Distribution of the Minimum Distance of Random Linear Codes

Jing Hao*, Han Huang†, Galyna Livshyts‡ and Konstantin Tikhomirov§
Georgia Institute of Technology
Georgia, USA
Email: *jhao6@gatech.edu, †hhuang421@gatech.edu, ‡glivshyts6@math.gatech.edu, §konstantin.tikhomirov@math.gatech.edu

***Abstract*—In this paper, we study the distribution of the minimum distance (in the Hamming metric) of a random linear code of dimension k in $\mathbb{F}_q^n$. We provide quantitative estimates showing that the distribution function of the minimum distance is close (superpolynomially in n) to the cumulative distribution function of the minimum of $(q^k-1)/(q-1)$ independent binomial random variables with parameters $\frac{1}{q}$ and n. The latter, in turn, converges to a Gumbel distribution at integer points when $\frac{k}{n}$ converges to a fixed number in $(0,1)$. In a sense, our result shows that apart from identification of the weights of parallel codewords, the probabilistic dependencies introduced by the linear structure of the random code, produce a negligible effect on the minimum code weight. As a corollary of the main result, we obtain an asymptotic improvement of the Gilbert–Varshamov bound for $2 < q < 49$.**

A full version of this paper is accessible at: https://arxiv.org/abs/1912.12833/

I. Introduction

Let $\mathbb{F}_q$ be a finite field. A *linear code* C is a subspace of $\mathbb{F}_q^n$ where n is the length of the code. The parameter q of the field is referred to as the alphabet size. The *size* of C is the number of elements in C. For a (not necessarily linear) code with size M, alphabet size q, and length n, the *rate* R is defined to be $\log_q(M)/n$. For a linear code this number is equal to k/n, where k is the dimension of the code as a vector space.

Another fundamental parameter is the relative minimum distance. Let *the Hamming distance* between any two codewords $u = (u_1, \cdots, u_n)$ and $v = (v_1, \cdots, v_n)$ in $\mathbb{F}_q^n$ be given by

$$d(u,v) := |\{1 \leq i \leq n,\ u_i \neq v_i\}|,$$

and *the Hamming weight* of a codeword u be defined as $\mathrm{wt}\,(u) := d(u,0)$. For linear codes, the minimum distance between two distinct codewords in a code is equal to the minimum weight over all nonzero codewords. The *relative minimum distance* δ is defined as the ratio $\frac{d}{n}$.

In coding theory, the trade-off between the code rate R and error-correcting ability δ is a central topic of study. Let q be fixed. For linear codes, It has been proved that there exists a function $\alpha_q(\cdot)$ with the following property: for any $\delta_0 \in (0, 1-1/q)$ and any $R_0 \leq \alpha_q(\delta_0)$, there is an infinite sequence of linear codes with the relative minimum distance converging to δ_0 and the rate converging to R_0; on the other hand, for every $R_0 > \alpha_q(\delta_0)$, such a sequence does not exist. (See [1] and [2].) An explicit description of $\alpha_q(\cdot)$ remains a major open problem (see [3]–[5], as well as [6] for an upper bound for α_q). (See [1]) Considerable work has been done to obtain explicit constructions for linear codes with good rate and relative minimum distance (we refer, in particular, to [7]).

Rather than considering special codes, one may be interested in studying the statistical properties on the space of all linear codes, using probabilistic methods. A classical result in this direction is the *Gilbert–Varshamov argument*. Gilbert [8] and Varshamov [9] gave lower bound for the size of a (not necessarily linear) code given n and d. Let $A_q(n,d)$ be the maximal size of a code of length n over $\mathbb{F}_q$ and with minimum distance d. Then

$$A_q(n,d) \geq \frac{q^n}{\sum_{j=0}^{d-1} \binom{n}{j} (q-1)^j},$$

and, moreover, there are *linear* codes that can achieve this bound i.e. there exists a linear code over $\mathbb{F}_q$ with dimension at least $n - \lfloor \log_q \sum_{j=0}^{d-1} (q-1)^j \rfloor$. The proof of the result can be obtained by a union bound argument.

Recall that *the q-ary entropy function* is defined by

$$H_q(x) := x \log_q(q-1) - x \log_q(x) - (1-x)\log_q(1-x).$$

In [10], it was shown that for $q = 2$ and given a rate R_0 and $\varepsilon > 0$, the probability that a random linear code of length n and rate R_0, uniformly distributed on the set of linear codes of the given length and rate, has the minimum distance $d < n(\delta_0 - \varepsilon)$, is exponentially small in n. Here $0 < \delta_0 < \frac{1}{2}$ is the solution of the equation $R_0 = 1 - H_2(\delta_0)$. On the other hand, if we fix any δ_0 satisfying $0 \leq \delta_0 < 1 - \frac{1}{q}$ and $0 < \varepsilon \leq 1 - H_q(\delta_0)$, then the Gilbert–Varshamov argument implies that there exist infinitely many linear codes with a rate $R \geq 1 - H_q(\delta_0) - \varepsilon$. By taking $\varepsilon \to 0$, one would obtain a lower bound for the function $\alpha_q(\delta)$ mentioned above:

$$\alpha_q(\delta) \geq 1 - H_q(\delta).$$

In fact, as was proved in [11], the following law of large numbers holds for the minimum distance of a sequence of random linear codes: if $n \to \infty$ and the rate k/n converges to a number $R_0 \in (0,1)$ then the relative minimum distance converges (almost surely) to the number δ_0 given by the equation $R_0 = 1 - H_q(\delta_0)$. Moreover, the probability that a random linear code of length n has the relative minimum distance outside of the interval $[\delta_0 - \varepsilon, \delta_0 + \varepsilon]$, is exponentially small in n (we remark here that in the same paper it was

978-1-7281-6432-8/20/$31.00 ©2020 IEEE

shown that the minimum distance of random non-linear codes is asymptotically worse than in the linear setting).

Our goal in this paper is to obtain a more precise description of the distribution of the minimum distance of random linear codes. The main statement is the following

Theorem I.1. *For any prime power q and any real numbers $R_1 < R_2$ in $(0,1)$ there is $c(R_1,R_2,q) > 0$ with the following property. Let positive integers k, n satisfy $R_1 \le k/n \le R_2$, and let $\mathcal{C}$ be the random linear code uniformly distributed on the set of all linear codes in $\mathbb{F}_q^n$ of dimension k. Denote by F_{dmin} the cumulative distribution function of the minimum distance of $\mathcal{C}$. Further, let $w_{\min}$ be the minimum weight of $\frac{q^k-1}{q-1}$ i.i.d. uniform random vectors in $\mathbb{F}_q^n$, and F_{wmin} be its cumulative distribution function. Then*

$$\sup_{x\in\mathbb{R}} \big|F_{\mathrm{dmin}}(x) - F_{\mathrm{wmin}}(x)\big| = O\big(\exp(-c(R_1,R_2,q)\sqrt{n})\big).$$

A surprising feature of this result is that the distribution of the minimum distance can be approximated by a c.d.f. of the minimum of i.i.d. binomial variables with precision superpolynomial in n. In a sense, this result asserts that dependencies between codeword weights introduced by the linear structure of the code, produce a negligible effect on the distribution of the minimum weight.

The proof of the result is based on analysis of moments of certain functionals associated with the code. We remark that in a recent work by Linial and Mosheiff [12], the authors calculated centered moments for the number of codewords of a random linear code with a given weight. The approach used in that paper influenced our work.

As an immediate corollary of our result, we obtain the following statement which gives an $\Theta(n^{1/2})$ improvement over the classical Gilbert–Varshamov bound:

Corollary I.1. *For any prime power q, any $\alpha \in (0,1)$, any integer n, and $d \in [\alpha n, (1-\alpha)(n - n/q)]$ there is a linear code with minimum distance d of size at least*

$$cn^{1/2}\,\frac{q^n}{\sum_{j=0}^{d-1}\binom{n}{j}(q-1)^j},$$

where $c > 0$ may only depend on α and q.

We note that existence of *non-linear* codes of size at least $cn\,\frac{q^n}{\sum_{j=0}^{d-1}\binom{n}{j}(q-1)^j}$ has been previously established in [4], [13]. Linear double-circulant binary codes beating the Gilbert–Varshamov bound were considered in [14]. To our best knowledge, the above improvement for $2 < q < 49$ is new.

Further, we obtain an explicit limit theorem for the distribution of the minimum distance. Due to the discrete nature of our random variable, the convergence to a Gumbel distribution can only be established on the points along certain arithmetic progressions:

Theorem I.2 (The limit theorem for the minimum distance)**.** *Let q be a prime power, and let $R_1 < R_2$ be numbers in $(0,1)$. Let (k_n) be a sequence of positive integers such that $R_1 \le k_n/n \le R_2$ for all large n. For any n let $d_{\min}(n)$ be the minimum distance of the random linear code uniformly distributed on the set of linear codes of length n and dimension k_n. Further, for any n let $d_0(n)$ be the largest integer satisfying*

$$u(n) := \frac{q^{k_n}-1}{q-1}\sum_{i=0}^{d_0(n)}\binom{n}{j}\Big(1-\frac{1}{q}\Big)^i q^{i-n} \le 1.$$

Denote by ξ_n the random variable

$$\xi_n := \big(d_0(n) - d_{\min}(n)\big)\log\frac{(q-1)(n-d_0(n))}{d_0(n)} - \log u(n).$$

Then, as $n\to\infty$, we have

$$\sup\Big\{|\mathbb{P}\{\xi_n < t\} - G(t)| :\ t \in \log\frac{(q-1)(n-d_0(n))}{d_0(n)}\,\mathbb{Z} - \log u(n)\Big\} \longrightarrow 0,$$

where G is the Gumbel law given by $G(t) = e^{-e^{-t}}$.

The paper is organized as follows. In Section II, we consider some auxiliary results for the binomial distribution, including a limiting result for the minimum of i.i.d. binomial random variables. At the end of the section, we show that the main result of the paper implies Theorem I.2.

In Section III, we consider the set of random vectors $\{Y_a : a \in \mathbb{F}_q^k\backslash\{0\}\}$ uniformly distribtued on $\mathbb{F}_q^n$ that are mutually independent up to the constraint that $Y_a = Y_b$ whenever a and b are parallel. We study moments of the random variable that counts number of codewords with weights less than or equal to d in this configuration as well as that of random linear code ensemble and give a quantitative comparison between them.

Finally, in Section IV we give the comparison of the c.d.f. of minimum distance between these two ensembles. Due to the discrete nature of this problem, either c.d.f. can be obtained by solving a set of linear equations involving quantities we computed in previous sections. Then we give a quantitative comparison by estimating the truncation errors and moment differences.

Most of the proofs will be omitted. Link for a longer version of our paper is given in the abstract. Interested readers can refer to that for more details.

II. Auxiliary results for the binomial distribution

Our goal in this section is to obtain quantitative estimates for the distribution of the minimum of i.i.d. binomial random variables (with specially chosen parameters).

Let $1 \le m \le (q-1)^n$ and let $X_1,\dots,X_m$ be i.i.d. vectors uniformly distributed in $\mathbb{F}_q^n$. Here, we are interested in estimates of the quantities

$$\mathbb{P}\big\{\min_{i\le m}\mathrm{wt}\,(X_i) \le d\big\}, \quad d \ge 0,$$

where $\mathrm{wt}\,(X_i)$ is the number of non-zero components of X_i. Denote

$$\rho_d := \mathbb{P}\big\{\mathrm{wt}\,(X_1) \le d\big\} = \sum_{i=0}^{d} \binom{n}{i} \Big(1 - \frac{1}{q}\Big)^i q^{i-n}.$$

We start by recording the following approximations to ρ_d:

Proposition II.1. *For any $\alpha \in (0,1)$ there is $C_\alpha > 0$ with the following property. Assume that $n \ge 1$ and $C_\alpha \log(n) \le d \le (1-\alpha)(1-1/q)n$. Then we have*

$$\frac{\rho_d}{\binom{n}{d} q^{-n}(q-1)^d} = \big(1 + O_\alpha(\log n/n)\big) \frac{n-d+1}{n - \left(\frac{q}{q-1}\right) d + 1}.$$

Furthermore, for any positive integer $t \le \sqrt{d}$, we have

$$\frac{\rho_{d+t}}{\rho_d} = \left(1 + O_\alpha\left(\frac{\log n}{n} + \frac{t^2}{d}\right)\right) \left(\frac{(q-1)(n-d)}{d}\right)^t.$$

This leads to the next proposition, which provides an approximation of the minimum of independent binomial variables in terms of the Gumbel distribution.

Recall that a Gumbel distribution is used to model the distribution of maximum value of various distributions, which is useful in predicting chances of rare events like natural disasters.

Proposition II.2. *Fix $q \ge 2$ and $\alpha \in (0,1)$. Let $q^{\alpha n} \le m \le q^{(1-\alpha)n}$ and let d_0 be the largest integer such that $\rho_{d_0} m \le 1$. Let $X_1, \dots, X_m$ be i.i.d. binomial random variables with parameters n and $\frac{1}{q}$, i.e.*

$$\mathbb{P}\{X_j = a\} = \binom{n}{a} \Big(1 - \frac{1}{q}\Big)^a q^{a-n}, \quad a = 0, 1, \dots, n,$$

and set $Y := \min_{j=1,\dots,m} X_j$. Then

$$\mathbb{P}\left\{Y - d_0 > \frac{-t}{\log \frac{(q-1)(n-d_0)}{d_0}} - \frac{\log(\rho_{d_0} m)}{\log \frac{(q-1)(n-d_0)}{d_0}}\right\}$$
$$= o_{\alpha,q}(1) + \exp\big(-e^{-t}\big),$$

for all $t \in \log \frac{(q-1)(n-d_0)}{d_0} \mathbb{Z} - \log(\rho_{d_0} m)$.

It is not difficult to see that the above proposition and the main theorem of the paper imply Theorem I.2.

III. Moments comparison for parallel codes

Fix $a \in \mathbb{R}^k$ and $d \ge 0$. Given the independent random vectors $X_1, \dots, X_k$ uniform on $\mathbb{F}_q^n$, we define

$$Z_d := \sum_{a \in \mathbb{F}_q^k \setminus \{0\}} W_a(d), \quad d \ge 0,$$

where $W_a(d)$ is the indicator of the event

$$\Big\{\mathrm{wt}\Big(\sum_{i=1}^{k} a_i X_i\Big) \le d\Big\}.$$

For any $a, b \in \mathbb{F}_q^k \setminus \{0\}$, we say a and b are parallel if there exists $f \in \mathbb{F}_q \setminus \{0\}$ such that $a = f\,b$ (here the multiplication is in the field $\mathbb{F}_q$). Notice that if a and b are parallel, then the supports of the linear combinations are the same, and thus $W_a(d) = W_b(d)$ whenever a and b are parallel.

Let $\{Y_a\}_{a \in \mathbb{F}_q \setminus \{0\}}$ be random vectors uniformly distributed on $\mathbb{F}_q^n$ and mutually independent up to the constraint that $Y_a = Y_b$ whenever a and b are parallel. Define

$$\widetilde{Z_d} := \sum_{a \in \mathbb{F}_q^k \setminus \{0\}} \widetilde{W_a}(d)$$

where $\widetilde{W_a}(d)$ is the indicator function of the event $\{\mathrm{wt}\,(Y_a) \le d\}$.

The goal of this section is to compare the moments of $\widetilde{Z_d}$ and Z_d assuming certain constraints on the parameters n, k and d. The main statement of the section is

Proposition III.1. *For any $\lambda_0 \in (0,1)$ there are $c_{\mathit{III.1}}(\lambda_0, q) > 0$ and $C_{\mathit{III.1}}(\lambda_0, q) > 0$ with the following property. Suppose $d, n \in \mathbb{N}$ satisfy $\frac{d}{n} \le \lambda_0(1 - \frac{1}{q})$, and $d^2/n^{3/2} \ge C_{\mathit{III.1}}(\lambda_0, q)$. Then for any positive integer $m \le c_{\mathit{III.1}}(\lambda_0, q) d^2/n^{3/2}$ such that $q^k \rho_d \ge \exp\big(-\frac{c_{\mathit{III.1}}(\lambda_0,q) d^4}{n^3 m}\big)$, we have*

$$\mathbb{E}{Z_d}^m = \big(1 + O(\exp(-c_{\mathit{III.1}}(\lambda_0, q) d^4/n^3)) + O(2^{-k/2})\big)\, \mathbb{E}\widetilde{Z_d}^m.$$

We briefly discuss the proof outline below.

First, let $\ell \le m \le k$ be positive integers. Suppose $I_1, \dots, I_\ell$ is a partitioning of $[m]$ into non-empty sets. Denote by $\Omega(I_1, \dots, I_\ell)$ the collection of all m–tuples $(a^1, \dots, a^m) \in \big(\mathbb{F}_q^k \setminus \{0\}\big)^m$ such that a^i is parallel to a^j if and only if $i, j \in I_t$ for some $t \in [\ell]$. Further, define

$$\Omega_\ell := \big\{(v^1, \dots, v^\ell) \in \big(\mathbb{F}_q^k \setminus \{0\}\big)^\ell :$$
$$v^1, \dots, v^\ell \text{ are pairwise non-parallel}\big\}.$$

Note that there is a natural $(q-1)^{m-\ell}$–to–one mapping from $\Omega(I_1, \dots, I_\ell)$ onto Ω_ℓ which assigns $(a_{\min\{j \in I_t\}})_{t=1}^{\ell}$ to each $(a^1, \dots, a^m)$.

Now, in view of the above remarks,

$$Z_d^m = \sum_{a^1,\dots,a^m \in \mathbb{F}_q^k \setminus \{0\}} \prod_{i=1}^{m} W_{a^i}(d)$$
$$= \sum_{\ell=1}^{m} \sum_{I_1,\dots,I_\ell} \sum_{v^1,\dots,v^\ell \in \Omega_\ell} (q-1)^{m-\ell} \prod_{i=1}^{\ell} W_{v^i}^{|I_i|}(d),$$

where the second summation is taken over all partitions $I_1, \dots, I_\ell$ of $[m]$ into non-empty sets. Notice that $W_{v^i}^{|I_i|}(d) = W_{v^i}(d)$, so we can simplify the above representation to

$$Z_d^m = \sum_{\ell=1}^{m} S(m,\ell)\,(q-1)^{m-\ell} \left(\sum_{v^1,\dots,v^\ell \in \Omega_\ell} \prod_{i=1}^{\ell} W_{v^i}(d)\right),$$

where $S(m, \ell)$ is the number of ways to partition $[m]$ into ℓ non-empty sets (a *Stirling number of the second kind*). The above formula works for $\widetilde{Z_d^m}$ as well, up to replacing $W_{v^i}(d)$ with $\widetilde{W_{v^i}}(d)$.

The central technical statement of the section is the following

Proposition III.2. *For any $\lambda_0 \in (0,1)$ there are $c_{III.2}(\lambda_0, q) > 0$ and $C_{III.2}(\lambda_0, q) > 0$ with the following property. Suppose $d, n \in \mathbb{N}$ satisfy $\frac{d}{n} \leq \lambda_0(1 - \frac{1}{q})$ and $d \geq C_{III.2}(\lambda_0, q)$. Suppose further that $s \leq k$, and $\left(v^1, v^2, \ldots, v^s\right)$ are linearly independent vectors in $\mathbb{F}_q^k$, and that $v^{s+1} = \sum_{i=1}^s c_i v^i$ for some $c_i \in \mathbb{F}_q \setminus \{0\}$. Then*

$$\mathbb{E} \prod_{i=1}^{s+1} W_{v^i}(d) \leq C \rho_d^s \exp(-c_{III.2}(\lambda_0, q) d^4/n^3)$$

where $C > 0$ is a universal constant.

As a corollary of the above statement, we have

Corollary III.3. *Suppose $d, n \in \mathbb{N}$ are as in Proposition III.2. Suppose further that $\ell \leq k$, and $v^1, v^2, \ldots, v^\ell$ are non-zero vectors in $\mathbb{F}_q^k$ such that the rank of $\left(v^1, v^2, \ldots, v^\ell\right)$ is $r < \ell$. Then*

$$\mathbb{E} \prod_{i=1}^{\ell} W_{v^i}(d) \leq C \rho_d^r \exp(-c_{III.2}(\lambda_0, q) d^4/n^3).$$

The following lemma gives the estimate of cardinality of the set of ℓ–tuples of vectors $\left(v^1, v^2, \ldots, v^\ell\right) \in \Omega_\ell$ with a given rank r. This completes the proof of III.1.

Lemma III.4. *For $r \leq \ell \leq k$, denote*

$$\Omega_{r,\ell} := \big\{\left(v^1, v^2, \ldots, v^\ell\right) \in \Omega_\ell : \\ \dim\left(\operatorname{span}\left(v^1, \ldots, v^\ell\right)\right) = r\big\}.$$

Then

$$|\Omega_{r,\ell}| \leq \binom{\ell}{r} q^{r(\ell - r)} \prod_{i=0}^{r-1} \left(q^k - q^i\right).$$

When $r = \ell$, equality holds, implying

$$\frac{|\Omega_{r,\ell}|}{|\Omega_{\ell,\ell}|} \leq \binom{\ell}{\ell - r} \frac{(q^r)^{\ell - r}}{\prod_{i=r}^{\ell-1} \left(q^k - q^i\right)}.$$

IV. Analysis of the distribution of the minimum distance

The goal of this section is to prove our main result comparing the distributions of the minimum distance of the random linear code, with the minimum $w_{\min}$ of the weights of the random vectors Y_a, $a \in \mathbb{F}_q^k \setminus \{0\}$ (defined earlier in the paper).

First, we state the "technical" version of the result:

Theorem IV.1. *For any $\lambda_0 \in (0,1)$ there are $c_{IV.1}(\lambda_0, q) > 0$ and $C_{IV.1}(\lambda_0, q) > 0$ with the following property. Let $n \geq 1$, and take any $L \geq e$. Assume further that k satisfies $C_{IV.1}(\lambda_0, q) L \log L \leq k \leq n$, and take any d such that*

$$C_{IV.1}(\lambda_0, q) \sqrt{L} n^{3/4} \leq d \leq \lambda_0 \Big(1 - \frac{1}{q}\Big) n,$$

and $c_{IV.1}(\lambda_0, q) L \geq q^k \rho_d \geq \exp\left(-\frac{c_{IV.1}(\lambda_0, q) d^2}{n^{3/2}}\right)$. Let, as before, $X_1, \ldots, X_k$ be i.i.d. random vectors uniformly distributed in $\mathbb{F}_q^n$, and denote

$$d_{\min} := \min\left\{\operatorname{wt}\left(\sum\nolimits_{i=1}^k a_i X_i\right),\ a \in \mathbb{F}_q^k \setminus \{0\}\right\}.$$

Then

$$\left|\mathbb{P}\{d_{\min} \leq d\} - \mathbb{P}\{w_{\min} \leq d\}\right| = O(\exp(-L)).$$

The theorem provides some freedom of the choice of the parameters, and includes a regime when the ratio k/n converges to one when $n \to \infty$. At the same time, we would like to provide a cleaner statement for the most important regime when k/n is "separated" from both 0 and 1. We obtain Theorem I.1 as a corollary of Theorem IV.1.

We give the proof outline of IV.1 as below.

For each $r \geq 0$, we let

$$M_d(r) := \mathbb{P}\big\{Z_d = r\big\}, \quad \widetilde{M}_d(r) := \mathbb{P}\big\{\widetilde{Z}_d = r\big\},$$

so that

$$\mathbb{P}\big\{d_{\min} \leq d\big\} = \mathbb{P}\big\{Z_d > 0\big\} = \sum_{r=1}^{\infty} M_d(r);$$

$$\mathbb{P}\big\{w_{\min} \leq d\big\} = \mathbb{P}\big\{Z_d > 0\big\} = \sum_{r=1}^{\infty} \widetilde{M}_d(r).$$

Observe further that the numbers $M_d(r)$ and $\widetilde{M}_d(r)$ satisfy the relations

$$\sum_{r=1}^{\infty} M_d(r) r^m = \mathbb{E} Z_d^m, \quad \sum_{r=1}^{\infty} \widetilde{M}_d(r) r^m = \mathbb{E} \widetilde{Z}_d^m, \quad m \geq 1.$$

These identities, together with the relations between $\mathbb{E} Z_d^m$ and $\mathbb{E} \widetilde{Z}_d^m$ obtained in the previous section, will allow us to compare $M_d(r)$ with $\widetilde{M}_d(r)$, and hence bound the distance between the distributions of $d_{\min}$ and $w_{\min}$. Let us start by recording a moment growth estimate for $\widetilde{Z}_d$:

Lemma IV.2. *We have*

$$\left(\mathbb{E} \widetilde{Z}_d^\ell\right)^{1/\ell} \leq C_{IV.2} \begin{cases} \frac{q^k \rho_d}{q-1}, & \text{if } \ell \leq \frac{q^k \rho_d}{q-1}, \\ \frac{\ell}{\log(e\ell(q-1)/(q^k \rho_d))}, & \text{if } \ell \geq \frac{q^k \rho_d}{q-1}. \end{cases}$$

Here, $C_{IV.2} > 0$ is a universal constant.

Next, fix an integer parameter $h \geq 1$ (its value will be defined later), and define the $h \times h$ square matrix $B = (b_{ij})$ as

$$b_{ij} = j^i, \quad i, j = 1, \ldots, h.$$

Then B is a Vandermonde matrix and the next lemma can be easily checked by a straightforward computation.

Lemma IV.3. *Let $B = (b_{ij})$ be as above. Then B is invertible, and the entries of the inverse matrix $B^{-1} = (b'_{ij})$ are given by*

$$b'_{ij} = \begin{cases} \dfrac{(-1)^{j-1} \sum\limits_{\substack{1 \leq m_1 < \cdots < m_{h-j} \leq h, \\ m_1, \ldots, m_{h-j} \neq i}} m_1 \ldots m_{h-j}}{i \prod\limits_{1 \leq m \leq h,\, m \neq i} (m - i)}, & \text{if } j < h; \\ \dfrac{1}{i \prod\limits_{1 \leq m \leq h,\, m \neq i} (i - m)}, & \text{if } j = h. \end{cases}$$

In what follows, we will not need a precise formula for the entries of the inverse; just a crude upper bound will be sufficient:

Corollary IV.4. *With the above notation, we have*

$$|b'_{ij}| \le \frac{\binom{h}{j} h^{h-j}}{((\lfloor h/2 \rfloor - 1)!)^2} \le C^h_{IV.4} h^{-j},$$

where $C_{IV.4} > 0$ is a universal constant.

Denote the vector $(M_d(1), \dots, M_d(h))^\top$ by V, and the vector $(\widetilde{M}_d(1), \dots, \widetilde{M}_d(h))^\top$ by $\widetilde{V}$. Further, let $U := (\mathbb{E}Z_d, \dots, \mathbb{E}Z_d^h)^\top$, and $\widetilde{U} := (\mathbb{E}\widetilde{Z}_d, \dots, \mathbb{E}\widetilde{Z}_d^h)^\top$, and, finally, define the "error vectors"

$$E := \Big(\sum_{r=h+1}^{\infty} r^i M_d(r) \Big)_{i=1}^h, \quad \widetilde{E} := \Big(\sum_{r=h+1}^{\infty} r^i \widetilde{M}_d(r) \Big)_{i=1}^h.$$

In view of the above,

$$BV + E = U, \quad B\widetilde{V} + \widetilde{E} = \widetilde{U},$$

whence the difference $V - \widetilde{V}$ can be expressed as

$$V - \widetilde{V} = B^{-1}(U - \widetilde{U}) - B^{-1}(E - \widetilde{E}).$$

The following lemma gives the estimate of $B^{-1}(U - \widetilde{U})$:

Lemma IV.5. *Suppose $d, n \in \mathbb{N}$ satisfy $\frac{d}{n} \le \lambda_0(1 - \frac{1}{q})$, and $d^2/n^{3/2} \ge C_{III.1}(\lambda_0, q)$. Assume additionally that $h \ge q^k \rho_d \ge \exp\big(-\frac{c_{III.1}(\lambda_0,q)d^4}{n^3 h}\big)$, $h \log_2 C_{IV.4} + h \log_2 C_{IV.2} + h + h \log(hp - h) \le k/4$ and $h \le \frac{c_{III.1}(\lambda_0,q)}{\log_2 C_{IV.4} + \log_2 C_{IV.2} + 2 + \log(q-1)} \frac{d^2}{n^{3/2}}$. Then the absolute value of each component of the vector $B^{-1}(U - \widetilde{U})$ is bounded above by*

$$O\bigg(\exp\Big(-\frac{1}{2} c_{III.1}(\lambda_0, q) d^4/n^3 \Big) + 2^{-k/4} \bigg).$$

By a slightly more careful argument, we get an estimate on the term $B^{-1}(E - \widetilde{E})$, which concludes the proof of IV.1.

Lemma IV.6. *Suppose $d, n \in \mathbb{N}$ satisfy $\frac{d}{n} \le \lambda_0(1 - \frac{1}{q})$, and $d^2/n^{3/2} \ge C_{III.1}(\lambda_0, q)$. Assume additionally that*

$$e^{-8C_{IV.2}C_{IV.4}(q-1)} h \ge q^k \rho_d \ge \exp\bigg(-\frac{c_{III.1}(\lambda_0, q) d^4}{4n^3 h} \bigg),$$

and $h \le \frac{c_{III.1}(\lambda_0,q)}{4} \frac{d^2}{n^{3/2}}$. Then

$$\sum_{r=h+1}^{\infty} M_d(r), \sum_{r=h+1}^{\infty} \widetilde{M}_d(r) = O(2^{-h}),$$

and the absolute value of each component of the vector $B^{-1}(E - \widetilde{E})$ is bounded above by $O(2^{-h})$.

Finally, we consider the improvement of the Gilbert–Varshamov bound implied by our argument. We shall state the result in a probabilistic form:

Corollary IV.7. *Let q be a prime power and $\alpha \in (0, \frac{1}{2})$. There exists constants $c, C > 0$ depending on q and α such that, for a sufficiently large integer n and $\alpha n \le d \le (1-\alpha)(1 - \frac{1}{q})n$, with probability greater than $\exp(-c\sqrt{n})$, a uniform random $\lfloor k + \frac{1}{2}\log_q(n) - C \rfloor$–dimensional linear code has the minimum distance at least d where k is the largest integer such that*

$$\frac{1}{q} \frac{q^n}{\sum_{j=0}^{d-1} \binom{n}{j} (q-1)^j} < q^k \le \frac{q^n}{\sum_{j=0}^{d-1} \binom{n}{j} (q-1)^j}.$$

(i.e. the dimension in Gilbert–Varshamov's bound)

Proof. Notice that k is the largest integer satisfying $q^k \rho_{d-1} \le 1$. The Gilbert–Varshamov result states that there exists a k–dimensional linear code with the minimum distance at least d.

Let $t \ge 0$ be a positive integer which we will determine later. Further, let $w_{\min}$ be the minimum weight of $\frac{q^{k+t}-1}{q-1}$ i.i.d. random vectors uniformly distributed over $\mathbb{F}_q^n$, and let $d_{\min}$ be the minimum distance of the uniform random $(k+t)$–dimensional linear code in $\mathbb{F}_q^n$. We have

$$\begin{aligned} \mathbb{P}\{w_{\min} \ge d\} &= (1 - \rho_{d-1})^{\frac{q^{k+t}-1}{q-1}} \\ &\ge \exp\Big(-2\rho_{d-1} \frac{q^{k+t}-1}{q-1} \Big) \\ &\ge \exp(-2\rho_{d-1} q^{k+t}) \\ &\ge \exp(-2q^t). \end{aligned}$$

Recall the q-ary entropy function

$$H_q(x) = x \log_q(q-1) - x \log_q(x) - (1-x)\log_q(1-x)$$

which appears in the Gilbert–Varshamov bound. It is a monotone increasing function on $(0, 1 - \frac{1}{q})$ with $H_q(0) = 0$ and $H_q(1) = 1$. Furthermore, for $x \in (1, 1 - \frac{1}{q})$,

$$\begin{aligned} H_q(x) &= \frac{1}{n} \log_q \left(\sum_{i=0}^{xn} \binom{n}{i} (q-1)^i \right) + o(1) \\ &= \frac{1}{n} \log_q (\rho_{xn} q^n) + o(1) \end{aligned}$$

whenever xn is an integer. (See [15, Proposition 3.3.1])

With $q^k \rho_d \le 1 < q^{k+1}\rho_d$, we have

$$H_q\Big(\frac{d}{n}\Big) = 1 - \frac{k}{n} + o(1).$$

Therefore, there exist $0 < R_1 < R_2 < 1$ depending only on q, α such that

$$R_1 \le \frac{k}{n} \le R_2.$$

Now we apply Theorem I.1 to get

$$\begin{aligned} \mathbb{P}\left\{ d_{\min} \ge d + t \right\} &\ge \mathbb{P}\left\{ w_{\min} \ge d + t \right\} \\ &\ge \exp(-2q^t) - \exp(-c_{\alpha,q}\sqrt{n}) \end{aligned}$$

where $c_{\alpha,q} = c(R_1, R_2, q)$. Choosing

$$t = \frac{1}{2} \log_q n + \log_q\big(\frac{c_{\alpha,q}}{4}\big)$$

we obtain the desired bound.

□

It is not difficult to check that the above corollary implies Corollary I.1 from the introduction.

References

[1] Y. I. Manin, “What is the maximum number of points on a curve over F_2,” *Journal of the Faculty of Science, the University of Tokyo. Sect. 1 A, Mathematics*, vol. 28, no. 3, pp. 715–720, 1982.

[2] M. Aaltonen, “Notes on the asymptotic behavior of the information rate of block codes (corresp.),” *IEEE Transactions on Information Theory*, vol. 30, no. 1, pp. 84–85, 1984.

[3] V. D. Goppa, “Bounds for codes,” *Dokl. Acad. Nauk.*, vol. 333, 1993.

[4] T. Jiang and A. Vardy, “Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes,” *IEEE Transactions on Information Theory*, vol. 50, no. 8, pp. 1655–1664, Aug 2004.

[5] D. Joyner and J.-L. Kim, *Selected Unsolved Problems in Coding Theory*. Springer Science & Business Media, 2011.

[6] R. McEliece, E. Rodemich, H. Rumsey, and L. Welch, “New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities,” *IEEE Transactions on Information Theory*, vol. 23, no. 2, pp. 157–166, 1977.

[7] M. A. Tsfasman, S. G. VIăduţ, and T. Zink, “Modular curves, shimura curves, and goppa codes, better than Varshamov-Gilbert bound,” *Mathematische Nachrichten*, vol. 109, no. 1, pp. 21–28, 1982. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mana.19821090103

[8] E. N. Gilbert, “A comparison of signalling alphabets,” *Bell System Technical Journal*, vol. 31, no. 3, pp. 504–522, may 1952.

[9] R. R. Varshamov, “Estimate of the number of signals in error correcting codes,” *Dokl. Acad. Nauk SSSR*, vol. 117, pp. 739–741, 1957.

[10] R. Gallagar, *Low-Density Parity-Check Codes*. MIT Press, 1963.

[11] A. Barg and G. D. Forney, “Random codes minimum distances and error exponents,” *IEEE Transactions on Information Theory*, vol. 48, no. 9, 2002.

[12] N. Linial and J. Mosheiff, “On the weight distribution of random binary linear codes,” *Random Structures & Algorithms*, vol. 56, no. 1, pp. 5–36, 2020.

[13] V. Vu and Lei Wu, “Improving the Gilbert-Varshamov bound for q-ary codes,” *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3200–3208, Sep. 2005.

[14] P. Gaborit and G. Zemor, “Asymptotic improvement of the Gilbert-Varshamov bound for linear codes,” *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 3865–3872, Sep. 2008.

[15] V. Guruswami, A. Rudra, and M. Sudan, *Essential Coding Theory*. Draft, Mar 15,2019.