

Optimum Linear Codes with Support Constraints over Small Fields

Hikmet Yildiz, Babak Hassibi
 California Institute of Technology
 Email: {hyildiz, hassibi}@caltech.edu

Abstract—We consider the problem of designing optimal linear codes (in terms of having the largest minimum distance) subject to a support constraint on the generator matrix. We show that the largest minimum distance can be achieved by a subcode of a Reed-Solomon code of small field size. As a by-product of this result, we settle the GM-MDS conjecture of Dau *et al.* in the affirmative.

I. INTRODUCTION

The problem of designing a linear code with the largest possible minimum distance, subject to support constraints on the generator matrix, has recently found several applications. These include multiple access networks [3], [5] as well as weakly secure data exchange [4], [8]. A simple upper bound on the maximum minimum distance can be obtained from a sequence of Singleton bounds (see eq. (3) below) and can further be achieved by randomly choosing the nonzero elements of the generator matrix from a field of a large enough size.

A natural question to ask is whether the above maximum minimum distance can be achieved with a field of small size, and in particular with a structured, possible algebraic, construction. This question is equivalent to a recently proposed conjecture by Dau *et al.* [2], which is commonly referred to as the GM-MDS conjecture.

In the past couple of years, progress has been reported on this conjecture. Heidarzadeh *et al.* [9] have proved it for dimensions $k \leq 5$. Halbawi *et al.* [5] have proved the statement for $m \leq 3$ if there are m distinct support sets on the rows of the generator matrix. In the authors' previous work [1], the statement has been proved for $m \leq 6$. Halbawi *et al.* [6], [7] and Song *et al.* [10], have studied the problem when the generator matrix is sparsest and balanced and established the conjecture in this special case. Yan *et al.* [8] give some partial results.

In this paper we show that the largest minimum distance can be achieved by a subcode of a Reed-Solomon code of small field size, in fact as low as $2n - d$, where n is the code length, and d is the maximum minimum distance dictated by the support constraints. As a by-product of this result, we settle the GM-MDS conjecture in the affirmative.

The remainder of the paper is organized as follows. In Section II, we characterize the generator matrices of subcodes of Reed-Solomon codes. Section III defines the problem (of maximizing d_{\min} subject to support constraints) and shows that it can be reduced to the GM-MDS conjecture. Section

IV proposes a more general statement of the problem, that is not directly related to the coding problem, but that more readily lends itself to an induction argument. This is the result we prove which, by fiat, solves the problem of maximizing d_{\min} and the GM-MDS conjecture. The proof is detailed in the Appendix.

A. Notation

Matrices are shown by bold capital letters and vectors are shown by bold lower case letters. For $n \geq 0$, we denote by $[n]$ the set $\{1, 2, \dots, n\}$ by admitting $[0] = \emptyset$. For $n \geq 1$, we write $[\theta_i]_{i=1}^m$ to represent the ordered list of objects $\theta_1, \dots, \theta_n$. For a finite nonempty $S \subset \mathbb{Z}$, $[\theta_i]_{i \in S}$ is the ordered list of θ_i 's for $i \in S$ in the ascending order of their indices.

$\mathbb{F}[x]$ represents the polynomial ring over the field \mathbb{F} , i.e. the set of polynomials with coefficients in \mathbb{F} . $\mathbb{F}(x)$ represents the field of rational functions in x over the field \mathbb{F} , i.e. the set of functions that can be written as a ratio of two polynomials in $\mathbb{F}[x]$ such that the denominator is not the zero polynomial.

$[n, k]_q$ and $[n, k, d]_q$ represent a linear code over \mathbb{F}_q with length n , dimension k , and minimum distance d .

II. SUBCODES OF REED-SOLOMON CODES

An $[n, \ell, n - \ell + 1]_q$ Reed-Solomon code can be generated by a Vandermonde matrix

$$\mathbf{V} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{\ell-1} & \alpha_2^{\ell-1} & \cdots & \alpha_n^{\ell-1} \end{pmatrix} \in \mathbb{F}_q^{\ell \times n} \quad (1)$$

for distinct $\alpha_1, \dots, \alpha_n \in \mathbb{F}_q$. Reed-Solomon codes have efficient decoders that can correct up to $\lfloor \frac{n-\ell+1}{2} \rfloor$ errors.

For $n \geq \ell \geq k$, $[n, k]_q$ subcodes of $[n, \ell]_q$ Reed-Solomon codes have generator matrices of the following form:

$$\mathbf{G} = \mathbf{T} \cdot \mathbf{V} \quad (2)$$

where $\mathbf{T} \in \mathbb{F}_q^{k \times \ell}$ is full rank and \mathbf{V} is given in (1).

The minimum distance d is equal to the minimum weight of \mathbf{mG} over all nonzero row vectors $\mathbf{m} \in \mathbb{F}_q^k$. Since \mathbf{V} is a Vandermonde matrix, if we treat the entries of \mathbf{mT} as coefficients of a nonzero polynomial $p \in \mathbb{F}_q[x]$, then, the entries of \mathbf{mG} will be $p(\alpha_1), \dots, p(\alpha_n)$. As $\deg p \leq \ell - 1$, the number of nonzero entries in \mathbf{mG} is at least $n - \ell + 1$. Therefore, the minimum distance is bounded by $d \geq n - \ell + 1$.

Theorem 3 gives necessary and sufficient conditions on the parameters $[(S_i, r_i)]_{i=1}^m$ for $\det \mathbf{M}$ to be nonzero.

Theorem 3. *Let $k \geq m \geq 1$, $n \geq 0$, $[(S_i, r_i)]_{i=1}^m \in \mathcal{S}_{k,m,n}$. Then, $\det \mathbf{M}[(S_i, r_i)]_{i=1}^m \neq 0$ if and only if for any nonempty $\Omega \subset [m]$,*

$$\left| \bigcap_{i \in \Omega} S_i \right| + \sum_{i \in \Omega} r_i \leq \max_{i \in \Omega} |S_i| + r_i \quad (9)$$

◇

Proof. See Appendix A. □

In Theorem 3, if we let $k = m$, $r_i = 1$, and $|S_i| = k - 1$, we will get Corollary 1.

Corollary 1. *Let $S_1, S_2, \dots, S_k \subset [n]$ such that $|S_i| = k - 1$. Then, the determinant of*

$$\mathbf{M}[(S_i, 1)]_{i=1}^k = \begin{pmatrix} 1 & \sum_{j \in S_1} \alpha_j & \cdots & \prod_{j \in S_1} \alpha_j \\ 1 & \sum_{j \in S_2} \alpha_j & \cdots & \prod_{j \in S_2} \alpha_j \\ \vdots & \vdots & & \vdots \\ 1 & \sum_{j \in S_m} \alpha_j & \cdots & \prod_{j \in S_m} \alpha_j \end{pmatrix} \quad (10)$$

is nonzero if and only if for any nonempty $\Omega \subset [k]$,

$$\left| \bigcap_{i \in \Omega} S_i \right| \leq k - |\Omega| \quad (11)$$

◇

APPENDIX A

PROOF OF THEOREM 3

Suppose that for some nonempty $\Omega \subset [m]$, the condition (9) is not true. Let $S_0 = \bigcap_{i \in \Omega} S_i$, $r_0 = \sum_{i \in \Omega} r_i$, $k' = \max_{i \in \Omega} |S_i| + r_i$. Then, $|S_0| + r_0 > k'$. Consider the r_0 rows of \mathbf{M} in the blocks indexed in Ω . They all have zeros in their last $k - k'$ entries. Let $\mathbf{M}_0 \in \mathbb{K}_n^{r_0 \times k'}$ be the submatrix consisting of these rows without including the last $k - k'$ columns. We will prove that $\text{rank } \mathbf{M}_0 < r_0$, which implies $\det \mathbf{M} = 0$. Let $\mathbf{W} = ((-\alpha_j)^{1-i})_{i \in [k'], j \in S_0}$ be $k' \times |S_0|$ Vandermonde matrix. Then, $\mathbf{M}_0 \cdot \mathbf{W} = 0$ because the polynomials with the coefficients in the rows of \mathbf{M}_0 vanish at $-\alpha_j$ for $j \in S_0$. Hence,

$$\text{rank } \mathbf{M}_0 \leq k' - \text{rank } \mathbf{W} \leq k' - \min\{k', |S_0|\} < r_0 \quad (12)$$

which proves the first direction.

For the other direction, we will apply induction on the parameters (k, m, n) considered in the lexicographical order. For $m = 1$, $\mathcal{S}_{k,1,n} = \{[(\emptyset, k)]\}$ and $\det \mathbf{M}[(\emptyset, k)] = \det \mathbf{I}_k = 1$. For $n = 0$, all of S_i 's are empty; hence, for $\Omega = [m]$, (9) yields $m = 1$, for which, we already showed $\det \mathbf{M} = 1$. For $k \geq m \geq 2$ and $n \geq 1$, assume that the statement is true for parameters (k', m', n') that are smaller than (k, m, n) with respect to lexicographical order. Take any $[(S_i, r_i)]_{i=1}^m \in \mathcal{S}_{k,m,n}$ that satisfies the condition (9). We will prove that $\det \mathbf{M}[(S_i, r_i)]_{i=1}^m \neq 0$ under three cases:

1) There exists $\Omega_1 \subset [m]$ such that $2 \leq |\Omega_1| \leq m - 1$ and

$$\left| \bigcap_{i \in \Omega_1} S_i \right| + \sum_{i \in \Omega_1} r_i = \max_{i \in \Omega_1} |S_i| + r_i \quad (13)$$

- 2) There exists a unique $i \in [m]$ such that $|S_i| + r_i = k$.
3) Else (i.e. 1 and 2 are false).

Case 1

Let $\Omega_2 = \{0\} \cup [m] - \Omega_1$. Note that $2 \leq |\Omega_1|, |\Omega_2| \leq m - 1$. Define

$$S_0 = \bigcap_{i \in \Omega_1} S_i, \quad r_0 = \sum_{i \in \Omega_1} r_i \quad (14)$$

Then, (13) becomes

$$|S_0| + r_0 = \max_{i \in \Omega_1} |S_i| + r_i \quad (15)$$

Define $S'_i = S_i - S_0$ for $i \in \Omega_1$. Then,

$$[(S'_i, r_i)]_{i \in \Omega_1} \in \mathcal{S}_{r_0, |\Omega_1|, n}, \quad [(S_i, r_i)]_{i \in \Omega_2} \in \mathcal{S}_{k, |\Omega_2|, n} \quad (16)$$

The first one is true because $r_0 = \sum_{i \in \Omega_1} r_i$ and for any $i \in \Omega_1$, by (15),

$$|S'_i| + r_i = |S_i| + r_i - |S_0| \leq r_0 \quad (17)$$

The second one is true because

$$k = \sum_{i=1}^m r_i = \sum_{i \in \Omega_1} r_i + \sum_{i \in [m] - \Omega_1} r_i = \sum_{i \in \Omega_2} r_i, \quad (18)$$

$|S_i| + r_i \leq k$ for $i \in [m] - \Omega_1$ and $|S_0| + r_0 \leq k$ due to (15).

By the induction hypothesis, the statement is true for $[(S'_i, r_i)]_{i \in \Omega_1}$ and $[(S_i, r_i)]_{i \in \Omega_2}$. We will show that both satisfy the condition (9):

1) For any nonempty $\Omega \subset \Omega_1$,

$$\left| \bigcap_{i \in \Omega} S'_i \right| + \sum_{i \in \Omega} r_i = \left| \bigcap_{i \in \Omega} S_i \right| - |S_0| + \sum_{i \in \Omega} r_i \quad (19)$$

$$\leq \max_{i \in \Omega} |S_i| + r_i - |S_0| \quad (20)$$

$$= \max_{i \in \Omega} |S'_i| + r_i \quad (21)$$

2) For any nonempty $\Omega \subset \Omega_2$, if $0 \notin \Omega_2$, then $\Omega \subset [m]$ and (9) holds trivially. Assume $\Omega = \{0\} \cup \Omega'$ for some $\Omega' \subset [m] - \Omega_1$. Then,

$$\left| \bigcap_{i \in \Omega} S_i \right| + \sum_{i \in \Omega} r_i = \left| \bigcap_{i \in \Omega_1 \cup \Omega'} S_i \right| + \sum_{i \in \Omega_1 \cup \Omega'} r_i \quad (22)$$

$$\leq \max_{i \in \Omega_1 \cup \Omega'} |S_i| + r_i \quad (23)$$

$$= \max_{i \in \Omega} |S_i| + r_i \quad (24)$$

Hence, we have that

$$\det \mathbf{M}[(S'_i, r_i)]_{i \in \Omega_1} \neq 0, \quad \det \mathbf{M}[(S_i, r_i)]_{i \in \Omega_2} \neq 0 \quad (25)$$

Now, we will use Proposition 1. Define for $i \in \{0\} \cup [m]$,

$$p_i = x^{k - |S_i| - r_i} \prod_{j \in S_i} (x + \alpha_j) \quad (26)$$

and for $i \in \Omega_1$,

$$p'_i = x^{r_0 - |S'_i| - r_i} \prod_{j \in S'_i} (x + \alpha_j) \quad (27)$$

Note that for $i \in \Omega_1$, $p_i = p'_i p_0$.

Consider any $q_1, \dots, q_m \in \mathbb{K}_n[x]$ such that $\deg q_i \leq r_i - 1$ for $i \in [m]$ and $\sum_{i=1}^m p_i q_i = 0$. We need to prove that $q_i = 0$ for all $i \in [m]$. Define

$$q_0 = \sum_{i \in \Omega_1} p'_i q_i \quad (28)$$

Note that $\deg q_0 \leq r_0 - 1$:

$$\deg q_0 \leq \max_{i \in \Omega_1} (\deg p'_i + \deg q_i) \quad (29)$$

$$\leq \max_{i \in \Omega_1} ((r_0 - r_i) + (r_i - 1)) \quad (30)$$

$$= r_0 - 1 \quad (31)$$

Also, we can write that

$$0 = \sum_{i=1}^m p_i q_i = p_0 \sum_{i \in \Omega_1} p'_i q_i + \sum_{i \in [m] - \Omega_1} p_i q_i = \sum_{i \in \Omega_2} p_i q_i \quad (32)$$

Then, by Proposition 1, we get $q_i = 0$ for all $i \in \Omega_2$. Then, $q_0 = \sum_{i \in \Omega_1} p'_i q_i = 0$. Then, by Proposition 1, $q_i = 0$ for all $i \in \Omega_1$. Hence, $q_i = 0$ for all $i \in [m]$. By Proposition 1, $\det \mathbf{M}[(S_i, r_i)]_{i=1}^m \neq 0$.

Case 2

W.l.o.g., let m be the maximizer. Then, for $i \in [m - 1]$,

$$k = |S_m| + r_m > |S_i| + r_i \quad (33)$$

Then, the last column of $\mathbf{M}[(S_i, r_i)]_{i=1}^m$ is all zero except the last entry, which is $\prod_{j \in S_m} \alpha_j$.

Hence, we have

$$\det \mathbf{M}[(S_i, r_i)]_{i=1}^m = \det \mathbf{M}[(S_i, r'_i)]_{i=1}^m \cdot \prod_{j \in S_m} \alpha_j \quad (34)$$

where $r'_m = r_m - 1$ and $r'_i = r_i$ for $i \in [m - 1]$ assuming that $r_m \geq 2$. (If $r_m = 1$, the first multiplier would be $\det \mathbf{M}[(S_i, r_i)]_{i=1}^{m-1}$, which is nonzero by the induction hypothesis.)

Note that $[(S_i, r'_i)]_{i=1}^m \in \mathcal{S}_{k-1, m, n}$ since $\sum_{i=1}^m r'_i = k - 1$ and $|S_i| + r'_i \leq k - 1$ for any $i \in [m]$ due to the unique maximizer assumption.

By the induction hypothesis, the statement is true for $[(S_i, r'_i)]_{i=1}^m$. If we prove that it satisfies the condition (9), then $\det \mathbf{M}[(S_i, r_i)]_{i=1}^m \neq 0$.

For any nonempty $\Omega \subset [m]$, if $m \notin \Omega$, then (9) holds trivially. Assume $m \in \Omega$.

$$\left| \bigcap_{i \in \Omega} S_i \right| + \sum_{i \in \Omega} r'_i = \left| \bigcap_{i \in \Omega} S_i \right| - 1 + \sum_{i \in \Omega} r_i \quad (35)$$

$$\leq \max_{i \in \Omega} |S_i| + r_i - 1 \quad (36)$$

$$= k - 1 \quad (37)$$

$$= \max_{i \in \Omega} |S_i| + r'_i \quad (38)$$

Case 3

For any nonempty $\Omega \subset [m]$ such that $|\Omega| \neq 1, m$, we have

$$\left| \bigcap_{i \in \Omega} S_i \right| + \sum_{i \in \Omega} r_i \leq \max_{i \in \Omega} |S_i| + r_i - 1 \quad (39)$$

Also, there exist at least two maximizers of $|S_i| + r_i$. W.l.o.g., assume that

$$k = |S_m| + r_m = |S_{m-1}| + r_{m-1} \quad (40)$$

If $S_m = S_{m-1}$, we get a contradiction in (9):

$$r_m + r_{m-1} \leq \max\{r_m, r_{m-1}\} \quad (41)$$

Then, either $S_{m-1} \neq [n]$ or $S_m \neq [n]$. W.l.o.g., we can assume that $n \notin S_m$. Substitute $\alpha_n = 0$:

$$\det \mathbf{M}[(S_i, r_i)]_{i=1}^m |_{\alpha_n=0} = \det \mathbf{M}[(S'_i, r_i)]_{i=1}^m \quad (42)$$

where $S'_i = S_i - \{n\}$.

Note that $[(S'_i, r_i)]_{i=1}^m \in \mathcal{S}_{k, m, n-1}$ since $S'_i \subset [n - 1]$ and $|S'_i| + r_i \leq |S_i| + r_i \leq k$ for $i \in [m]$.

By the induction hypothesis, the statement is true for $[(S'_i, r_i)]_{i=1}^m$. If we prove that it satisfies the condition (9), then $\det \mathbf{M}[(S_i, r_i)]_{i=1}^m \neq 0$.

For $|\Omega| = 1$, (9) holds trivially. For $|\Omega| \neq 1, m$, we have

$$\left| \bigcap_{i \in \Omega} S'_i \right| + \sum_{i \in \Omega} r_i \leq \left| \bigcap_{i \in \Omega} S_i \right| + \sum_{i \in \Omega} r_i \quad (43)$$

$$\leq \max_{i \in \Omega} |S_i| + r_i - 1 \quad (44)$$

$$\leq \max_{i \in \Omega} |S'_i| + r_i \quad (45)$$

For $\Omega = [m]$, it is enough to show that $k = \max_{i \in [m]} |S'_i| + r_i$, which is true because

$$|S'_m| + r_m = |S_m| + r_m = k \quad (46)$$

REFERENCES

- [1] H. Yildiz and B. Hassibi, "Further progress on the gm-mds conjecture for reed-solomon codes," *arXiv preprint arXiv:1801.07865*, 2018.
- [2] S. H. Dau, W. Song, and C. Yuen, "On the existence of mds codes over small fields with constrained generator matrices," in *International Symposium on Information Theory (ISIT)*. IEEE, 2014, pp. 1787–1791.
- [3] S. H. Dau, W. Song, and C. Yuen, "On simple multiple access networks," *IEEE Journal on Selected Areas in Communications*. vol. 33, no. 2, pp. 236–249, 2015.
- [4] M. Yan and A. Sprintson, "Algorithms for weakly secure data exchange," in *International Symposium on Network Coding (NetCod)*. IEEE, 2013, pp. 1–6.
- [5] W. Halbawi, T. Ho, H. Yao, and I. Duursma, "Distributed reed-solomon codes for simple multiple access networks," in *International Symposium on Information Theory (ISIT)*. IEEE, 2014, pp. 651–655.
- [6] W. Halbawi, Z. Liu, and B. Hassibi, "Balanced reed-solomon codes," in *International Symposium on Information Theory (ISIT)*. IEEE, 2016, pp. 935–939.
- [7] W. Halbawi, Z. Liu, and B. Hassibi, "Balanced Reed-Solomon codes for all parameters," in *Information Theory Workshop (ITW)*. IEEE, 2016, pp. 409–413.
- [8] M. Yan, A. Sprintson, and I. Zelenko, "Weakly secure data exchange with generalized reed solomon codes," in *International Symposium on Information Theory (ISIT)*. IEEE, 2014, pp. 1366–1370.
- [9] A. Heidarzadeh and A. Sprintson, "An algebraic-combinatorial proof technique for the gm-mds conjecture," in *International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 11–15.
- [10] W. Song and K. Cai, "Generalized reed-solomon codes with sparsest and balanced generator matrices," *arXiv preprint arXiv:1801.02315*, 2018.