On the MDS Conjecture

Aiwei Liu

Shanghai Jiao Tong University

December 31, 2025

Table of Contents

1. Introduction
2. MDS code
3. Main proof
4. Consequences

MDS Conjecture

Conjecture 1 (MDS Conjecture)

A set S of vectors of the vector space $\mathbb{F}_q^k$ such that every subset of S of size $k \leq q$ is a basis, has size at most $q+1$, unless q is even and $k=3$ or $k=q-1$, in which case it has size at most $q+2$.

Example 2

$k=4, q=5$

$$S = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 3 & 4 & 0 \\ 0 & 1 & 4 & 4 & 1 & 0 \\ 0 & 1 & 3 & 2 & 4 & 1 \end{pmatrix}$$

Main Theorem

Theorem 3 (Ball)

Let $q = p^h$. A set S of vectors of the vector space $\mathbb{F}_q^k$ such that every subset of S of size $k \leq q$ is a basis, has size at most $q + k + 1 - \min(k, p)$.

This proves the MDS Conjecture in the case $k \leq p$, which includes the entire prime case.

Table of Contents

1. Introduction
2. MDS code
3. Main proof
4. Consequences

MDS code

Theorem 4 (Singleton Bound)

Let $M_q(n, d)$ denote the maximum possible size such that there is an $(n, M_q(n, d))$ code. Then

$$M_q(n, d) \leq q^{n-d+1}.$$

MDS code

Theorem 4 (Singleton Bound)

Let $M_q(n,d)$ denote the maximum possible size such that there is an $(n, M_q(n,d))$ code. Then

$$M_q(n,d) \leq q^{n-d+1}.$$

Definition 5 (MDS code)

A code attaining this bound is called an MDS code.

The following definitions are equivalent.

- $\mathfrak{C}$ is an MDS $[n,k,d=n-k+1]_q$ linear code.
- Every k-columns of the generator matrix G are $\mathbb{F}_q$ linearly independent.
- Every $n-k$-columns of the parity check matrix H are $\mathbb{F}_q$ linearly independent.
- $\mathfrak{C}^{\perp}$ is an MDS $[n, n-k, n-d=k+1]_q$ linear code.

Table of Contents

1 Introduction

2 MDS code

3 Main proof

4 Consequences

An easy bound

Let S be a set of vectors of $\mathbb{F}_q^k$ such that every subset of S of size k is a basis.

Lemma 6

$$|S| \leq q + k - 1$$

An easy bound

Let S be a set of vectors of $\mathbb{F}_q^k$ such that every subset of S of size k is a basis.

Lemma 6

$$|S| \leq q + k - 1$$

Consider the $k-2$-dimensional subspace U spanned by $k-2$ vectors of S. Each of the $q+1$ hyperplanes containing U contains at most one other vector of S. Thus, $|S| \leq k - 2 + q - 1$.

Tangent Function

Given subset C of S with size $k-2$. Let $t = q + k - 1 - |S|$. It should be the number of hyperplanes Σ with $\Sigma \cap S = C$.

$T_C(u) = \prod_{f(u)=0 \text{ defines } \Sigma} f(u)$ is called the *tangent function*. Note that this is defined up to scalar factor.

Example 7

$k = 3, q = 5$

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$T_{\{x\}}(a,b,c) = (2b-c)(3b-c)(4b-c)$$

$$T_{\{y\}}(a,b,c) = (2c-a)(3c-a)(4c-a)$$

$$T_{\{z\}}(a,b,c) = (2a-b)(3a-b)(4a-b)$$

Interpolation of Tangent Function

Lemma 8

If $|S| \geq k+t$ then for any disjoint subsets $Y = \{y_1, y_2, \cdots, y_{k-2}\}$ and $E = \{a_1, a_2, \cdots, a_{t+2}\}$ of S,

$$0 = \sum_{a \in E} T_Y(a) \prod_{z \in E \setminus \{a\}} \det(a, z, y_1, y_2, \cdots, y_{k-2})^{-1}.$$

Proof.

We consider this formula in $\mathbb{F}_q^k/\mathrm{span}(Y)/\sim$, where $x \sim y$ iff $x = \lambda y$.

$$T_Y(x) = \sum_{j=1}^{t+1} T_Y(a_j) \prod_{l=1, l \neq j}^{t+1} \frac{\det(x, a_l, y_1 \cdots, y_{k-2})}{\det(a_j, a_l, y_1 \cdots, y_{k-2})}.$$

Let $x = a_{t+2}$. $\square$

"Higher" Interpolation of Tangent Function

Exchange $b \in E$ and $y_1 \in Y$:

$$0 = T_{(Y\setminus\{y_1\})\cup\{b\}}(y_1) \prod_{z\in E\setminus\{b\}} \det(y_1, z, b, y_2, \cdots, y_{k-2})^{-1}$$
$$+ \sum_{a\in E\setminus\{b\}} T_{(Y\setminus\{y_1\})\cup\{b\}}(a) \prod_{z\in(E\setminus\{a,b\})\cup\{y_1\}} \det(a, z, b, y_2, \cdots, y_{k-2})^{-1}$$

Multiply $\frac{T_Y(b)}{T_{(Y\setminus\{y_1\})\cup\{b\}}(y_1)}$ and sum over $b \in E$:

$$0 = \sum_{a_1,a_2\in E} \frac{T_{\theta_1}(a_1)}{T_{\theta_2}(y_1)} T_{\theta_2}(a_2) \prod_{z\in(E\cup Y)\setminus(\{a_2\}\cup\theta_2)} \det(a_2, z, \theta_2)^{-1}$$

...

Continue exchanging $b \in E$ and $y_r \in Y$, we'll get

“Higher” Interpolation of Tangent Function

Lemma 9

If $|S| \geq k+t$ *then for any disjoint subsets* $Y = \{y_1, \cdots, y_{k-2}\}$ *and* E *of* S *with* $|E| = t+2$ *and* $r \leq \min(k-1, t+2)$,

$$0 = \sum_{a_1,\cdots,a_r \in E} \left(\prod_{i=1}^{r-1} \frac{T_{\theta_i}(a_i)}{T_{\theta_{i+1}}(y_i)}\right) T_{\theta_r}(a_r) \prod_{z \in (E \cup Y) \setminus (\theta_r \cup \{a_r\})} \det(a_r, z, \theta_r)^{-1},$$

where $\theta_i = (a_1, \cdots, a_{i-1}, y_i, \cdots, y_{k-2})$ *is an ordered sequence.*

"Higher" Interpolation of Tangent Function

Moreover, by transposing a_j and a_{j+1}, we can prove that the r! terms in the sum for a fixed r-element subset of E are the same.

Lemma 10

If $|S| \geq k+t$ then for any disjoint subsets $Y = \{y_1, \cdots, y_{k-2}\}$ and E of S with $|E| = t+2$ and E an ordered sequence, and $r \leq \min(k-1, t+2)$,

$$0 = r! \sum_{a_1 < \cdots < a_r \in E} \left(\prod_{i=1}^{r-1} \frac{T_{\theta_i}(a_i)}{T_{\theta_{i+1}}(y_i)}\right) T_{\theta_r}(a_r) \prod_{z \in (E \cup Y) \setminus (\theta_r \cup \{a_r\})} \det(a_r, z, \theta_r)^{-1}$$

Proof of Theorem 3

Remark 11

The following statements are equivalent:

- *Every k-columns of the generator matrix G are $\mathbb{F}_q$ linearly independent.* $[S, k, t]$
- *Every $n-k$-columns of the parity check matrix H are $\mathbb{F}_q$ linearly independent.* $[S', k', t']$

Proof of Theorem 3

Proof.

- $|S| \leq k + t$ and $|S'| \leq k' + t'$
  $\Rightarrow k' \leq t$ and $k \leq t'$
  $\Rightarrow |S| \leq q - 1 + (t' - t)$ and $|S| = |S'| \leq q - 1 + (t - t')$
  $\Rightarrow |S| \leq q - 1$
- $|S| \geq k + t$. Either $t \geq k - 2$, or $t + 2 \leq k - 1$, then
  $$0 = (t+2)!\left(\prod_{i=1}^{t+1} \frac{T_{\theta_i}(a_i)}{T_{\theta_{i+1}}(y_i)}\right) T_{\theta_{t+2}}(a_{t+2}) \prod_{z \in Y \setminus \theta_{t+2}} \det\left(a_{t+2}, z, \theta_{t+2}\right)^{-1}$$
  $\Rightarrow (t+2)! \equiv 0 \pmod p$
  $\Rightarrow t \geq p - 2.$
  Therefore,
  $|S| \leq k + q - 1 - \min\left(p - 2, k - 2\right) = k + q + 1 - \min\left(p, k\right).$
- $k' + t' \leq |S| \leq k + t \Rightarrow k' \leq k.$
  $|S'| \leq k' + q + 1 - \min\left(p, k'\right) \leq k + q + 1 - \min\left(p, k\right).$

Classification of the largest subsets

Theorem 12 (Ball)

If $p \geq k$ then a set S of $q+1$ vectors of $\mathbb{F}_q^k$ such that every subset of S of size k is a basis, is equivalent to

$$S = \{(1, t, t^2, \cdots, t^{k-1}) | t \in \mathbb{F}_q\} \cup \{(0, \cdots, 0, 1)\}.$$

Table of Contents

1 Introduction

2 MDS code

3 Main proof

4 Consequences

Consequences

- MDS code
- Matroid

Corollary 13

If, for a matroid $M = (E, F)$ and prime $p \geq r(E)$, there is a subset $S \subset E$ of size $p + 2$ in which every subset of size $r(E)$ is a basis, then M is not representable over $\mathbb{F}_p$.

Some thoughts about the main proof

$t+2$ **might be more crucial rather than** t**:**

Why considering $|S| \leq k+t$ and $|S| \geq k+t$?

- $t+2$ might be the more crucial thing
- Let $m := t+2$, consider $|S| \geq k+m$ and $|S'| \geq k'+m'$ in the first part of the proof $\Rightarrow |S| = q+1$

What is $k+t$ (or $k+m$)?

- $k+m = k+\frac{1}{2}(q+1)$

References

[1] Simeon Ball, *On sets of vectors of a finite vector space in which every subset of basis size is a basis. J. Eur. Math. Soc. 14 (2012), no. 3, pp. 733–748.*