The Critical Problem for Combinatorial Geometries and Coding Theory

Alberto Ravagnani

Eindhoven University of Technology

Finite Geometries 2022, Irsee

The Critical Problem (Crapo&Rota, 1970)

Setup:

- X be a vector space of dimension $N \geq 3$ over $\mathbb{F}_q$;
- $\mathscr{G}_q(X,1)$ be the set of 1-dimensional subspaces of X;
- $\mathscr{A} \subseteq \mathscr{G}_q(X,1)$ a non-empty subset;
- $1 \leq k \leq N-1$ an integer.

Questions:

- Is there a k-dimensional subspace $\mathscr{C} \leq X$ such that $\mathscr{C} \cap L = \{0\}$ for all $L \in \mathscr{A}$?
- How many are these spaces?

"Answer": It depends on the combinatorics of $\mathscr{A}$, in a precise sense.

The Critical Problem (Crapo&Rota, 1970)

Recall: $\mathscr{A} \subseteq \mathscr{G}_q(X,1)$.

Let $\mathscr{L}$ be the lattice of subspaces of X that are spanned by some elements of $\mathscr{A}$, ordered by inclusion $\leq$.

Proposition (Folklore)

$\mathscr{L}$ is a geometric lattice and its rank function is the $\mathbb{F}_q$-dimension of spaces.

The ith **Whitney number** of $\mathscr{L}$ is

$$w_i(\mathscr{L}) = \sum_{\substack{V \in \mathscr{L} \\ \dim(V)=i}} \mu_{\mathscr{L}}(V).$$

The **characteristic polynomial** of $\mathscr{L}$ is

$$\chi(\mathscr{L},\lambda) = \sum_i w_i(\mathscr{L})\lambda^{\mathrm{rk}(\mathscr{L})-i} \in \mathbb{Z}[\lambda].$$

Theorem (Crapo&Rota, 1970)

The largest k for which there exists a k-subspace of X *avoiding* all the elements of $\mathscr{A}$ is

$$\mathrm{rk}(\mathscr{L}) - \min\{r \mid \chi(\mathscr{L},q^r) \neq 0\}.$$

The value of the minimum is called **critical exponent**.

The Critical Problem (Crapo&Rota, 1970)

Refining the result of Crapo&Rota:

Theorem (R., 2020)

The following are *equivalent*:

- (partial) knowledge of the number of "avoiders"
- (partial) knowledge of the Whitney numbers

More precisely, let $\alpha_k(\mathscr{A}) = \#\{\mathscr{C} \leq X \mid \dim(\mathscr{C}) = k, \mathscr{C} \cap L = \{0\} \text{ for all } L \in \mathscr{A}\}$. Then

$$\alpha_k(\mathscr{A}) = \sum_{i=0}^{k} w_i(\mathscr{L}) \begin{bmatrix} N-i \\ k-i \end{bmatrix}_q \quad \text{for } 0 \leq k \leq N,$$

$$w_i(\mathscr{L}) = \sum_{k=0}^{i} \alpha_k(\mathscr{A}) \begin{bmatrix} N-k \\ i-k \end{bmatrix}_q (-1)^{i-k} q^{\binom{i-k}{2}} \quad \text{for } 0 \leq i \leq N.$$

The Critical Problem and Coding Theory

Having large minimum distance is an "avoiding-type" property:

Remark

Let $X = \mathbb{F}_q^n$ and let $2 \leq d \leq n$.

Let $\mathscr{A}$ be the collection of 1-dimensional subspaces of X generated by a vector of Hamming weight $< d$.

Then the avoiders of $\mathscr{A}$ are the codes $\mathscr{C} \leq \mathbb{F}_q^n$ of minimum Hamming distance $\geq d$.

The lattices that correspond to Hamming-metric codes are called **higher-weight Dowling lattices** ($\sim$ 1970).

Notation

$\mathscr{H}(q, n, r)$ is the lattice of subspaces of $\mathbb{F}_q^n$ that are generated by some vectors of Hamming weight $\leq r$. The ith Whitney number is $w_i(q, n, j)$.

The techniques for computing the Whitney numbers of these lattices have not been discovered yet (some progress made by Dowling, Zaslavsky, Bonin, Kung, Brini, Games, ...) $\rightarrow$ wide open problem, equivalent to counting codes.

Formulas can be nasty and with no obvious "structure":

Theorem (R., 2020)

For all $n \geq 9$ we have

$$
\begin{aligned}
-w_3(2,n,3) = & \sum_{1\leq \ell_1<\ell_2<\ell_3\leq n-2} \left(\prod_{j=1}^{3} \binom{n-\ell_j-9+3j)}{2} \right) + 8\binom{n}{3} \sum_{s=3}^{8} \binom{n-3}{n-s}(-1)^{s-3} \\
& + 106\binom{n}{4} \sum_{s=4}^{8} \binom{n-4}{n-s}(-1)^{s-4} + 820\binom{n}{5} \sum_{s=5}^{8} \binom{n-5}{n-s}(-1)^{s-5} \\
& + 4565\binom{n}{6} \sum_{s=6}^{8} \binom{n-6}{n-s}(-1)^{s-6} \\
& + 19810\binom{n}{8} \sum_{s=7}^{8} \binom{n-7}{n-s}(-1)^{s-7} + 70728\binom{n}{8}.
\end{aligned}
$$

Higher-Weight Dowling Lattices

For some parameters, Bernoulli numbers show up:

Theorem (R., 2020)

For all integers $n \geq d \geq 2$ and any prime power q,

$$w_2(q,n,d) = (q^{n-1}-1)\sum_{j=1}^{d}\binom{n}{j}(q-1)^{j-2} - \sum_{1\leq \ell_1<\ell_2\leq n}\Bigg[q^{n-\ell_1-1}\left(\sum_{j=0}^{d-1}\binom{n-\ell_2}{j}(q-1)^j\right)$$
$$+\sum_{j=d}^{n-\ell_2}\sum_{h=0}^{d-1}\binom{n-\ell_2}{j}\binom{n-\ell_1-1}{h}(q-1)^{j+h}$$
$$+\sum_{s=d}^{n-\ell_2}\sum_{t=0}^{d-2}\binom{n-\ell_2}{s}\binom{n-\ell_1-1-s}{t}(q-1)^{s+t}\sum_{\nu=d-t}^{s}\gamma_q(s,s-d+t+2,\nu)\Bigg],$$

where the $\gamma_a(b,c,\nu)$'s are the *agreement numbers*.

$\gamma_a(b,c,\nu)$ is a polynomial in a (for any b, c and ν) whose coefficients are expressions involving the Bernoulli numbers:

$$\frac{x}{e^x-1} = \sum_{n=0}^{+\infty} B_n \frac{x^n}{n!}.$$

The Hardness of Counting Codes

Whitney numbers have a very good track record of resisting explicit computations (NP-hard, in Sheekey's sense).

And counting codes is as hard as computing Whitney numbers explicitly.

$\rightarrow$ Counting codes is a hard problem.

$\rightarrow$ Look at approximations: estimate codes having a certain property.

The Hardness of Counting Codes

Whitney numbers have a very good track record of resisting explicit computations (NP-hard, in Sheekey's sense).

And counting codes is as hard as computing Whitney numbers explicitly.

$\rightarrow$ Counting codes is a hard problem.

$\rightarrow$ Look at approximations: estimate codes having a certain property.

Proposition (Folklore)

A uniformly random k-dimensional code $\mathscr{C} \leq \mathbb{F}_q^n$ is MDS with probability that approaches 1 as $q \to +\infty$.

In a language that is better aligned with this conference:

Proposition (Folklore)

n uniformly random projective points in $\mathrm{PG}(k-1, q)$ form an arc with probability that goes to 1 as $q \to +\infty$.

Rank-metric codes

Definition

A **rank-metric code** is a non-zero $\mathbb{F}_q$-subspace $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$. Its **minimum distance** is

$$d_{\mathrm{rk}}(\mathscr{C}) = \min\{\mathrm{rk}(X) \mid X \in \mathscr{C},\, X \neq 0\}.$$

Known since 1978 in various contexts:

- 1978: Delsarte (association schemes, bilinear forms)
- 1985: Gabidulin (vectors over field extension)
- 1991: Roth (crisscross error correction)
- 1998: Cooperstein (external flats to determinantal varieties)
- 2008: Silva, Koetter, Kschischang (fixing error amplification in networks)

Studied in connection with a number of topics:

- association schemes
- semifields
- linear sets
- posets and lattices
- q-analogues of matroids
- zeta functions
- q-rook theory
- ...

The Singleton-type bound

We assume $m \geq n$ without loss of generality.

Theorem (Delsarte)

Let $\mathscr{C} \leq \mathbb{F}_q^{n \times m}$ be a rank-metric code with $d_{\mathrm{rk}}(\mathscr{C}) \geq d$. Then

$$\dim(\mathscr{C}) \leq m(n-d+1).$$

This follows from the fact that the projection

$$\pi : \mathscr{C} \to \mathbb{F}_q^{(n-d+1) \times m}$$

onto the last $n-d+1$ rows must be injective.

Definition

$\mathscr{C}$ is **MRD** if $\dim(\mathscr{C}) = m(n-d_{\mathrm{rk}}(\mathscr{C})+1)$.

Unlike MDS codes, MRD codes exist for all parameters (existence question solved by Delsarte himself).

How many MRD codes are out there?

Equivalently, what is the value of

$$\delta_q(n\times m,d) := \frac{\#\{\mathscr{C}\leq \mathbb{F}_q^{n\times m} \mid \dim(\mathscr{C})=m(n-d+1), \mathscr{C} \text{ is MRD of distance } d\}}{\begin{bmatrix} mn \\ m(n-d+1)\end{bmatrix}_q} \ ?$$

This is a hard problem, equivalent to computing the Whitney numbers of certain lattices.

Recall:

The (Generalized) Critical Problem by Crapo&Rota, 1970

Let X be a linear space over $\mathbb{F}_q$ and let $\mathscr{A}$ be a collection of 1-dimensional subspaces of X. How many k-dimensional subspaces of X avoid every element of $\mathscr{A}$?

Specializing:

Remark (MRD codes)

Let $X=\mathbb{F}_q^{n\times m}$ and let $\mathscr{A}$ be the nonzero matrices of rank $\leq d-1$ up to multiples. Then the "avoiders" of $\mathscr{A}$ of dimension $m(n-d+1)$ are precisely the MRD codes of distance d.

$2 \times m$ MRD codes of distance 2

Expanding the theory of spectrum-free matrices:

Theorem (Antrobus, Gluesing-Luerssen, 2018)

We have

$$\lim_{q \to +\infty} \delta_q(2 \times m, 2) = \sum_{i=0}^{m} \frac{(-1)^i}{i!}.$$

This number is close to $1/e$. Explanation? Is this situation "typical"?

$2 \times m$ MRD codes of distance 2

Expanding the theory of spectrum-free matrices:

Theorem (Antrobus, Gluesing-Luerssen, 2018)

We have

$$\lim_{q \to +\infty} \delta_q(2 \times m, 2) = \sum_{i=0}^{m} \frac{(-1)^i}{i!}.$$

This number is close to $1/e$. Explanation? Is this situation "typical"?

- There are $\sim q^m$ matrices of rank 1 up to multiples in $\mathbb{F}_q^{2 \times m}$
- If $\mathscr{A}$ is a uniformly random set of projective points in $\mathbb{F}_q^{2 \times m}$ of size $\sim q^m$, what is the density of m-dimensional avoiders of $\mathscr{A}$ for q large?

Theorem (Gruica, R., Sheekey, Zullo, 2022)

The average density is precisely $1/e$ for $q \to +\infty$.

In other words, the asymptotic density of $2 \times m$ MRD codes of distance 2 is quite typical, given the number of projective points to be avoided.

The asymptotic density of MRD codes

Question

What is the asymptotic behavior of $\delta_q(n \times m, d)$ as $q \to +\infty$?

MRD codes are the rank-analogues of MDS codes, which are *dense* for $q \to +\infty$. But...

The asymptotic density of MRD codes

Question

What is the asymptotic behavior of $\delta_q(n \times m, d)$ as $q \to +\infty$?

MRD codes are the rank-analogues of MDS codes, which are *dense* for $q \to +\infty$. But...

Theorem (Gruica, R., 2020)

We have

$$\delta_q(n \times m, d) \in O\left(q^{-(d-1)(n-d+1)+1}\right) \quad \text{as } q \to +\infty.$$

This results uses the interpretation of MRD codes as "avoiders" (Crapo&Rota, 1970). Tools in the proof: graph theory.

Note: non-density of MRD codes is known from 2018 (Antrobus&Gluesing-Luerssen, Byrne&R.)

The asymptotic density of MRD codes

Question

What is the asymptotic behavior of $\delta_q(n \times m, d)$ as $q \to +\infty$?

MRD codes are the rank-analogues of MDS codes, which are *dense* for $q \to +\infty$. But...

Theorem (Gruica, R., 2020)

We have

$$\delta_q(n \times m, d) \in O\left(q^{-(d-1)(n-d+1)+1}\right) \quad \text{as } q \to +\infty.$$

This results uses the interpretation of MRD codes as "avoiders" (Crapo&Rota, 1970). Tools in the proof: graph theory.

Note: non-density of MRD codes is known from 2018 (Antrobus&Gluesing-Luerssen, Byrne&R.)

Corollary (Antrobus, Gluesing-Luerssen, Gruica, R., 2018/20)

$$\lim_{q \to +\infty} \delta_q(n \times m, d) = \begin{cases} 1 & \text{if } d = 1, \\ \sum_{i=0}^{m} \frac{(-1)^i}{i!} & \text{if } n = d = 2, \\ 0 & \text{otherwise.} \end{cases}$$

What is the "exact" asymptotic density?

This is known only in very few cases (we have seen one already: $2 \times m$, distance 2).

Some cases can be studied using the link between MRD codes and semifields:

Theorem (various authors, various forms)

There is a 1-to-1 correspondence between equivalence classes of $n \times n$ full-rank MRD codes and isotopy classes of semifields.

De la Cruz, Kiermaier, Wassermann, Willems, **Algebraic structures of MRD codes**, Advances in Mathematics of Communications, 2016.

What is the "exact" asymptotic density?

This is known only in very few cases (we have seen one already: $2 \times m$, distance 2).

Some cases can be studied using the link between MRD codes and semifields:

Theorem (various authors, various forms)

There is a 1-to-1 correspondence between equivalence classes of $n \times n$ full-rank MRD codes and isotopy classes of semifields.

De la Cruz, Kiermaier, Wassermann, Willems, **Algebraic structures of MRD codes**, Advances in Mathematics of Communications, 2016.

Using the connection with semifields:

Theorem (Gluesing-Luerssen, 2019)

$$\delta_q(3 \times 3, 3) = \frac{(q-1)(q^3-1)(q^3-q)^3(q^3-q^2)^2(q^3-q^2-q-1)}{3(q^7-1)(q^9-1)(q^9-q)}.$$

For q large, this number is $\sim \frac{1}{3}q^{-3}$.

The number of $n \times n$ full-rank MRD codes

Building on results by Menichetti and Biliotti&Jha&Johnson:

Theorem (Gruica, R., Sheekey, Zullo, 2022)

The number of full-rank MRD codes $\mathscr{C} \leq \mathbb{F}_q^{n \times n}$ is at least

$$\frac{|\mathrm{GL}_n(q)|^2}{n(q^n-1)^2}\left(1+\binom{n-1}{2}\frac{(q^n-1)(q-2)}{q-1}\right).$$

Moreover, the bound is sharp for n prime and q sufficiently large (and for any q if $n=3$).

This recovers the sparseness result for 3×3 full-rank MRD codes.

The connection between spaces of matrices and semifields is based on:

$$(\mathscr{L}_{n,q},+,\circ) \cong (\mathrm{End}_{\mathbb{F}_q}(\mathbb{F}_q^n),+,\circ) \cong (\mathbb{F}_q^{n\times n},+,\cdot)$$

The number of $n \times n$ full-rank MRD codes

Corollary (Gruica, R., Sheekey, Zullo, 2022)

For n prime we have

$$\delta_q^{\mathrm{rk}}(n \times n, n) \sim \frac{(n-1)(n-2)}{2n} q^{-n^3+3n^2-n} \quad \text{as } q \to +\infty.$$

Is this "typical"?

The number of $n \times n$ full-rank MRD codes

Corollary (Gruica, R., Sheekey, Zullo, 2022)

For n prime we have

$$\delta_q^{\mathrm{rk}}(n \times n, n) \sim \frac{(n-1)(n-2)}{2n} q^{-n^3+3n^2-n} \quad \text{as } q \to +\infty.$$

Is this "typical"?

- There are $\sim q^{n^2-2}$ matrices of rank $\leq n-1$ up to multiples in $\mathbb{F}_q^{n \times n}$
- If $\mathscr{A}$ is a uniformly random set of projective points in $\mathbb{F}_q^{n \times n}$ of size $\sim q^{n^2-2}$, what is the density of n-dimensional avoiders of $\mathscr{A}$ for q large?

The number of $n \times n$ full-rank MRD codes

Corollary (Gruica, R., Sheekey, Zullo, 2022)

For n prime we have

$$\delta_q^{\mathrm{rk}}(n \times n, n) \sim \frac{(n-1)(n-2)}{2n} q^{-n^3+3n^2-n} \quad \text{as } q \to +\infty.$$

Is this "typical"?

- There are $\sim q^{n^2-2}$ matrices of rank $\leq n-1$ up to multiples in $\mathbb{F}_q^{n \times n}$
- If $\mathscr{A}$ is a uniformly random set of projective points in $\mathbb{F}_q^{n \times n}$ of size $\sim q^{n^2-2}$, what is the density of n-dimensional avoiders of $\mathscr{A}$ for q large?

Theorem (Gruica, R., Sheekey, Zullo, 2022)

The average density is

$$\sim \frac{1}{e^{q^{n-2}}} \quad \text{for } q \to +\infty.$$

Conclusions

For n prime, full-rank $n \times n$ MRD codes are sparse, but way less sparse than "average".

The Average Critical Problem

The previous results have been obtained by specializing:

Theorem (Gruica, R., Sheekey, Zullo, 2022)

Let $X = \mathbb{F}_q^N$. The average density of the k-dimensional avoiders of a uniformly random set of projective points in $\mathbb{F}_q^N$ of size $\sim q^s$ is

$$\sim e^{-q^{k+s-N}}.$$

The case of $2 \times m$, distance 2, MRD codes corresponds to $s = N - k \quad \rightsquigarrow \quad e^{-1}$.

The Average Critical Problem

The previous results have been obtained by specializing:

Theorem (Gruica, R., Sheekey, Zullo, 2022)

Let $X = \mathbb{F}_q^N$. The average density of the k-dimensional avoiders of a uniformly random set of projective points in $\mathbb{F}_q^N$ of size $\sim q^s$ is

$$\sim e^{-q^{k+s-N}}.$$

The case of $2 \times m$, distance 2, MRD codes corresponds to $s = N - k \quad \rightsquigarrow \quad e^{-1}$.

Question

Which properties of the points to be avoided play a role? How does the "dependency" on these properties look like?

The Average Critical Problem

Fixing both cardinality and dimension of the span:

Theorem (Gruica, R., Sheekey, Zullo, 2022)

Let $X = \mathbb{F}_q^N$. The average density of the k-dimensional avoiders of a uniformly random set of projective points in $\mathbb{F}_q^N$ of size $\sim q^s$ and spanning a space of dimension $2 \leq \rho \leq N$ is

$$\frac{\lambda_q(N,k,\ell,\rho)}{\lambda_q(N,0,\ell,\rho)},$$

where

$$\lambda_q(N,s,\ell,\rho) = \sum_{i=0}^{\rho} (-1)^{\rho-i} q^{\binom{\rho-i}{2}} \begin{bmatrix} N-i \\ \rho-i \end{bmatrix}_q \sum_{t=0}^{s} \binom{\frac{q^i-q^t}{q-1}}{\ell} \begin{bmatrix} s \\ t \end{bmatrix}_q \begin{bmatrix} N-s \\ i-t \end{bmatrix}_q q^{(s-t)(i-t)}.$$

Experimentally: On average, sets of points that span larger spaces have more avoiders than those spanning small spaces. But only on average (counterexamples).

To find out more: arXiv 2201.07193

RANK-METRIC CODES, SEMIFIELDS, AND THE AVERAGE CRITICAL PROBLEM

ANINA GRUICA[1], ALBERTO RAVAGNANI[1], JOHN SHEEKEY[2], AND FERDINANDO ZULLO[3]

ABSTRACT. We investigate two fundamental questions intersecting coding theory and combinatorial geometry, with emphasis on their connections. These are the problem of computing the asymptotic density of MRD codes in the rank metric, and the Critical Problem for combinatorial geometries by Crapo and Rota. Using methods from semifield theory, we derive two lower bounds for the density function of full-rank, square MRD codes. The first bound is sharp when the matrix size is a prime number and the underlying field is sufficiently large, while the second bound applies to the binary field. We then take a new look at the Critical Problem for combinatorial geometries, approaching it from a qualitative, often asymptotic, viewpoint. We illustrate the connection between this very classical problem and that of computing the asymptotic density of MRD codes. Finally, we study the asymptotic density of some special families of codes in the rank metric, including the symmetric, alternating and Hermitian ones. In particular, we show that the optimal codes in these three contexts are sparse.

INTRODUCTION

This paper focuses on two fundamental open problems in coding theory and combinatorial geometry, namely: (1) Computing the asymptotic density of MRD codes in the rank-metric; and (2) Solving new instances of the Critical Problem for combinatorial geometries, proposed by Crapo and Rota. As we will illustrate, the former problem can be regarded as an "asymptotic instance" of the latter.

Thank you!