Construcciones de Códigos MSRD (Maximum Sum-Rank Distance)

Umberto Martínez-Peñas

Geometría Algebraica: Singularidades e Invariantes,

Universidad de La Laguna, 2021

The Hamming metric

- Let $\mathbb{F}_q$ denote the finite field of size q.
- Let
$$\mathbf{c} = (c_1, c_2, \ldots, c_n) \in \mathbb{F}_q^n.$$
- We define the Hamming weight of $\mathbf{c}$ as
$$\mathrm{wt}_H(\mathbf{c}) = |\{i \in [n] \mid c_i \neq 0\}|.$$
- We define the Hamming distance between $\mathbf{c}, \mathbf{d} \in \mathbb{F}_q^n$ as
$$\mathrm{d}_H(\mathbf{c}, \mathbf{d}) = \mathrm{wt}_H(\mathbf{c} - \mathbf{d}) = |\{i \in [n] \mid c_i \neq d_i\}|.$$
- Singleton bound: For a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ with $k = \log_q |\mathcal{C}|$,
$$\mathrm{d}_H(\mathcal{C}) \leq n - k + 1.$$

Reed-Solomon codes

- Let $1 \leq k \leq n \leq q$, and let $a_1, a_2, \ldots, a_n \in \mathbb{F}_q$ be pair-wise distinct.
- Reed-Solomon codes: (n, k) codes generated by

$$G_{n,k}^{RS}(\mathbf{a}) = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ a_1 & a_2 & a_3 & \ldots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \ldots & a_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1^{k-1} & a_2^{k-1} & a_3^{k-1} & \ldots & a_n^{k-1} \end{pmatrix} \in \mathbb{F}_q^{k \times n}.$$

- As evaluation codes: k-dimensional linear codes

$$\mathcal{C}_{n,k}^{RS}(\mathbf{a}) = \{F(\mathbf{a}) \mid F \in \mathbb{F}_q[x], \deg(F) < k\} \subseteq \mathbb{F}_q^n.$$

- They are MDS (but $q \geq n$):

$$\mathrm{d}_H(\mathcal{C}_{n,k}^{RS}(\mathbf{a})) = n - k + 1.$$

The rank metric

- Let $M : \mathbb{F}_{q^m}^n \longrightarrow \mathbb{F}_q^{m\times n}$ denote the matrix representation map:

$$M(c_1, c_2, \ldots, c_n) = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix}.$$

- We define the rank distance between $\mathbf{c}, \mathbf{d} \in \mathbb{F}_{q^m}^n$ as

$$\mathrm{d}_R(\mathbf{c}, \mathbf{d}) = \mathrm{Rk}(M(\mathbf{c}) - M(\mathbf{d})).$$

- The same Singleton bound applies ($k = \log_{q^m} |\mathcal{C}|$):

$$\mathrm{d}_R(\mathcal{C}) \leq n - k + 1.$$

Gabidulin Codes

- Let $1 \leq k \leq n \leq m$ and let $\beta_1, \beta_2, \ldots, \beta_n \in \mathbb{F}_{q^m}$ be $\mathbb{F}_q$-lin. indep.
- Gabidulin codes: (n, k) codes generated by

$$G_{n,k}^{Gab}(\boldsymbol{\beta}) = \begin{pmatrix} \beta_1 & \beta_2 & \beta_3 & \cdots & \beta_n \\ \beta_1^q & \beta_2^q & \beta_3^q & \cdots & \beta_n^q \\ \beta_1^{q^2} & \beta_2^{q^2} & \beta_3^{q^2} & \cdots & \beta_n^{q^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_1^{q^{k-1}} & \beta_2^{q^{k-1}} & \beta_3^{q^{k-1}} & \cdots & \beta_n^{q^{k-1}} \end{pmatrix} \in \mathbb{F}_{q^m}^{k \times n}.$$

- As evaluation codes: k-dimensional linear codes

$$\mathcal{C}_{n,k}^{Gab}(\boldsymbol{\beta}) = \{F(\boldsymbol{\beta}) \mid F \in \mathcal{L}_q\mathbb{F}_{q^m}[x], \deg_q(F) < k\} \subseteq \mathbb{F}_{q^m}^n.$$

- They are MRD (but $m \geq n$, thus $q^m \geq q^n$):

$$\mathrm{d}_R(\mathcal{C}_{n,k}^{Gab}(\boldsymbol{\beta})) = n - k + 1.$$

The Sum-Rank Metric

- Let

$$\mathbf{c} = (\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \ldots, \mathbf{c}^{(g)}) \in \mathbb{F}_{q^m}^n,$$
$$n = n_1 + n_2 + \cdots + n_g,$$

  where $\mathbf{c}^{(i)} \in \mathbb{F}_{q^m}^{n_i}$, for $i = 1, 2, \ldots, g$.

- We define the sum-rank weight of $\mathbf{c} \in \mathbb{F}_{q^m}^n$ as

$$\mathrm{wt}_{SR}(\mathbf{c}) = \sum_{i=1}^{g} \mathrm{Rk}(M_i(\mathbf{c}^{(i)})).$$

- We define the sum-rank distance between $\mathbf{c}, \mathbf{d} \in \mathbb{F}_{q^m}^n$ as

$$\mathrm{d}_{SR}(\mathbf{c}, \mathbf{d}) = \mathrm{wt}_{SR}(\mathbf{c} - \mathbf{d}).$$

The Sum-Rank Metric

- Let

$$\mathbf{c} = (\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \ldots, \mathbf{c}^{(g)}) \in \mathbb{F}_{q^m}^n,$$
$$n = n_1 + n_2 + \cdots + n_g,$$

  where $\mathbf{c}^{(i)} \in \mathbb{F}_{q^m}^{n_i}$, for $i = 1, 2, \ldots, g$. Note that:

- If $m = n_1 = n_2 = \ldots = n_g = 1$, then $c^{(i)} \in \mathbb{F}_q^1$, $n = g$ and

$$\mathrm{wt}_{SR}(\mathbf{c}) = \sum_{i=1}^{n} \mathrm{Rk}(c^{(i)}) = |\{i \in [n] \mid c^{(i)} \neq 0\}| = \mathrm{wt}_H(\mathbf{c}).$$

- If $g = 1$, then $n = n_1$ and

$$\mathrm{wt}_{SR}(\mathbf{c}) = \sum_{i=1}^{1} \mathrm{Rk}(M_i(\mathbf{c}^{(i)})) = \mathrm{Rk}(M_1(\mathbf{c}^{(1)})) = \mathrm{wt}_R(\mathbf{c}).$$

The Sum-Rank Metric

A few remarks:

- The sum-rank metric in $\mathbb{F}_{q^m}^n$ depends on the base field and length partition:
$$\mathbb{F}_q \quad \text{and} \quad n = n_1 + n_2 + \cdots + n_g.$$
- We could (but we won't) consider different numbers of rows too:
$$\mathbb{F}_q^{m_1 \times n_1} \times \mathbb{F}_q^{m_2 \times n_2} \times \cdots \times \mathbb{F}_q^{m_g \times n_g} \quad \ncong \mathbb{F}_{q^m}^n.$$
- The same Singleton bound applies ($k = \log_{q^m} |\mathcal{C}|$):
$$\mathrm{d}_{SR}(\mathcal{C}) \leq n - k + 1.$$
Codes attaining this bound will be called MSRD codes.
- Any MRD code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ is MSRD for any length partition, but
$$q^m \geq q^n.$$

Field sizes

- Any MRD code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ is MSRD for any length partition, but

$$q^m \geq q^n.$$

- So, what is the problem?

Field sizes

Example: Locally repairable codes (LRCs)

- Started by the people at Microsoft for Windows Azure:


P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Info. Theory*, vol. 58, no. 11, pp. 6925–6934, Nov 2012.

- PMDS codes are the best, information-theoretically, and were started by IBM and Microsoft:


M. Blaum, J. L. Hafner, and S. Hetzler, "Partial-MDS codes and their application to RAID type of architectures," *IEEE Trans. Info. Theory*, vol. 59, no. 7, pp. 4510–4519, July 2013.


P. Gopalan, C. Huang, B. Jenkins, and S. Yekhanin, "Explicit maximally recoverable codes with locality," *IEEE Trans. Info. Theory*, vol. 60, no. 9, pp. 5245–5256, Sept 2014.

Field sizes

Example: Locally repairable codes (LRCs)

- Gabidulin codes (or any MRD code) were known to give optimal LRCs and PMDS codes for arbitrary parameters:

A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath, "Optimal locally repairable and secure codes for distributed storage systems," *IEEE Trans. Info. Theory*, vol. 60, no. 1, pp. 212–236, 2014.

- However, these are not implemented, and considered impractical.
- Tamo and Barg received an IEEE IT-Soc. paper award for a "weaker" family of optimal LRCs (not PMDS), which were tested at Microsoft:

I. Tamo and A. Barg, "A family of optimal locally recoverable codes," *IEEE Trans. Info. Theory*, vol. 60, no. 8, pp. 4661–4676, Aug 2014.

Field sizes

- The big difference is in the field sizes.
- In Distributed Storage (similarly in Network Coding), messages or files are not exactly codewords in the code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$.
- Our k information symbols form quite large vectors

$$\mathbf{x} \in \mathbb{F}_{2^M}^k,$$

where $M = 10^3 \cdot 8$ (MB), $M = 10^6 \cdot 8$ (GB), etcetera.

- Our code lives in $\mathbb{F}_{2^N}^n$, with $N|M$, thus the encoding is

$$\mathbf{c} = \mathbf{x} \cdot G \in \mathbb{F}_{2^M}^n.$$

- If our code is $\mathbb{F}_{2^N}$-linear, then erasure correction requires about

$$\frac{M}{N} \cdot n^2 \cdot N^2$$

operations in $\mathbb{F}_2$ (XORs).

Field sizes

- The big difference is in the field sizes.
- In Distributed Storage (similarly in Network Coding), messages or files are not exactly codewords in the code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$.
- Our k information symbols form quite large vectors

$$\mathbf{x} \in \mathbb{F}_{2^M}^k,$$

where $M = 10^3 \cdot 8$ (MB), $M = 10^6 \cdot 8$ (GB), etcetera.

- Our code lives in $\mathbb{F}_{2^N}^n$, with $N|M$, thus the encoding is

$$\mathbf{c} = \mathbf{x} \cdot G \in \mathbb{F}_{2^M}^n.$$

- If our code is $\mathbb{F}_{2^N}$-linear, then erasure correction requires about

$$M \cdot N \cdot n^2$$

operations in $\mathbb{F}_2$ (XORs).

Field sizes

- MRD-based LRCs $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n = \mathbb{F}_{2^N}^n$ require

$$N \geq n.$$

- Therefore, their computational complexity over $\mathbb{F}_2$ is about

$$M \cdot N \cdot n^2 \approx M \cdot n^3.$$

- Reed-Solomon codes and Tamo-Barg codes $\mathcal{C} \subseteq \mathbb{F}_{2^N}^n$ only require

$$2^N \approx n, \quad \text{that is} \quad N \approx \log_2(n).$$

- Therefore, their computational complexity over $\mathbb{F}_2$ is about

$$M \cdot N \cdot n^2 \approx M \cdot n^2 \cdot \log_2(n).$$

Field sizes

- To simplify things, ideally, the size of field of linearity

$$q^m$$

  of our codes $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ should be polynomial in n.

- Linear field sizes $q^m \approx n$ are preferable.
- Exponential field sizes $q^m \approx q^n$ are not practical.
- MRD codes require exponential field sizes.
- To construct PMDS codes, linear field sizes are not possible:

S. Gopi, V. Guruswami, and S. Yekhanin.
On maximally recoverable local reconstruction codes.
Electr. Colloq. Comp. Complexity (ECCC), 24(183), 2017.

Linearized Reed-Solomon Codes

- MSRD codes yield PMDS codes:

  

  U. Martínez-Peñas and F. R. Kschischang.
  Universal and dynamic locally repairable codes with maximal recoverability via sum-rank codes.
  IEEE Trans. Info. Theory, 65(12):7790–7805, 2019.

- There exist MSRD codes with polynomial field sizes (linearized Reed-Solomon codes):

  

  U. Martínez-Peñas,
  Skew and linearized Reed-Solomon codes and maximum sum rank distance codes over any division ring.
  Journal of Algebra, Vol. 504, pp. 587–612, 2018.

Linearized Reed-Solomon Codes

- Fix $a \in \mathbb{F}_{q^m}$, let $\sigma(a) = a^q$, and define its *i*th norm as

$$N_i(a) = \sigma^{i-1}(a)\sigma^{i-2}(a)\cdots\sigma(a)a = a^{\frac{q^i-1}{q-1}}.$$

- Define the $\mathbb{F}_q$-linear operator $\mathcal{D}_a^i : \mathbb{F}_{q^m} \longrightarrow \mathbb{F}_{q^m}$ by

$$\mathcal{D}_a^i(\beta) = \sigma^i(\beta)N_i(a) = \beta^{q^i} a^{\frac{q^i-1}{q-1}}.$$

- We evaluate a skew polynomial $F = \sum_{i=0}^{k-1} F_i x^i \in \mathbb{F}_{q^m}[x;\sigma]$ at $(a,\beta) \in \mathbb{F}_{q^m}^2$ as

$$F^{\mathcal{D}_a}(\beta) = \sum_{i=0}^{k-1} F_i \mathcal{D}_a^i(\beta) = \sum_{i=0}^{k-1} F_i \beta^{q^i} a^{\frac{q^i-1}{q-1}}.$$

Linearized Reed-Solomon Codes

- Assume $\beta_1, \beta_2, \ldots, \beta_m \in \mathbb{F}_{q^m}$ form a basis over $\mathbb{F}_q$, and
- $N_{\mathbb{F}_{q^m}/\mathbb{F}_q}(a_i a_j^{-1}) \neq 1$ for $1 \leq i < j \leq g$ (pair-wise non-conjugate).
- Define the (n, k) linearized Reed-Solomon code as

$$\mathcal{C}^L_{n,k}(\mathbf{a}, \boldsymbol{\beta}) = \{F^{\mathcal{D}_\mathbf{a}}(\boldsymbol{\beta}) \in \mathbb{F}_{q^m}^n \mid F \in \mathbb{F}_{q^m}[x;\sigma], \deg(F) < k\}.$$

- A generator matrix is $G^L_{n,k}(\mathbf{a}, \boldsymbol{\beta}) =$

$$\left(\begin{array}{cccc|c|cccc}
\beta_1 & \beta_2 & \cdots & \beta_{n_1} & \cdots & \beta_1 & \beta_2 & \cdots & \beta_{n_g} \\
\mathcal{D}_{a_1}(\beta_1) & \mathcal{D}_{a_1}(\beta_2) & \cdots & \mathcal{D}_{a_1}(\beta_{n_1}) & \cdots & \mathcal{D}_{a_g}(\beta_1) & \mathcal{D}_{a_g}(\beta_2) & \cdots & \mathcal{D}_{a_g}(\beta_{n_g}) \\
\mathcal{D}^2_{a_1}(\beta_1) & \mathcal{D}^2_{a_1}(\beta_2) & \cdots & \mathcal{D}^2_{a_1}(\beta_{n_1}) & \cdots & \mathcal{D}^2_{a_g}(\beta_1) & \mathcal{D}^2_{a_g}(\beta_2) & \cdots & \mathcal{D}^2_{a_g}(\beta_{n_g}) \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\mathcal{D}^{k-1}_{a_1}(\beta_1) & \mathcal{D}^{k-1}_{a_1}(\beta_2) & \cdots & \mathcal{D}^{k-1}_{a_1}(\beta_{n_1}) & \cdots & \mathcal{D}^{k-1}_{a_g}(\beta_1) & \mathcal{D}^{k-1}_{a_g}(\beta_2) & \cdots & \mathcal{D}^{k-1}_{a_g}(\beta_{n_g})
\end{array}\right).$$

- They require $q > g$ and $m \geq \max_i n_i$, and they are MSRD:

$$\mathrm{d}_{SR}(\mathcal{C}^L_{n,k}(\mathbf{a}, \boldsymbol{\beta})) = n - k + 1.$$

Linearized Reed-Solomon Codes

A closer look at the generator matrix $G^L_{n,k}(\mathbf{a}, \boldsymbol{\beta})$:

$$\left(\begin{array}{cccc|c|cccc} \beta_1 & \beta_2 & \cdots & \beta_{n_1} & \cdots & \beta_1 & \beta_2 & \cdots & \beta_{n_g} \\ \mathcal{D}_{a_1}(\beta_1) & \mathcal{D}_{a_1}(\beta_2) & \cdots & \mathcal{D}_{a_1}(\beta_{n_1}) & \cdots & \mathcal{D}_{a_g}(\beta_1) & \mathcal{D}_{a_g}(\beta_2) & \cdots & \mathcal{D}_{a_g}(\beta_{n_g}) \\ \mathcal{D}^2_{a_1}(\beta_1) & \mathcal{D}^2_{a_1}(\beta_2) & \cdots & \mathcal{D}^2_{a_1}(\beta_{n_1}) & \cdots & \mathcal{D}^2_{a_g}(\beta_1) & \mathcal{D}^2_{a_g}(\beta_2) & \cdots & \mathcal{D}^2_{a_g}(\beta_{n_g}) \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{D}^{k-1}_{a_1}(\beta_1) & \mathcal{D}^{k-1}_{a_1}(\beta_2) & \cdots & \mathcal{D}^{k-1}_{a_1}(\beta_{n_1}) & \cdots & \mathcal{D}^{k-1}_{a_g}(\beta_1) & \mathcal{D}^{k-1}_{a_g}(\beta_2) & \cdots & \mathcal{D}^{k-1}_{a_g}(\beta_{n_g}) \end{array}\right).$$

We may write $G^L_{n,k}(\mathbf{a}, \boldsymbol{\beta}) = (G_1|G_2|\ldots|G_g)$, where

$$G_i = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_{n_i} \\ \beta_1^q a_i & \beta_2^q a_i & \cdots & \beta_{n_i}^q a_i \\ \beta_1^{q^2} a_i^{q+1} & \beta_2^{q^2} a_i^{q+1} & \cdots & \beta_{n_i}^{q^2} a_i^{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{q^{k-1}} a_i^{q^{k-2}+\cdots+q+1} & \beta_2^{q^{k-1}} a_i^{q^{k-2}+\cdots+q+1} & \cdots & \beta_{n_i}^{q^{k-1}} a_i^{q^{k-2}+\cdots+q+1} \end{pmatrix}.$$

Linearized Reed-Solomon Codes

Note that:

- If $m = n_1 = n_2 = \ldots = n_g = 1$, then

$$\mathcal{D}_a^i(\beta) = \sigma^i(\beta) N_i(a) = \beta a^i.$$

($\Longrightarrow$ Classical and generalized Reed-Solomon codes)

- If $g = 1$ and $a = 1$, then

$$\mathcal{D}_a^i(\beta) = \sigma^i(\beta) N_i(1) = \beta^{q^i}.$$

($\Longrightarrow$ Gabidulin codes)

Bounds for MSRD codes

- Assuming $r = n_1 = n_2 = \ldots = n_g$, we have MSRD codes with field sizes
$$q^m \approx (g+1)^r \approx (g+1)^{(n/g)}.$$
- Do we have any lower bounds on q^m?
- We have upper bounds on g based on q, m and n:

E. Byrne, H. Gluesing-Luerssen, and A. Ravagnani. Fundamental properties of sum-rank metric codes. 2020. Preprint: arXiv:2010.02779.

Bounds for MSRD codes

Assume an MSRD code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ exists, with $r = n_1 = n_2 = \ldots = n_g \leq m$ and $h = \mathrm{d}_{SR}(\mathcal{C}) - 1$. Then

- In general,

$$g \leq \left\lfloor \frac{h-2}{r} \right\rfloor + \left\lfloor (q-1) \cdot \frac{q^m}{q^r - 1} \right\rfloor + 1. \tag{1}$$

- If $h = 2$, then

$$g \leq \left\lfloor (q-1) \cdot \frac{q^m + 1}{q^r - 1} \right\rfloor. \tag{2}$$

- If $m = r$, then

$$g \leq \left\lfloor \frac{h-2}{r} \right\rfloor + q + 1. \tag{3}$$

Bounds for MSRD codes

Assume an MSRD code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ exists, with $r = n_1 = n_2 = \ldots = n_g \leq m$ and $h = \mathrm{d}_{SR}(\mathcal{C}) - 1$. Then

- In general,

$$q^m \geq \frac{q^r - 1}{q - 1} \cdot \left(g - \left\lfloor \frac{h-2}{r} \right\rfloor - 1 \right). \tag{4}$$

- If $h = 2$, then

$$q^m \geq \frac{q^r - 1}{q - 1} \cdot g - 1. \tag{5}$$

- If $m = r$, then

$$q \geq g - \left\lfloor \frac{h-2}{r} \right\rfloor - 1. \tag{6}$$

New MSRD codes

- We are going to present now several new linear MSRD codes with smaller field sizes.
- We are going to consider different evaluation vectors

$$\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r | \ldots | \beta_{(\mu-1)r+1}, \ldots, \beta_{\mu r}) \in \mathbb{F}_{q^m}^{\mu r}.$$

- For these, we want to see when we obtain a linear MSRD code from $G_{n,h}^{L}(\mathbf{a}, \boldsymbol{\beta}) = (G_1 | G_2 | \ldots | G_\ell)$, where

$$G_i = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_{\mu r} \\ \beta_1^q a_i & \beta_2^q a_i & \cdots & \beta_{\mu r}^q a_i \\ \beta_1^{q^2} a_i^{q+1} & \beta_2^{q^2} a_i^{q+1} & \cdots & \beta_{\mu r}^{q^2} a_i^{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{q^{h-1}} a_i^{q^{h-2}+\cdots+q+1} & \beta_2^{q^{h-1}} a_i^{q^{h-2}+\cdots+q+1} & \cdots & \beta_{\mu r}^{q^{h-1}} a_i^{q^{h-2}+\cdots+q+1} \end{pmatrix}.$$

New MSRD codes

- If we look again at $G^L_{n,h}(\mathbf{a}, \boldsymbol{\beta}) = (G_1|G_2|\ldots|G_\ell)$, where

$$G_i = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_{\mu r} \\ \beta_1^q a_i & \beta_2^q a_i & \cdots & \beta_{\mu r}^q a_i \\ \beta_1^{q^2} a_i^{q+1} & \beta_2^{q^2} a_i^{q+1} & \cdots & \beta_{\mu r}^{q^2} a_i^{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^{q^{h-1}} a_i^{q^{h-2}+\cdots+q+1} & \beta_2^{q^{h-1}} a_i^{q^{h-2}+\cdots+q+1} & \cdots & \beta_{\mu r}^{q^{h-1}} a_i^{q^{h-2}+\cdots+q+1} \end{pmatrix},$$

- we see that we are considering the sum-rank metric for the length partition:

$$n = (\ell\mu)r = \underbrace{r + r + \cdots + r}_{\ell\mu \text{ times}}$$

- Thus, we have $g = \ell\mu$ rank blocks, where

$$\ell = q - 1.$$

New MSRD codes

Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be the linear code with parity-check matrix $G_{n,h}^L(\mathbf{a}, \boldsymbol{\beta}) = (G_1|G_2|\ldots|G_\ell)$ as above.

Theorem

Define the $\mathbb{F}_q$-linear subspace

$$\mathcal{H}_i = \left\langle \beta_{(i-1)r+1}, \beta_{(i-1)r+2}, \ldots, \beta_{ir} \right\rangle_{\mathbb{F}_q} \subseteq \mathbb{F}_{q^m},$$

for $i = 1, 2, \ldots, \mu$. Then $\mathcal{C}$ is an MSRD code if, and only if, the following two conditions hold:

- $\dim_{\mathbb{F}_q}(\mathcal{H}_i) = r$, *for* $i = 1, 2, \ldots, \mu$, *and*
- $\dim_{\mathbb{F}_q}\left(\sum_{i\in\Gamma} \mathcal{H}_i\right) = r|\Gamma|$, *for any set* $\Gamma \subseteq [\mu]$ *of size* $|\Gamma| = \min\{h, \mu\}$.

New MSRD codes

- If $t = \min\{h, \mu\}$, we want μ subspaces of $\mathbb{F}_q^m$ of dimension r such that any t of them are in direct sum. ($t = 2 \Longrightarrow$ partial spreads.)
- Recall that the field of linearity is $\mathbb{F}_{q^m}$.
- So we want small q^m relative to

$$g = \ell\mu = (q-1)\mu \quad \text{and} \quad r.$$

New MSRD codes

- Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_r) \in \mathbb{F}_{q^r}^r$ be a basis of $\mathbb{F}_{q^r}$ over $\mathbb{F}_q$.
- Assume that $m = r\rho$ (hence $\mathbb{F}_{q^r} \subseteq \mathbb{F}_{q^m}$).
- Choose $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_\mu) \in \mathbb{F}_{q^m}^\mu$, and define

$$(\beta_1, \beta_2, \ldots, \beta_{\mu r}) = (\alpha_1\gamma_1, \ldots, \alpha_r\gamma_1 | \ldots | \alpha_1\gamma_\mu, \ldots, \alpha_r\gamma_\mu) \in \mathbb{F}_{q^m}^{\mu r}.$$

Theorem

The vector $(\beta_1, \beta_2, \ldots, \beta_{\mu r}) \in \mathbb{F}_{q^m}^{\mu r}$ *satisfies Conditions 1 and 2 in the previous Theorem if, and only if, the vector* $(\gamma_1, \gamma_2, \ldots, \gamma_\mu)$ *is t-wise independent over* $\mathbb{F}_{q^r}$*, for* $t = \min\{h, \mu\}$.

New MSRD codes

- Let $\boldsymbol{\delta} \in \mathbb{F}_{q^m}^{\rho}$ be a basis of $\mathbb{F}_{q^m} = \mathbb{F}_{q^{r\rho}}$ over $\mathbb{F}_{q^r}$.
- Consider the matrix representation map $M_{\boldsymbol{\delta}} : \mathbb{F}_{q^{r\rho}}^{\mu} \longrightarrow \mathbb{F}_{q^r}^{\rho \times \mu}$.
- For the previous $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_\mu) \in \mathbb{F}_{q^m}^{\mu}$, define

$$H_{\boldsymbol{\gamma}} = M_{\boldsymbol{\delta}}(\boldsymbol{\gamma}) \in \mathbb{F}_{q^r}^{\rho \times \mu}.$$

Lemma

The vector $\boldsymbol{\gamma} \in \mathbb{F}_{q^m}^{\mu}$ is t-wise independent over $\mathbb{F}_{q^r}$ if, and only if, $\mathrm{d}_H(\mathcal{C}_{\boldsymbol{\gamma}}) \geq t+1$, for the $\mathbb{F}_{q^r}$-linear code

$$\mathcal{C}_{\boldsymbol{\gamma}} = \left\{ \mathbf{y} \in \mathbb{F}_{q^r}^{\mu} \mid \mathbf{y} H_{\boldsymbol{\gamma}} = \mathbf{0} \right\} \subseteq \mathbb{F}_{q^r}^{\mu}.$$

New MSRD codes

- What does all this mean?
- Choose an $\mathbb{F}_{q^r}$-linear code $\mathcal{C}_\gamma \subseteq \mathbb{F}_{q^r}^{\mu}$ with

$$\mathrm{d}_H(\mathcal{C}_\gamma) \geq t+1, \quad \text{where } t = \min\{h, \mu\},$$

- with small redundancy (co-dimension), which is ρ, since

$$q^m = q^{\rho r}.$$

- The number of rank blocks is

$$g = \ell\mu = (q-1)\mu,$$

and μ is the length of the code.

New MSRD codes

First choice: Trivial codes.

- Choose $\mu = \rho = 1$, thus $\gamma = 1 \in \mathbb{F}_{q^r}^1$, and

$$\mathcal{C}_\gamma = \{0\} \subseteq \mathbb{F}_{q^r}^1.$$

- The obtained MSRD code is a linearized Reed-Solomon code.
- The size of the field of linearity is

$$q^m = q^r = (g+1)^r, \quad \text{that is,} \quad m = r.$$

- Meets *Byrne et al.*'s bounds (asymptotically) for $m = r$ and

$$h = \mathcal{O}(rq).$$

New MSRD codes

Second choice: MDS codes.

- Choose $\mu \leq q^r + 1$ and $\rho = t = \min\{h, \mu\}$.
- Take an MDS code of dimension $\mu - t$,

$$\mathcal{C}_\gamma \subseteq \mathbb{F}_{q^r}^{\mu},$$

  thus $\mathrm{d}_H(\mathcal{C}_\gamma) = t + 1$.
- The obtained MSRD code satisfies that

$$g = \ell\mu = (q-1)(q^r+1).$$

- The size of the field of linearity is

$$q^m = \left(\frac{g}{q-1} - 1\right)^{\min\left\{h, \frac{g}{q-1}\right\}}, \quad \text{that is,} \quad m = r\min\{h, \mu\}.$$

New MSRD codes

Third choice: Hamming codes ($h = 2$).

- Consider $1 \leq \rho < \mu$ and choose a Hamming code

$$\mathcal{C}_\gamma \subseteq \mathbb{F}_{q^r}^{\mu}.$$

- In other words, we take the whole projective space

$$\mathbb{P}_{\mathbb{F}_{q^r}}(\mathbb{F}_{q^{r\rho}}) = \{[\gamma_1], [\gamma_2], \ldots, [\gamma_\mu]\}.$$

- The size of the field of linearity is

$$q^m = q^{r\rho} = \frac{q^r - 1}{q - 1} \cdot g + 1.$$

- Meets *Byrne et al.*'s bounds exactly for $h = 2$ and $m = \rho r$.

New MSRD codes

Fourth choice: BCH codes.

- Set a prescribed distance $\delta = t + 1$ and choose a BCH code

$$\mathcal{C}_\gamma \subseteq \mathbb{F}_{q^r}^{\mu}.$$

- The size of the field of linearity is

$$q^m = q^{r\rho} \leq q^r \cdot \left(\frac{g}{q-1} + 1 \right)^{\left\lceil \frac{q^r-1}{q^r}(h-1) \right\rceil}.$$

New MSRD codes

Fifth choice: Algebraic-Geometry codes.

- Using a Hermitian curve

$$x^{q^{\frac{r}{2}}+1} - y^{q^{\frac{r}{2}}} z - y z^{q^{\frac{r}{2}}} = 0,$$

the size of the field of linearity is

$$q^m = (q^r)^{h+\mathfrak{g}} = \mu^{\frac{2}{3}(h+\mathfrak{g})} = \mu^{\frac{1}{3}(2h+\mu^{2/3}-\mu^{1/3})}.$$

- Using a Suzuki curve

$$x^{2^s}\left(y^{q^r} + y x^{q^r-1}\right) = z^{2^s}\left(z^{q^r} + z x^{q^r-1}\right),$$

the size of the field of linearity is

$$q^m = (q^r)^{h+\mathfrak{g}} = \mu^{\frac{1}{2}(h+\mathfrak{g})} \leq \mu^{\frac{1}{2}\left(h+\mu^{\frac{3}{4}}-\mu^{\frac{1}{4}}\right)}.$$

New MSRD codes

Fifth choice: Algebraic-Geometry codes.

- Using García and Stichtenoth's second tower of curves,

$$x_{i+1}^{q^{\frac{r}{2}}} + x_{i+1} = \frac{x_i^{q^{\frac{r}{2}}}}{x_i^{q^{\frac{r}{2}}-1} + 1},$$

the size of the field of linearity is

$$q^m = (q^r)^{h_i+\mathfrak{g}_i} \leq (q^r)^{h_i+q^{\frac{ir}{2}}} = \left(\frac{\mu_i}{q^{\frac{r}{2}}-1}\right)^{\frac{2}{i}\left(h_i+\frac{\mu_i}{q^{\frac{r}{2}}-1}\right)},$$

for $i \in \mathbb{Z}_+$.

Field size tables

Table for $g = 7$, q even	$r = 2$		$r = 3$		$r = 4$		$r = 5$		$r = 6$	
Code $\mathcal{C}_\gamma$	q^m	q	q^m	q	q^m	q	q^m	q	q^m	q
Trivial $\mathcal{C}_\gamma = \{0\}$ (Lin. RS), $\forall h \geq 1$	$\mathbf{2^6}$	2^3	$\mathbf{2^9}$	2^3	$\mathbf{2^{12}}$	2^3	$\mathbf{2^{15}}$	2^3	$\mathbf{2^{18}}$	2^3
MDS, $h = 2$	2^8		$\mathbf{2^6}$		$\mathbf{2^8}$		$\mathbf{2^{10}}$		$\mathbf{2^{12}}$	
$h = 3$	2^{12}	2^2	$\mathbf{2^9}$	2	$\mathbf{2^{12}}$	2	$\mathbf{2^{15}}$	2	$\mathbf{2^{18}}$	2
$h = 4$	2^{16}		2^{12}		2^{16}		2^{20}		2^{24}	
Hamming, $\rho = 3$, $h = 2$	$\mathbf{2^6}$	2	2^9	2	2^{12}	2	2^{15}	2	2^{18}	2
Pr. BCH, $s = 1, 2$, $h = 2$	$\mathbf{2^6}$		$\mathbf{2^6}$		$\mathbf{2^8}$		$\mathbf{2^{10}}$		$\mathbf{2^{12}}$	
$h = 3$	2^{10}	2	$\mathbf{2^9}$	2	$\mathbf{2^{12}}$	2	$\mathbf{2^{15}}$	2	$\mathbf{2^{18}}$	2
$h = 4$	2^{14}		2^{12}		2^{16}		2^{20}		2^{24}	
Best MRD code possible, $\forall h \geq 1$	2^{14}	2	2^{28}	2	2^{42}	2	2^{56}	2	2^{70}	2

Field size tables

Table for $g = 15$, q even	$r = 2$		$r = 3$		$r = 4$		$r = 5$		$r = 6$	
Code $\mathcal{C}_\gamma$	q^m	q	q^m	q	q^m	q	q^m	q	q^m	q
Trivial $\mathcal{C}_\gamma = \{0\}$ (Lin. RS), $\forall h \geq 1$	$\mathbf{2^8}$	2^4	$\mathbf{2^{12}}$	2^4	$\mathbf{2^{16}}$	2^4	$\mathbf{2^{20}}$	2^4	$\mathbf{2^{24}}$	2^4
MDS, $h = 2$	2^8		2^{12}		$\mathbf{2^8}$		$\mathbf{2^{10}}$		$\mathbf{2^{12}}$	
$h = 3$	2^{12}	2^2	2^{18}	2^2	$\mathbf{2^{12}}$	2	$\mathbf{2^{15}}$	2	$\mathbf{2^{18}}$	2
$h = 4$	2^{16}		2^{24}		$\mathbf{2^{16}}$		$\mathbf{2^{20}}$		$\mathbf{2^{24}}$	
Hamming, $\rho = 3$, $h = 2$	$\mathbf{2^6}$	2	$\mathbf{2^9}$	2	2^{12}	2	2^{15}	2	2^{18}	2
Pr. BCH, $s = 1, 2$, $h = 2$	$\mathbf{2^6}$		$\mathbf{2^9}$		$\mathbf{2^8}$		$\mathbf{2^{10}}$		$\mathbf{2^{12}}$	
$h = 3$	2^{10}	2	2^{15}	2	$\mathbf{2^{12}}$	2	$\mathbf{2^{15}}$	2	$\mathbf{2^{18}}$	2
$h = 4$	2^{14}		2^{21}		$\mathbf{2^{16}}$		$\mathbf{2^{20}}$		$\mathbf{2^{24}}$	
Best MRD code possible, $\forall h \geq 1$	2^{30}	2	2^{45}	2	2^{60}	2	2^{75}	2	2^{90}	2

Field size tables

Table for $g = 31$, q even	$r = 2$		$r = 3$		$r = 4$		$r = 5$		$r = 6$	
Code $\mathcal{C}_\gamma$	q^m	q	q^m	q	q^m	q	q^m	q	q^m	q
Trivial $\mathcal{C}_\gamma = \{0\}$ (Lin. RS), $\forall h \geq 1$	$\mathbf{2^{10}}$	2^5	$\mathbf{2^{15}}$	2^5	$\mathbf{2^{20}}$	2^5	$\mathbf{2^{25}}$	2^5	$\mathbf{2^{30}}$	2^5
MDS, $h = 2$	$\mathbf{2^8}$		2^{12}		2^{16}		$\mathbf{2^{10}}$		$\mathbf{2^{12}}$	
$h = 3$	2^{12}	2^2	2^{18}	2^2	2^{24}	2^2	$\mathbf{2^{15}}$	2	$\mathbf{2^{18}}$	2
$h = 4$	2^{16}		2^{24}		2^{32}		$\mathbf{2^{20}}$		$\mathbf{2^{24}}$	
Hamming, $\rho = 3$, $h = 2$	$\mathbf{2^8}$	2	$\mathbf{2^9}$	2	$\mathbf{2^{12}}$	2	2^{15}	2	2^{18}	2
Pr. BCH, $s = 1, 2, 3$, $h = 2$	$\mathbf{2^8}$		$\mathbf{2^9}$		$\mathbf{2^{12}}$		$\mathbf{2^{10}}$		$\mathbf{2^{12}}$	
$h = 3$	2^{14}	2	$\mathbf{2^{15}}$	2	$\mathbf{2^{20}}$	2	$\mathbf{2^{15}}$	2	$\mathbf{2^{18}}$	2
$h = 4$	2^{20}		2^{21}		2^{28}		$\mathbf{2^{20}}$		$\mathbf{2^{24}}$	
Best MRD code possible, $\forall h \geq 1$	2^{62}	2	2^{93}	2	2^{124}	2	2^{155}	2	2^{186}	2

Back to PMDS codes

Right now, all the PMDS codes, for general parameters, with smallest field sizes, are given by MSRD codes:


U. Martínez-Peñas and F. R. Kschischang.
Universal and dynamic locally repairable codes with maximal recoverability via sum-rank codes.
IEEE Trans. Info. Theory, 65(12):7790–7805, 2019.


H. Cai, Y. Miao, M. Schwartz, and X. Tang.
A construction of maximally recoverable codes with order-optimal field size. arXiv:2011.13606.


U. Martínez-Peñas.
A general family of MSRD codes and PMDS codes with smaller field sizes from extended Moore matrices. arXiv:2011.14109.


S. Gopi and V. Guruswami.
Improved maximally recoverable LRCs using skew polynomials. arXiv:2012.07804.