consider two code words $\mathbf{c} = \mathbf{a}^\top\mathbf{G}$ and $\mathbf{c}' = \mathbf{b}^\top\mathbf{G}$ in the code C generated by $\mathbf{G}$, and let $\mathbf{h} = \mathbf{a} - \mathbf{b}$. Then $d(\mathbf{c}, \mathbf{c}') = w(\mathbf{h}^\top\mathbf{G}) \geq t_j$ if $h_j \neq 0$. So we can conclude that $s_j(C) \geq t_j$ for all j.

The Griesmer Bound for linear UEP Codes from the ILP Problem

Conversely, let $(n_{\mathbf{i}})_{\mathbf{i}\in\mathbb{F}_q\setminus\{\mathbf{0}\}}$ be a feasible solution to our ILP and $n = \sum n_{\mathbf{i}}$.

Consider two code words $\mathbf{c} = \mathbf{a}^\top\mathbf{G}$ and $\mathbf{c}' = \mathbf{b}^\top\mathbf{G}$ in the code C generated by $\mathbf{G}$, and let $\mathbf{h} = \mathbf{a} - \mathbf{b}$. Then $d(\mathbf{c}, \mathbf{c}') = w(\mathbf{h}^\top\mathbf{G}) \geq t_j$ if $h_j \neq 0$. So we can conclude that $s_j(C) \geq t_j$ for all j.

So the ILP problem is equivalent to finding a linear UEP code with the smallest length for which $\mathbf{s} \geq (t_1, \dots, t_k)$.

Thank you!