

The Proportion of (Non-)Linear MRD Codes

Anina Gruica and Alberto Ravagnani

Eindhoven University of Technology
{a.gruica,a.ravagnani}@tue.nl

Abstract. In this talk, we will show that the proportion of MRD codes within the set of codes of the same cardinality is very small when the field size q is large, both in the linear and the non-linear setting. More precisely, we will prove that the asymptotic density of MRD codes is 0 as q tends to infinity. The approach is mainly based on the study of isolated vertices in bipartite graphs. It shows a strong divergence between codes in the Hamming and in the rank metric.

Keywords: rank metric code, MRD code, graph theory, density.

1 Introduction

A *rank metric code* is a set of matrices in $\mathbb{F}_q^{n \times m}$ over a finite field \mathbb{F}_q , in which the difference of any two (distinct) matrices in this set have rank distance bounded from below by a given integer. The largest such number, often denoted by d , is called the *minimum distance* of the code. Rank metric codes were first introduced by Delsarte for combinatorial interest [1], and since then they have gained interest in different areas of mathematics and information theory; see [2–6] for example. The rank metric analogue of the well-known *Singleton-bound* gives the class of *Maximum Rank Distance (MRD)* codes, which are optimal codes and exist for any m, n and $d \leq n$.

In this extended abstract, we analyse the asymptotic densities of MRD codes (both in the non-linear and linear setting). More in detail, we fix a value for the minimum distance d and we compute the asymptotics, as $q \rightarrow +\infty$, of the proportion of MRD codes of that distance, respectively, within the set of codes that share the same cardinality. This question has been studied before in the linear setting, where in [7–9] it was shown that MRD codes are *not* dense within the set of codes having the same dimension. In [10] it was shown that MRD codes are actually always (very) sparse as the field size q tends to infinity except for very few exceptions. We revisit the results of [10] and we also show how to approach the problem in the non-linear setting.

2 Preliminaries

Throughout the paper, q is a prime power, \mathbb{F}_q is the finite field with q elements, and $m \geq n \geq 2$ denote integers. All asymptotic estimates are for $q \rightarrow \infty$. In this paper we work with rank-metric codes; see e.g. [1, 6, 11].

Definition 1. A (*rank metric*) *code* over \mathbb{F}_q is a non-empty subset $\mathcal{C} \subseteq \mathbb{F}_q^{n \times m}$. When $|\mathcal{C}| \geq 2$ its *minimum (rank) distance* is the integer

$$d(\mathcal{C}) = \min\{\text{rk}(M - N) \mid M, N \in \mathcal{C}, M \neq N\}.$$

A code \mathcal{C} is *linear* if it is an \mathbb{F}_q -linear subspace of $\mathbb{F}_q^{n \times m}$. We will write $\mathcal{C} \leq \mathbb{F}_q^{n \times m}$ in this case.

A rank metric code cannot both have large dimension and minimum distance. The trade-off between these quantities is captured by the following result of Delsarte.

Theorem 1 (Singleton-like bound; see [1, Theorem 5.4]). Let $\mathcal{C} \subseteq \mathbb{F}_q^{n \times m}$ be a rank metric code of minimum distance d . We have $|\mathcal{C}| \leq q^{m(n-d^{\text{rk}(\mathcal{C})+1})}$.

A rank metric code is called **MRD** if it meets the bound of Theorem 1 with equality. In contrast with MDS codes, MRD codes exist for all parameter sets and field sizes; see [1].

Definition 2. The *ball* of radius $0 \leq r \leq n$ in $\mathbb{F}_q^{n \times m}$ is the set of matrices $M \in \mathbb{F}_q^{n \times m}$ with $\text{rk}(M) \leq r$. It is well-known that its size is

$$\mathbf{b}_q(n \times m, r) := \sum_{i=0}^r \binom{n}{i}_q \prod_{j=0}^{i-1} (q^m - q^j) \sim q^{r(m+n-r)}. \quad (1)$$

It is natural to ask whether the typical rank metric code of a given cardinality is MRD or not. In this extended abstract, we concentrate on the scenario where the alphabet size q is large. To address this question formally, we consider the problem of estimating the proportion of MRD codes (linear and non-linear) of a certain cardinality within the family of codes having the same cardinality. Recall that

$$\begin{bmatrix} a \\ b \end{bmatrix}_q = \prod_{i=0}^{b-1} \frac{(q^a - q^i)}{(q^b - q^i)}$$

denotes the number of b -dimensional subspaces of an a -dimensional space over \mathbb{F}_q . This number is called the q -**binomial coefficient** of a and b . In order to simplify arguments in the sequel, we introduce the following terminology.

Definition 3. For $1 \leq d \leq n$, let $k = m(n - d + 1)$ and

$$\delta_q(n \times m, d) = \frac{|\{\mathcal{C} \subseteq \mathbb{F}_q^{n \times m} \mid |\mathcal{C}| = q^k, d^{\text{rk}(\mathcal{C})} = d\}|}{\binom{q^{mn}}{q^k}}$$

denote the *density (function)* of (possibly) non-linear MRD codes in $\mathbb{F}_q^{n \times m}$ with minimum distance at least d among all codes of cardinality q^k . Their **asymptotic**

density is $\lim_{q \rightarrow +\infty} \delta_q(n \times m, d)$, when the limit exists. Analogously, we denote the density function of \mathbb{F}_q -linear MRD codes of minimum distance $1 \leq d \leq n$ as

$$\delta_q[n \times m, d] = \frac{|\{\mathcal{C} \leq \mathbb{F}_q^{n \times m} \mid \dim(\mathcal{C}) = k, d^{\text{rk}}(\mathcal{C}) = d\}|}{\begin{bmatrix} mn \\ k \end{bmatrix}_q},$$

and their asymptotic density is $\lim_{q \rightarrow +\infty} \delta_q[n \times m, d]$.

Since we focus on large alphabets, we study the asymptotics of the previous problem for q going to infinity. More formally, we denote by Q the set of prime powers, fix m, n and d , and we want to study how the asymptotic density $\lim_{q \rightarrow +\infty} \delta_q(n \times m, d)$ (and $\lim_{q \rightarrow +\infty} \delta_q[n \times m, d]$) behaves. If the asymptotic density for $q \rightarrow +\infty$ is 0, then we say that MRD codes are **sparse**. If instead the asymptotic density is 1, we say that they are **dense**.

Notation 2. We use the Bachmann-Landau notation (“Little O ” and “ \sim ”) to describe the asymptotic growth of real-valued functions defined on Q ; see e.g. [12]. We omit “ $q \in Q$ ” when writing $q \rightarrow +\infty$ and often omit “as $q \rightarrow +\infty$ ” when writing, for example, “ $f(q) \in o(1)$ ”.

3 Graph Theory Tools

In this section we briefly state some graph theory tools we will need later. The results are taken from [10] and the proofs are omitted.

Definition 4. A (*directed*) **bipartite graph** is a 3-tuple $\mathcal{B} = (\mathcal{V}, \mathcal{W}, \mathcal{E})$, where \mathcal{V} and \mathcal{W} are finite non-empty sets and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{W}$. The elements of $\mathcal{V} \cup \mathcal{W}$ are the **vertices** of the graph. We say that a vertex $W \in \mathcal{W}$ is **isolated** if there is no $X \in \mathcal{V}$ with $(X, W) \in \mathcal{E}$. We say that \mathcal{B} is **left-regular of degree** $\delta \geq 0$ if for all $X \in \mathcal{V}$

$$|\{W \in \mathcal{W} \mid (X, W) \in \mathcal{E}\}| = \delta.$$

In order to give bounds for the number of non-isolated vertices in a bipartite graph, we need the notion of an association.

Definition 5. Let \mathcal{V} be a finite non-empty set and let $r \geq 0$ be an integer. An **association** on \mathcal{V} of **magnitude** r is a function $\alpha : \mathcal{V} \times \mathcal{V} \rightarrow \{0, \dots, r\}$ satisfying the following:

- (i) $\alpha(X, X) = r$ for all $X \in \mathcal{V}$;
- (ii) $\alpha(X, Y) = \alpha(Y, X)$ for all $X, Y \in \mathcal{V}$.

Let $\mathcal{B} = (\mathcal{V}, \mathcal{W}, \mathcal{E})$ be a finite bipartite graph and let α be an association on \mathcal{V} of magnitude r . We say that \mathcal{B} is **α -regular** if for all $(X, Y) \in \mathcal{V} \times \mathcal{V}$ the number of vertices $W \in \mathcal{W}$ with $(X, W) \in \mathcal{E}$ and $(Y, W) \in \mathcal{E}$ only depends on $\alpha(X, Y)$. If this is the case, we denote this number by $\mathcal{W}_\ell(\alpha)$, where $\ell = \alpha(X, Y)$.

Remark 1. Note that an α -regular bipartite graph for an association α is necessarily left-regular of degree $\hat{d} = \mathcal{W}_r(\alpha)$.

The main results stated in this extended abstract will be derived by the following bound.

Lemma 1 (see [10, Lemma 3.5]). *Let $\mathcal{B} = (\mathcal{V}, \mathcal{W}, \mathcal{E})$ be a finite bipartite α -regular graph, where α is an association on \mathcal{V} of magnitude r . Let $\mathcal{F} \subseteq \mathcal{W}$ be the collection of non-isolated vertices of \mathcal{W} . If $\mathcal{W}_r(\alpha) > 0$, then*

$$|\mathcal{F}| \geq \frac{\mathcal{W}_r(\alpha)^2 |\mathcal{V}|^2}{\sum_{\ell=0}^r \mathcal{W}_\ell(\alpha) |\alpha^{-1}(\ell)|}.$$

The previous lemma follows by combining the notion of an association and the Cauchy-Schwarz Inequality. We refer to [10] for the proof.

4 Density of (Possibly) Non-Linear MRD Codes

We show how to apply the results of Section 3 to derive estimates for the number of (non-linear) codes in the rank metric having minimum distance bounded from below.

Notation 3. *In this section, let m, n and d be fixed integers with $2 \leq d \leq n \leq m$ and let $k = m(n - d + 1)$. We work with the bipartite graphs*

$$\mathcal{B}_q = (\mathcal{V}_q, \mathcal{W}_q, \mathcal{E}_q),$$

where

$$\mathcal{V}_q = \{\{M, N\} \subseteq \mathbb{F}_q^{n \times m} \mid M \neq N, d(M, N) \leq d - 1\},$$

\mathcal{W}_q is the collection of codes in $\mathbb{F}_q^{n \times m}$ of cardinality q^k , and $(\{M, N\}, \mathcal{C}) \in \mathcal{E}_q$ if and only if $\{M, N\} \subseteq \mathcal{C}$.

It is easy to see that the isolated vertices in \mathcal{W}_q correspond to the MRD codes in $\mathbb{F}_q^{n \times m}$. Since we fixed d in this section, from now on, let \mathbf{b}_q denote the size of the ball in $\mathbb{F}_q^{n \times m}$ in the rank metric of radius $d - 1$ (see Definition 2). We have

$$|\mathcal{V}_q| = \frac{1}{2} q^{mn} (\mathbf{b}_q - 1), \quad |\mathcal{W}_q| = \binom{q^{mn}}{q^k}.$$

It follows from the definitions that \mathcal{B}_q is a left-regular graph of degree

$$\binom{q^{mn} - 2}{q^k - 2}.$$

We now use Lemma 1 to derive a lower bound for the number of non-MRD codes which gives an upper bound on the number of MRD codes.

Theorem 4. Let \mathcal{F}_q be the collection of codes $\mathcal{C} \subseteq \mathbb{F}_q^{n \times m}$ that have cardinality q^k and minimum distance at most $d - 1$. Define the quantities

$$\begin{aligned}\beta_q(0) &= \frac{1}{2}q^{mn}(\mathbf{b}_q - 1) - 2\mathbf{b}_q + 3, \\ \beta_q(1) &= 2\mathbf{b}_q - 4, \\ \Omega_q &= 1 + \beta_q(1) \frac{q^k - 2}{q^{mn} - 2} + \beta_q(0) \frac{(q^k - 2)(q^k - 3)}{(q^{mn} - 2)(q^{mn} - 3)}.\end{aligned}$$

For all $q \in Q$ we have

$$|\mathcal{F}_q| \geq \frac{q^{mn}(\mathbf{b}_q - 1) \binom{q^{mn} - 2}{q^k - 2}}{2\Omega_q}.$$

In particular,

$$\delta_q(n \times m, d) \leq 1 - \frac{(\mathbf{b}_q - 1)q^k(q^k - 1)}{2\Omega_q(q^{mn} - 1)}.$$

Proof. Let $\alpha : \mathcal{Y}_q \times \mathcal{Y}_q \rightarrow \{0, 1, 2\}$ be defined by

$$\alpha(\{M, N\}, \{K, L\}) := 4 - |\{M, N, K, L\}|$$

for all $M, N, K, L \in \mathbb{F}_q^{n \times m}$. We claim that for all $q \in Q$ we have

$$\begin{aligned}|\alpha^{-1}(2)| &= |\mathcal{Y}_q|, \\ |\alpha^{-1}(1)| &= 2|\mathcal{Y}_q|(\mathbf{b}_q - 2), \\ |\alpha^{-1}(0)| &= |\mathcal{Y}_q|(|\mathcal{Y}_q| - 2\mathbf{b}_q + 3).\end{aligned}$$

Indeed, it is not hard to see that $|\alpha^{-1}(2)| = |\mathcal{Y}_q|$. All the elements of $\alpha^{-1}(1)$ can be constructed by freely choosing $\{M, N\} \in \mathcal{Y}_q$ and then $\{K, L\} \in \mathcal{Y}_q$ with either $K = M$ or $K = N$ and

$$L \in \{X \in \mathbb{F}_q^{n \times m} \mid d(X, K) \leq d - 1\} \setminus \{M, N\}.$$

Therefore

$$|\alpha^{-1}(1)| = 2|\mathcal{Y}_q|(\mathbf{b}_q - 2).$$

To compute $|\alpha^{-1}(0)|$ we simply note that

$$|\mathcal{Y}_q|^2 = |\alpha^{-1}(0)| + |\alpha^{-1}(1)| + |\alpha^{-1}(2)|.$$

Therefore the value of $|\alpha^{-1}(0)|$ follows from the values of $|\alpha^{-1}(1)|$ and $|\alpha^{-1}(2)|$.

One easily checks that α is an association on \mathcal{Y}_q and that the bipartite graph \mathcal{B}_q is regular with respect to α . More precisely, for $(\{M, N\}, \{K, L\}) \in \mathcal{Y}_q \times \mathcal{Y}_q$ if we let $\ell = \alpha(\{M, N\}, \{K, L\})$ then

$$\begin{aligned}\mathcal{W}_{q,\ell}(\alpha) &:= |\{W \in \mathcal{Y}_q \mid \{M, N, K, L\} \subseteq W\}| \\ &= \binom{q^{mn} - 4 + \ell}{q^k - 4 + \ell}.\end{aligned}\tag{2}$$

We can now apply Lemma 1 obtaining that $|\mathcal{F}_q|$ is lower bounded by

$$\frac{\mathcal{W}_{q,2}(\alpha)^2 |\mathcal{V}_q|^2}{|\alpha^{-1}(2)|\mathcal{W}_{q,2}(\alpha) + |\alpha^{-1}(1)|\mathcal{W}_{q,1}(\alpha) + |\alpha^{-1}(0)|\mathcal{W}_{q,0}(\alpha)}.$$

Finally, combining the identity

$$\binom{m}{\ell} = \frac{m}{\ell} \binom{m-1}{\ell-1} \quad (3)$$

with the formulas for $|\alpha^{-1}(2)|$, $|\alpha^{-1}(1)|$ and $|\alpha^{-1}(0)|$ and Equation (2), easy computations yield the desired result.

In order to study the asymptotics of the previous bound on the density as the alphabet size q tends to infinity we will need the following estimate:

$$\mathbf{b}_q \sim q^{(d-1)(m+n-d+1)} \quad \text{as } q \rightarrow +\infty. \quad (4)$$

With this we get one of the main results presented in this extended abstract.

Theorem 5. *We have $\lim_{q \rightarrow +\infty} \delta_q(n \times m, d) = 0$. In particular, (non-linear) MRD codes are sparse for all parameter sets.*

Proof. It is not hard to see that

$$\Omega_q \sim \frac{1}{2} q^{-mn+(d-1)(m+n-d+1)+2k} \quad \text{as } q \rightarrow +\infty.$$

Therefore we have

$$\frac{(\mathbf{b}_q - 1)q^k(q^k - 1)}{2\Omega_q(q^{mn} - 1)} \sim 1 \quad \text{as } q \rightarrow +\infty.$$

This proves the statement.

5 Density of Linear MRD Codes

The problem of deciding whether MRD codes are dense or not has been studied before and we start this section with revisiting the approaches been developed so far, one of which is the subject of this extended abstract. We start by briefly recalling the results obtained with the other three approaches. Note that we only focus on the case where $q \rightarrow +\infty$ here, but in [7, 8, 10] the asymptotic behavior of the density function for $m \rightarrow +\infty$ was investigated as well.

In [8], a combinatorial approach to study asymptotic enumeration problems in coding theory was developed, based on the notion of a *partition-balanced* family of codes. Applying the machinery to the problem of estimating the number of linear MRD codes, one obtains the following.

Theorem 6 (see [8, Corollary 6.2]). *If $d \geq 2$, then*

$$\limsup_{q \rightarrow +\infty} \delta_q(n \times m, d) \leq 1/2.$$

Another (sharper) upper bound on the number of MRD codes is obtained in [7] using the theory of spectrum-free matrices. It reads as follows.

Theorem 7 (see [7, Theorem VII.6]). *We have*

$$\limsup_{q \rightarrow +\infty} \delta_q(n \times m, d) \leq \left(\sum_{i=0}^m \frac{(-1)^i}{i!} \right)^{(d-1)(n-d+1)}.$$

In [7], it is also shown that the bound of Theorem 7 is sharp if $d = n = 2$ and for all values of $m \geq 2$.

Finally, in [9] the exact number of MRD codes with the parameters $m = n = d = 3$ is computed, showing that these codes are sparse. The approach of [9] is based on the connection between full-rank MRD codes and semifields.

In this section we establish the analogue of Theorem 5 for \mathbb{F}_q -linear MRD codes. The results are taken from [10] and the proofs are omitted.

Notation 8. *We fix integers m , n and d with $1 \leq d \leq n \leq m$ and we let $k = m(n - d + 1)$. We consider the bipartite graphs*

$$\tilde{\mathcal{B}}_q = (\tilde{\mathcal{V}}_q, \tilde{\mathcal{W}}_q, \tilde{\mathcal{E}}_q),$$

where $\tilde{\mathcal{V}}_q$ is the set of matrices in $\mathbb{F}_q^{n \times m}$ of rank smaller or equal to $d - 1$ (up to multiples), $\tilde{\mathcal{W}}_q$ is the collection of \mathbb{F}_q -linear codes in $\mathbb{F}_q^{n \times m}$ with dimension k , and $(M, \mathcal{C}) \in \tilde{\mathcal{E}}_q$ if and only if $M \in \mathcal{C}$.

We again let \mathbf{b}_q denote the size of the ball in $\mathbb{F}_q^{n \times m}$ of radius $d - 1$. Note that we have

$$|\tilde{\mathcal{V}}_q| = \frac{\mathbf{b}_q - 1}{q - 1}, \quad |\tilde{\mathcal{W}}_q| = \begin{bmatrix} mn \\ k \end{bmatrix}_q.$$

It is easy to check that \tilde{B} is α -regular with respect to the association $\alpha : \tilde{\mathcal{V}} \times \tilde{\mathcal{V}} \rightarrow \{0, 1\}$ where $\alpha(V, V') = \dim(V \cap V')$ for $V, V' \in \tilde{\mathcal{V}}_q$. Simple computations show that by applying Lemma 1 we get the following result.

Theorem 9. *Let $\tilde{\mathcal{F}}_q$ be the collection of rank metric codes $\mathcal{C} \leq \mathbb{F}_q^{n \times m}$ of dimension k and minimum distance at most $d - 1$. We have*

$$|\tilde{\mathcal{F}}_q| \geq \frac{\left(\frac{\mathbf{b}_q - 1}{q - 1} \right) \begin{bmatrix} mn - 1 \\ k - 1 \end{bmatrix}_q^2}{\begin{bmatrix} mn - 1 \\ k - 1 \end{bmatrix}_q + \left(\left(\frac{\mathbf{b}_q - 1}{q - 1} \right) - 1 \right) \begin{bmatrix} mn - 2 \\ k - 2 \end{bmatrix}_q}.$$

In particular,

$$\delta_q[n \times m, d] \leq 1 - \frac{\left(\frac{\mathbf{b}_q - 1}{q - 1}\right) \left[\begin{matrix} mn - 1 \\ k - 1 \end{matrix} \right]_q^2}{\left[\begin{matrix} mn \\ k \end{matrix} \right]_q \left(\left[\begin{matrix} mn - 1 \\ k - 1 \end{matrix} \right]_q + \left(\left(\frac{\mathbf{b}_q - 1}{q - 1}\right) - 1 \right) \left[\begin{matrix} mn - 2 \\ k - 2 \end{matrix} \right]_q \right)}.$$

In order to compute the asymptotic density of linear MRD codes, we use the well-known estimate for q -ary binomial coefficient, which says that for non-negative integers $a \geq b$ we have

$$\left[\begin{matrix} a \\ b \end{matrix} \right]_q \sim q^{b(a-b)} \text{ as } q \rightarrow +\infty.$$

The following result shows the sparseness of linear MRD codes. We omit the proof but it can be found in [10].

Theorem 10. *We have*

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

In particular, linear MRD codes are sparse whenever $n \geq 3$ and $d \geq 2$.

6 Discussion and Future Work

We described the behavior of the density function of possibly non-linear and linear MRD codes. Both families of codes are very sparse within the set of codes of the same cardinality. This means, if one chooses uniformly at random a rank metric code of a certain cardinality, this code will most probably not be MRD. This is in stark contrast with the behavior of linear MDS codes in the Hamming metric for example but it coincides with the asymptotic behavior of possibly non-linear MDS codes.

A natural question inspired by the above results is that of understanding more about which structural invariants of a metric space determine whether or not a uniformly random subset has good distance properties. It seems that graph theory is a valid tool for approaching this problem.

References

1. P. Delsarte, “Bilinear forms over a finite field, with applications to coding theory,” *Journal of Combinatorial Theory, Series A*, vol. 25, no. 3, pp. 226–241, 1978.
2. R. Kötter and F. R. Kschischang, “Coding for errors and erasures in random network coding,” *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3579–3591, 2008.
3. J. Sheekey, “New semifields and new MRD codes from skew polynomial rings,” *Journal of the London Mathematical Society*, vol. 101, no. 1, pp. 432–456, 2020.

4. E. Gorla, R. Jurrius, H. H. López, and A. Ravagnani, “Rank-metric codes and q -polymatroids,” *Journal of Algebraic Combinatorics*, vol. 52, no. 1, pp. 1–19, 2020.
5. E. M. Gabidulin, “Theory of codes with maximum rank distance,” *Problemy Peredachi Informatsii*, vol. 21, no. 1, pp. 3–16, 1985.
6. R. M. Roth, “Maximum-rank array codes and their application to crisscross error correction,” *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 328–336, 1991.
7. J. Antrobus and H. Gluesing-Luerssen, “Maximal Ferrers diagram codes: Constructions and genericity considerations,” *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6204–6223, 2019.
8. E. Byrne and A. Ravagnani, “Partition-balanced families of codes and asymptotic enumeration in coding theory,” *Journal of Combinatorial Theory, Series A*, vol. 171, 2020.
9. H. Gluesing-Luerssen, “On the sparseness of certain linear MRD codes,” *Linear Algebra and its Applications*, vol. 596, pp. 145–168, 2020.
10. A. Gruica and A. Ravagnani, “Common complements of linear subspaces and the sparseness of MRD codes,” *arXiv preprint 2011.02993*, 2020.
11. E. M. Gabidulin, “Theory of codes with maximum rank distance,” *Problemy peredachi informatsii*, vol. 21, no. 1, pp. 3–16, 1985.
12. N. G. De Bruijn, *Asymptotic Methods in Analysis*. Courier Corporation, 1981, vol. 4.