Research Article

Márton Erdélyi, Pál Hegedüs, Sándor Z. Kiss, and Gábor P. Nagy*

On linear codes with random multiplier vectors and the maximum trace dimension property

https://doi.org/10.1515/jmc-2023-0022
received September 04, 2023; accepted October 31, 2023

Abstract: Let C be a linear code of length n and dimension k over the finite field $\mathbb{F}_{q^m}$. The trace code $\mathrm{Tr}(C)$ is a linear code of the same length n over the subfield $\mathbb{F}_q$. The obvious upper bound for the dimension of the trace code over $\mathbb{F}_q$ is mk. If equality holds, then we say that C has maximum trace dimension. The problem of finding the true dimension of trace codes and their duals is relevant for the size of the public key of various code-based cryptographic protocols. Let $C_{\boldsymbol{a}}$ denote the code obtained from C and a multiplier vector $\boldsymbol{a} \in (\mathbb{F}_{q^m})^n$. In this study, we give a lower bound for the probability that a random multiplier vector produces a code $C_{\boldsymbol{a}}$ of maximum trace dimension. We give an interpretation of the bound for the class of algebraic geometry codes in terms of the degree of the defining divisor. The bound explains the experimental fact that random alternant codes have minimal dimension. Our bound holds whenever $n \geq m(k + h)$, where $h \geq 0$ is the Singleton defect of C. For the extremal case $n = m(h + k)$, numerical experiments reveal a closed connection between the probability of having maximum trace dimension and the probability that a random matrix has full rank.

Keywords: trace codes, subfield subcodes, dimension of trace codes, random alternant codes, weight enumerator, Singleton defect

MSC 2020: 14G50, 15A03

1 Introduction

1.1 Code-based post-quantum cryptosystems

Recent research has focused extensively on quantum computers that use quantum mechanical techniques to solve difficult mathematical computational problems [1]. The existence of these potent devices poses a threat to numerous widely used public-key cryptosystems [2]. McEliece [3] introduced the first code-based public-key cryptosystem in 1978. One of the most pressing problems in cryptography today is to reduce the key size and enhance the security level of the McEliece cryptosystem, which is a promising cryptographic scheme for the post-quantum era [4]. Error-correcting codes used in code-based cryptographic protocols must be decoded

*** Corresponding author: Gábor P. Nagy,** Department of Algebra and Geometry, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary; Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H-6720 Szeged, Hungary, e-mail: nagyg@math.u-szeged.hu
Márton Erdélyi: Department of Algebra and Geometry, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary, e-mail: merdelyi@math.bme.hu
Pál Hegedüs: Department of Algebra and Geometry, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary, e-mail: hegpal@math.bme.hu
Sándor Z. Kiss: Department of Algebra and Geometry, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary, e-mail: ksandor@math.bme.hu

with efficient algorithms. The family of algebraic geometry (AG) codes and their subcodes and subfield subcodes constitute a rich class of such codes. These include the generalized Reed–Solomon, alternant, binary Goppa, and BCH codes. For a survey on decoding AG codes, see the research by Høholdt et al. [5].

Couvreur et al. [6–8] provided polynomial-time attacks against the McEliece cryptosystem that employs AG codes or their subcodes. In general, evaluation codes do not operate like random codes. This enables a wide variety of attacks against the McEliece cryptosystem based on AG codes. The technique described by Couvreur et al. [7,8] is inspired by the so-called *filtration attacks* that rely on computing the dimension of the Schur product that makes AG codes distinguishable from random ones. This observation was used by Wieschebrink [9] to provide an attack against the McEliece scheme based on subcodes of generalized Reed–Solomon codes [10]. Numerous attacks have employed a combination of powerful techniques, such as the filtration method, an error-correcting pair (ECP), or an error-correcting array (ECA), leading to a key recovery attack or a blind reconstruction of a decoding algorithm [7,8,11]. These vulnerabilities are based on the operation of the *Schur product* and a thorough examination of the dimensions of the Schur products for specific subcodes.

1.2 Key generation of code-based cryptosystems

The key generation process of the code-based scheme starts with a public code C_0 and a decoding algorithm Δ_0 which can efficiently correct a certain number of errors. Then, a random seed σ and a procedure Π are used to construct a code C with a decoding algorithm Δ.

Roughly speaking, the code $C = \Pi(C_0, \sigma)$ represents the public key, while the decoding algorithm $\Delta = \Pi(\Delta_0, \sigma)$ represents the private key. The class of *random alternant codes,* where the starting code C_0 is the full support Reed–Solomon code of dimension k over the field of order q^m, serves as an illustration. The random seed consists of a pair of vectors of length n over $\mathbb{F}_{q^m}$: the *multiplier* $\boldsymbol{a} = (a_1, \ldots, a_n)$, $(a_i \neq 0)$, and the *support* $\boldsymbol{x} = (x_1, \ldots, x_n)$, $(x_i \neq x_j \text{ for } i \neq j)$. The process Π has two main steps: first, compute the generalized Reed–Solomon code $C_1 = \mathbf{GRS}_k(\boldsymbol{x}, \boldsymbol{a})$, then compute the subfield subcode $\mathcal{A}_k(\boldsymbol{x}, \boldsymbol{a}) = C_1^{\perp} \cap \mathbb{F}_q^n$ of the dual of C_1. Due to Delsarte's theorem, the second step is equivalent to taking the dual of the trace code: $\mathcal{A}_k(\boldsymbol{x}, \boldsymbol{a}) = \mathrm{Tr}(C_1)^{\perp}$. (For more precise definitions and references, see Section 2).

While binary Goppa codes form a subclass of alternant codes, randomness for binary Goppa codes operates distinctly. One starts with the full support Reed–Solomon code C_0, where $q = 2$ and $k = 2t$. The seed consists of the support $\boldsymbol{x}$, and the monic irreducible polynomial $g(X)$ of degree t over $\mathbb{F}_{q^m}$. The multiplier $\boldsymbol{a}$ is defined by $a_i = 1/g(x_i)$, and the result is the alternant code $\mathcal{A}_k(\boldsymbol{x}, \boldsymbol{a})$. In both cases, the scheme's cryptographic strength depends on taking the subfield subcode or, equivalently, taking the trace code. Existing known mathematical techniques have yet failed to grasp the essence of these two operations. In particular, it is difficult to determine the true dimension of subfield subcodes and trace codes in general.

1.3 Random trace codes and their dimension

Subfield subcodes and trace codes are linked by duality. This study deals with the dimension problem of trace codes. Let q be a prime power, and m, k, n and positive integers. We extend the trace map $\mathrm{Tr}:\mathbb{F}_{q^m} \to \mathbb{F}_q$ to vectors and matrices. For a linear subspace $C \leq \mathbb{F}_{q^m}^n$, we write $\mathrm{Tr}(C) = \{\mathrm{Tr}(\boldsymbol{x}) | \boldsymbol{x} \in C\}$. For the linear code $C \leq \mathbb{F}_{q^m}^n$ of dimension k, we have the obvious upper bound $\dim(\mathrm{Tr}(C)) \leq mk$, and we say that C has *maximum trace dimension,* if the equality holds. Assume that the $\mathbb{F}_q$-linear code $C = \Pi(C_0, \sigma)$ is constructed using a $\mathbb{F}_{q^m}$-linear code C_0 and a random seed σ. Then, we may inquire about the probability

$$\mathrm{Prob}(C = \Pi(C_0, \sigma) \text{has maximum trace dimension}),$$

a value which depends solely on C_0. This probability has already been estimated using numerical experimentation for binary Goppa codes of the classic McEliece scheme (see Sections 2.2.2 and 4.2 of [12]), and for random alternant codes [13].

The focus on this probability is mainly theoretical; however, bounds on the proportion of random alternant codes with maximum trace dimension are beneficial in understanding the complexity of the algorithms used in the key generation process of code-based cryptography, as well as the size of public keys.

In this study, we prove a lower bound for the probability of maximum trace dimension in the probability model of random multipliers.

Theorem 1. *Let C be an $[n, k, d]_{q^m}$-code and let $h = n + 1 - k - d$ be its Singleton defect. Let P_C denote the proportion of multiplier vectors $\boldsymbol{a} = (a_1, \ldots, a_n) \in (\mathbb{F}_{q^m}^*)^n$ such that the linear code*

$$C_{\boldsymbol{a}} = \{(a_1x_1, \ldots, a_nx_n) | \boldsymbol{x} \in C\}$$

has maximum trace dimension. Then,

$$P_C \geq 1 - \frac{1 - q^{-m(h+k)}}{(q-1)q^{n-m(h+k)}}. \tag{1}$$

In particular, if $n \geq m(k + h)$, or equivalently $d \geq n(1 - 1/m) + 1$, then $P_C > 0$.

Our proof uses double counting methods that involve the weight distribution of the dual code $C^\perp$. We apply recent results of studies by Meneghetti et al. [14] that relate the weight distribution to numerical properties of the code that can be computed if the Singleton defect is small. For our purposes, the most important property is the number of $k \times v$ submatrices of rank r of the generator matrix.

Except for the case $q = 2$ and $n = m(h + k)$, Theorem 1 implies $P_C \geq 1/2$. This means that the Monte Carlo method of generating a random code $C_{\boldsymbol{a}}$ of maximum trace dimension is very effective. For $q = 2$ and $n = m(h + k)$, further research is necessary.

If $n < mk$, then clearly $\dim(\mathrm{Tr}(C)) \leq n < m \dim(C) = mk$, so C cannot be of maximum trace dimension for any C. Moreover, if C is an maximum distance separable (MDS) code of length $n - h$ extended with zeros in the last coordinates, then it is easy to see that $\dim(\mathrm{Tr}(C_{\boldsymbol{a}})) \leq n - h$. Thus, one might ask for the proportion of multiplier vectors for which $\dim(\mathrm{Tr}(C_{\boldsymbol{a}}))$ is close to the largest possible value n.

Theorem 2. *Let C be an $[n, k, d]_{q^m}$-code and let $h = n + 1 - k - d$ be its Singleton defect. Let P'_C denote the proportion of multiplier vectors $\boldsymbol{a} = (a_1, \ldots, a_n) \in (\mathbb{F}_{q^m}^*)^n$ such that $\dim(\mathrm{Tr}(C_{\boldsymbol{a}})) \geq n - mh$. Then,*

$$P'_C \geq \frac{q^{mh+1} - q^{mh} - q^{n-mk} + q^{-mk}}{q^{mh+1} - 1}. \tag{2}$$

In particular, if $n \leq m(k + h)$, or equivalently $d \leq n(1 - 1/m) + 1$, then $P'_C > 0$.

If $h = 0$ (thus C is MDS) and $n \leq mk$, then the formula in the above theorem obtains simpler and more similar to the one in Theorem 1:

$$P'_C \geq 1 - \frac{1 - q^{-n}}{(q-1)q^{mk-n}}. \tag{3}$$

1.4 Maximum trace dimension probabilities of AG codes

AG codes are linear error-correcting codes constructed from algebraic curves over finite fields, generalizing the Reed–Solomon code concept. They are defined by *evaluating functions* or by using *residues of differentials*. Their parameters can be derived from well-known theorems in AG. Our notation and terminology on algebraic plane curves over finite fields, their function fields, divisors, and Riemann–Roch spaces are conventional (see, for example, [15]).

Let $\mathcal{X}$ be a smooth algebraic curve over the finite field $\mathbb{F}_{q^m}$. Let $P_1, P_2, \ldots, P_n$ be pairwise distinct places of $\mathcal{X}$, and D is the divisor $D = P_1 + \cdots + P_n$. Let G be another divisor with support disjoint from D. The Riemann–Roch

theorem enables us to estimate the dimension, the minimum distance and the Singleton defect of AG codes. These, together with Theorem 1, imply a lower bound for the probability of maximum trace dimension of AG codes.

Theorem 3. *Let $C = C_L(D, G)$ be a functional AG code of length $n = \deg(D)$ over the finite field $\mathbb{F}_{q^m}$, $m > 1$. If $\deg(G) \le n/m - 1$, then*

$$P_C \ge 1 - \frac{1 - q^{-m(\deg(G)+1)}}{(q-1)q^{n-m(\deg(G)+1)}}. \tag{4}$$

1.5 Rank properties of random matrices in other probabilistic models

The rank properties of random matrices over finite fields have been extensively studied as a problem in combinatorial graph theory and other contexts, including coding theory and code-based post-quantum cryptography. For the probabilistic paradigm, there are a variety of alternatives [16–20]. One possibility is to choose each entry of the matrix independently and uniformly at random from the field. This can be extended to non-uniform distributions, which may or may not depend on the matrix entry's position. Studholme and Blake [20] studied *windowed random matrices,* where the nonzero elements of each column are restricted to fall within a window of length w, beginning at a randomly chosen row. Salmond et al. [18] proved that the probability that a random matrix has full rank cannot increase if we fix any number of additional elements to be identically zero.

Let A be an $n \times n$ matrix over the finite field $\mathbb{F}_q$, whose entries are chosen uniformly at random. As $n \to \infty$, the probability that A has rank n converges very fast to the value

$$S(q) = \prod_{i=1}^{\infty}\left(1 - \frac{1}{q^i}\right). \tag{5}$$

$S(q)$, which is independent of n, is also called the q-Pochhamer symbol $(1/q; 1/q)_\infty$, [21]. For $q = 2$, a good estimate for $S(2)$ is 0.2888. Let V be an $\mathbb{F}_q$-space of dimension n, and take n nonzero vectors uniformly at random from $V\setminus\{0\}$. The probability that the vectors are linearly independent also converges to $S(q)$ very fast if $n \to \infty$.

We performed numerical experiments for Reed–Solomon codes $C = \mathbf{RS}_k(\boldsymbol{x})$ over $\mathbb{F}_{q^m}$, where k, m are positive integers, $q = 2$ or $q = 3$, and $x_1, \ldots, x_{km}$ are random distinct elements of $\mathbb{F}_{q^m}$. Therefore, C has a length $n = km$, and Singleton defect $h = 0$. We observed that the probability that C has maximum trace dimension is near to the value $S(q)$.

1.6 Outline of the article

Notation and classical prerequisites on linear codes are given in Section 2. Section 3 collects basic properties and examples of codes which have maximum trace dimension. In Section 4, we deal with the dimension problem of random alternant codes. Sections 5 and 6 contain detailed proofs of the main theorems. The basic concepts of AG codes are also presented in Section 6.

2 Prerequisites from coding theory

Let q be a prime power and let m, n, k be positive integers such that $mk \le n \le q^m$. Let $x_1, \ldots, x_n$ be distinct elements of $\mathbb{F}_{q^m}$. The *Reed–Solomon code* $\mathbf{RS}_k(\boldsymbol{x})$ is defined by the generator matrix

$$G = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{k-1} & x_2^{k-1} & \cdots & x_n^{k-1} \end{bmatrix}. \tag{6}$$

The vector $\boldsymbol{x}$ is called the *support* of the Reed–Solomon code. $\mathbf{RS}_k(\boldsymbol{x})$ has dimension k and minimum distance $d = n - k + 1$. It is an MDS code with Singleton defect $h = 0$. Let $a_1, \ldots, a_n$ be nonzero elements of $\mathbb{F}_{q^m}$. The *generalized Reed–Solomon code* $\mathbf{GRS}_k(\boldsymbol{x}, \boldsymbol{a})$ has generator matrix

$$G' = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ a_1x_1 & a_2x_2 & \cdots & a_nx_n \\ \vdots & \vdots & & \vdots \\ a_1x_1^{k-1} & a_2x_2^{k-1} & \cdots & a_nx_n^{k-1} \end{bmatrix}.$$

Clearly, $\mathbf{GRS}_k(\boldsymbol{x}, \boldsymbol{a})$ and $\mathbf{RS}_k(\boldsymbol{x})$ have same parameters. In particular, generalized Reed–Solomon codes are MDS. If $n = q^m$, then $\{x_1, \ldots, x_n\} = \mathbb{F}_{q^m}$ and the codes are said to have *full support.* The dual code of $\mathbf{GRS}_k(\boldsymbol{x}, \boldsymbol{a})$ is again a generalized Reed–Solomon code $\mathbf{GRS}_{n-k}(\boldsymbol{x}, \boldsymbol{b})$, with the same support $\boldsymbol{x}$. The Berlekamp–Massey algorithm provides an efficient decoding algorithm for Reed–Solomon codes, which can correct up to $\lfloor \frac{d-1}{2} \rfloor = \lfloor \frac{n-k}{2} \rfloor$ errors. If the multiplier vector is given, then this algorithm can also be used to decode generalized Reed–Solomon codes.

Let C be a linear code of length n, dimension k, and minimum distance d, defined over the finite field $\mathbb{F}_{q^m}$. The *subfield subcode* or *restricted code* of C is

$$C|_{\mathbb{F}_q} = C \cap \mathbb{F}_q^n.$$

We extend the trace map $\mathrm{Tr}:\mathbb{F}_{q^m} \to \mathbb{F}_q$ to vectors and matrices entry-wise. We define the *trace code* of the linear $C \leq \mathbb{F}_{q^m}^n$ by

$$\mathrm{Tr}(C) = \{\mathrm{Tr}(\boldsymbol{x}) | \boldsymbol{x} \in C\}.$$

Clearly, $\mathrm{Tr}(C)$ is an $\mathbb{F}_q$-linear code of length n. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ be a basis of C, and let $\beta_1, \ldots, \beta_m$ be a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$. Then, the vectors $\mathrm{Tr}(\beta_i \boldsymbol{x}_j)$, $(1 \leq i \leq m, 1 \leq j \leq k)$ span the trace code $\mathrm{Tr}(C)$. This implies the obvious upper bound $\dim(\mathrm{Tr}(C)) \leq km$ for the dimension of the trace code. We say that C has *maximum trace dimension,* if

$$\dim_{\mathbb{F}_q}(\mathrm{Tr}(C)) = m \dim_{\mathbb{F}_{q^m}}(C).$$

According to Delsarte's theorem [22],

$$(\mathrm{Tr}(C))^\perp = (C^\perp)|_{\mathbb{F}_q},$$

which shows that subfield subcodes and trace codes are basically dual objects. This yields the obvious lower bound

$$\dim(C|_{\mathbb{F}_q}) \geq n - m(n - k)$$

for the dimension of the subfield subcode. The minimum distance of $C|_{\mathbb{F}_q}$ is at least the minimum distance of C. Moreover, subfield subcodes inherit the decoding algorithms of their parent code.

An *alternant code* is defined as the subfield subcode of a generalized Reed–Solomon code

$$\mathcal{A}_k(\boldsymbol{x}, \boldsymbol{a}) = (\mathbf{GRS}_k(\boldsymbol{x}, \boldsymbol{a})^\perp)|_{\mathbb{F}_q},$$

or equivalently, as the dual code of the trace code of a generalized Reed–Solomon code

$$\mathcal{A}_k(\boldsymbol{x}, \boldsymbol{a}) = \mathrm{Tr}(\mathbf{GRS}_k(\boldsymbol{x}, \boldsymbol{a}))^\perp.$$

The integer k is referred to as the *degree* of the alternant code, and m as its *extension degree.* The vector $\boldsymbol{x}$ is the *support*, and the vector $\boldsymbol{a}$ is the *multiplier* of the alternant code. In the sequel, even without explicitly saying it, we assume that the entries of the support vector are distinct, and the entries of the multiplier vector are different from zero.

The obvious lower bound for the dimension of the alternant code is

$$\dim(\mathcal{A}_k(\boldsymbol{x}, \boldsymbol{a})) \geq n - mk.$$

Given the support and the multiplier, the Berlekamp–Massey algorithm can correct up to $\left\lfloor \frac{k}{2} \right\rfloor$ errors for the alternant code $\mathcal{A}_k(\boldsymbol{x}, \boldsymbol{a})$.

Recall that the Schur product of the vectors $\mathbf{a} = (a_1, \ldots, a_n)$, $\mathbf{b} = (b_1, \ldots, b_n)$ is defined by

$$\mathbf{a} \star \mathbf{b} = (a_1 b_1, \ldots, a_n b_n).$$

3 The maximum trace dimension property

In this section, we prove a collection of properties of codes having the maximum trace dimension. At the end of the section, we present a class of examples which shows that Theorem 1 is close to being sharp asymptotically.

In the sequel, C denotes an $\mathbb{F}_{q^m}$-linear code of length n, dimension k, and minimum distance d.

Definition 4. Define the *support* of C as

$$\operatorname{supp}(C) = \{i \in \{1, 2, \ldots, n\} | \exists \boldsymbol{x} \in C : x_i \neq 0\}. \tag{7}$$

For an integer i, define

$$d_i(C) = \min_{\substack{D \leq C \\ \dim(D) = i}} |\operatorname{supp}(D)|. \tag{8}$$

Note that $d_1(C) = d$ is the minimum distance. Clearly, $\operatorname{supp}(C) = \operatorname{supp}(C_{\boldsymbol{a}})$ and $d_i(C) = d_i(C_{\boldsymbol{a}})$ for each multiplier vector $\boldsymbol{a}$. Furthermore, $\dim(C) \leq |\operatorname{supp}(C)|$.

The proofs of the following lemmas are straightforward consequences of the definitions.

Lemma 5. *The following are equivalent*:

(*i*) *The code C has maximum trace dimension.*

(*ii*) *All $\mathbb{F}_{q^m}$-linear subspaces of C have maximum trace dimension.*

(*iii*) *For all $\boldsymbol{x} \in C \setminus \{\boldsymbol{0}\}$, $\operatorname{Tr}(\boldsymbol{x}) \neq \boldsymbol{0}$.*

(*iv*) *$C \cap K = \{\boldsymbol{0}\}$, where K is the kernel of the trace map $\operatorname{Tr}: \mathbb{F}_{q^m}^n \to \mathbb{F}_q^n$.*

Lemma 6. *Assume that for some multiplier vector $\boldsymbol{a}$, $C_{\boldsymbol{a}}$ has maximum trace dimension. Then, we have $d_i(C) \geq im$ for all $1 \leq i \leq k$.*

Proof. If $D \leq C$, $\dim(D) = i$ such that $|\operatorname{supp}(D)| < im$, then

$$\operatorname{supp}(\operatorname{Tr}(D_{\boldsymbol{a}})) \subseteq \operatorname{supp}(D_{\boldsymbol{a}}) = \operatorname{supp}(D)$$

and

$$\dim(\operatorname{Tr}(D_{\boldsymbol{a}})) \leq |\operatorname{supp}(\operatorname{Tr}(D_{\boldsymbol{a}}))| = |\operatorname{supp}(\operatorname{Tr}(D))| < im = m \dim(D) = m \dim(D_{\boldsymbol{a}}).$$

Therefore, $D_{\boldsymbol{a}}$ and $C_{\boldsymbol{a}}$ have no maximum trace dimension. □

We conjecture that the converse of Lemma 6 holds as well.

As the following examples show, the proportion of multiplier vectors producing a maximum trace dimension code is related to the probability of a random matrix to have full rank. Let A be an $n \times n$ matrix whose entries are chosen from $\mathbb{F}_q$ uniformly at random. The probability for A to have maximum rank n is

$$S_1(n, q) = \prod_{i=1}^{n} \left(1 - \frac{1}{q^i}\right). \tag{9}$$

As $n \to \infty$, $S_1(n, q)$ converges very fast to the value

$$S(q) = \prod_{i=1}^{\infty}\left(1 - \frac{1}{q^i}\right). \tag{10}$$

Let V be an $\mathbb{F}_q$-space of dimension n, and take n nonzero vectors uniformly at random from $V\backslash\{0\}$. The probability for the vectors to be linearly independent is

$$S_2(n, q) = \prod_{j=0}^{n-1}\frac{q^n - q^j}{q^n - 1} = \prod_{i=1}^{n}\left(1 - \frac{1}{q^i}\right)\left(1 + \frac{1}{q^n - 1}\right)^n. \tag{11}$$

As the last factor converges to 1 very fast, $S_2(n, q) \to S(q)$ very fast (Figures 1 and 2). In fact, if $n > 20$, then $S(q)$ is a good practical approximation for $S_1(n, q)$ and $S_2(n, q)$.

Lemma 7. *Let C be the m-fold repetition code over $\mathbb{F}_{q^m}$. The probability that $C_{\boldsymbol{a}}$ has maximum trace dimension for a random multiplier vector $\boldsymbol{a}$ is $S_2(m, q)$. In practice, if $m \geq 20$, then $S(q)$ is a good approximation for this probability.*

Let C_i be linear $[n_i, k_i]_{q^m}$-codes, $i = 1, 2$. Their sum $C_1 + C_2$ is a linear $[n_1 + n_2, k_1 + k_2]_{q^m}$-code whose codewords are $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ with $\boldsymbol{x}_i \in C_i$. The minimum distance of the sum is $d(C_1 + C_2) = \min(d(C_1), d(C_2))$.

Lemma 8. *Let C, C' be $\mathbb{F}_{q^m}$-linear codes, and $D = C + C'$ their sum. Then, $P_D = P_C P_{C'}$, where P_C is as defined in Theorem* 1.

Let C be the k-fold sum of the m-fold repetition code. Clearly, C has length $n = mk$, dimension k, and minimum distance $d = m$. The last two lemmas imply that the proportion P_C of multiplier vectors with maximum trace dimension is approximately $P_C \approx S(q)^k$, which tends to zero if $k \to \infty$. In particular, we cannot expect P_C to be close to 1 just because k and m are large. However, $P_C > 0$, so there is a multiplier $\boldsymbol{a}$ such that $C_{\boldsymbol{a}}$ has the maximum trace dimension. On the other hand, $d_i(C) = im$ for all $1 \leq i \leq k$, showing that Lemma 6 is sharp.

Clearly, if $n < mk$, then $P_C = 0$. In Theorem 1, we see that $n \geq m(h + k)$ implies $P_C > 0$. The question whether P_C is zero or not is open for the interval $[mk, m(h + k) - 1]$. The following class of examples has Singleton defect $h \approx \log_q(k)$, hence the interval is small. Still, the condition $n = mk$ is not enough to ensure $P_C > 0$. In other words, Theorem 1 is close to being sharp asymptotically.

Proposition 9. *For all prime power q and integers $m > 2$, $2 \leq k \leq q^m/m$, there exists an $\mathbb{F}_{q^m}$-linear code $C' = C'(q, m, k)$ of length $n = mk$, dimension k, and Singleton bound $h = m$, such that $P_{C'} = 0$.*

n	$S_1(n,q)$	$S_2(n,q)$	$S_1(n,q) - S(q)$
5	0.298004150390625	0.349271971075915	0.00921605530402259
10	0.289070298419749	0.291908472309700	0.000282203333146547
15	0.288796908379162	0.288929141393520	8.81329255975061e-6
20	0.288788370496567	0.288793878752760	2.75409964223261e-7
25	0.288788103693158	0.288788318857146	8.60655607892724e-9
30	0.288788095355557	0.288788103424204	2.68954858384518e-10
35	0.288788095095007	0.288788095389177	8.40483238562229e-12
40	0.288788095086865	0.288788095097371	2.62623256475081e-13
45	0.288788095086611	0.288788095086980	8.16013923099490e-15
50	0.288788095086603	0.288788095086615	2.22044604925031e-16

Figure 1: $q = 2$.

n	$S_1(n,q)$	$S_2(n,q)$	$S_1(n,q) - S(q)$
5	**0.561280381843718**	**0.572973321315295**	**0.00115430391576976**
10	**0.560130820850226**	**0.560225688332595**	**4.74292227792272e-6**
15	**0.560126097446024**	**0.560126682988612**	**1.95180757112112e-8**
20	**0.560126078008270**	**0.560126081221122**	**8.03216382294636e-11**
25	**0.560126077928279**	**0.560126077944806**	**3.30735439035834e-13**
30	**0.560126077927950**	**0.560126077928031**	**1.44328993201270e-15**

Figure 2: $q = 3$.

Proof. Let $n' = m(k-1)-1$ and let $x_1, \ldots, x_n$ be distinct elements of $\mathbb{F}_{q^m}$ such that $x_{n'+1}, \ldots, x_n \neq 0$. Let $C' = C'(q, m, k)$ be the code with generator matrix

$$G' = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ x_1 & x_2 & \cdots & x_{n'} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_1^{k-2} & x_2^{k-2} & \cdots & x_{n'}^{k-2} & 0 & \cdots & 0 \\ x_1^{k-1} & x_2^{k-1} & \cdots & x_{n'}^{k-1} & x_{n'+1}^{k-1} & \cdots & x_n^{k-1} \end{bmatrix}.$$

Let D' be the subcode generated by the first $k-1$ rows. As $k-1 \leq n'$, we have $\dim(D') = k-1$. Moreover, D' has support $\{1, 2, \ldots, n'\}$, hence $|\mathrm{supp}(D')| = n' = m(k-1)-1 < m \dim(D')$. Lemma 6 implies that $P_{C'} = 0$.

Now we compute the minimum distance of C'. Take any linear combination $\boldsymbol{c}$ of the rows of G'. Write $\boldsymbol{c}' = (c_1, c_2, \ldots, c_{n'})$ and $\boldsymbol{x}' = (x_1, x_2, \ldots, x_{n'})$. If the last row has zero coefficient, then the last $m+1$ coordinates are 0 and $\boldsymbol{c}' \in \mathbf{RS}_{k-1}(\boldsymbol{x}')$. So

$$\mathrm{wt}(\boldsymbol{c}) \geq n' - (k-1) + 1 = (m-1)(k-1),$$

and equality occurs for some $\boldsymbol{c}$. If the last row has nonzero coefficient, then the last $m+1$ coordinates of $\boldsymbol{c}$ are nonzero and $\boldsymbol{c}' \in \mathbf{RS}_k(\boldsymbol{x}')$. So

$$\mathrm{wt}(\boldsymbol{c}) \geq n' - k + 1 + m + 1 > (m-1)(k-1).$$

Thus, the minimal distance of $C'(m,k)$ is indeed $d = (m-1)(k-1)$ and $h = m$. □

4 The dimension of random alternant codes

In numerical experiments, one observes that the dimension of random alternant codes typically attains the obvious lower bound [13]. In this short section, we derive a proof for this observation from Theorem 1. We show that if the length of the random alternant code exceeds mk, then the dimension is $n - mk$ with high probability. In particular, this is the case for most random alternant codes of full support.

Definition 10. Given the field of definition $\mathbb{F}_q$, the degree k, and the extension degree m, the random alternant code is a code $\mathcal{A}_k(\boldsymbol{x}, \boldsymbol{a})$, where the support $\boldsymbol{x}$ and the multiplier $\boldsymbol{a}$ are chosen uniformly at random.

Proposition 11. *Let q be a prime power and m, n, k be positive integers such that $mk \leq n \leq q^m$. The random alternant code of length n, degree k, extension degree m over $\mathbb{F}_q$ has dimension $n - mk$ with probability at least*

$$1 - \frac{1 - q^{-mk}}{(q-1)q^{n-mk}}.$$

Proof. The dual of the alternant code is $\mathrm{Tr}(\mathbf{GRS}_k(\boldsymbol{x}, \boldsymbol{a}))$. Since $\mathbf{GRS}_k(\boldsymbol{x}, \boldsymbol{a})$ is MDS of dimension k, Theorem 3 implies the proposition. □

5 Proof of Theorems 1 and 2

In this section, we use the notation of Theorem 1. We describe the average cardinality of $\mathrm{Tr}(C_{\boldsymbol{a}})^{\perp}$, $\boldsymbol{a} \in (\mathbb{F}_{q^m}^*)^n$, with the help of the weight distribution of the dual code $C^{\perp}$. Let us introduce the following notation:

Definition 12. Let $\mathrm{wt} : C \to \mathbb{N}$ denote the Hamming weight and

$$B_w = |\{\boldsymbol{c} \in C^{\perp} | \mathrm{wt}(c) = w\}| \tag{12}$$

for $0 \le w \le n$ the weight distribution of C. Then, let

$$\lambda(C) = \sum_{w=0}^{n} B_w \left(\frac{q-1}{q^m - 1} \right)^w. \tag{13}$$

For $0 \le r \le v \le n$, let

$$N_G(v, r) = |\{k \times v \text{ submatrices of } G \text{ with rank } r\}|, \tag{14}$$

where $G \in \mathbb{F}_{q^m}^{k \times n}$ is a generator matrix of C.

Proposition 13. *We have the following average form*:

$$\lambda(C) = \frac{1}{|(\mathbb{F}_{q^m}^*)^n|} \sum_{\boldsymbol{a} \in (\mathbb{F}_{q^m}^*)^n} q^{n - \dim(\mathrm{Tr}(C_{\boldsymbol{a}}))}. \tag{15}$$

Proof. For $\boldsymbol{a} \in (\mathbb{F}_{q^m}^*)^n$, we write $\boldsymbol{a}^{-1} = (a_j^{-1})_{1 \le j \le n}$. We double-count the set

$$H = \{(\boldsymbol{a}, \boldsymbol{c}) | \boldsymbol{a}^{-1} \star \boldsymbol{c} \in \mathbb{F}_q^n, \boldsymbol{a} \in (\mathbb{F}_{q^m}^*)^n, \boldsymbol{c} \in C^{\perp}\}. \tag{16}$$

For any fixed $\boldsymbol{a}$, $(\boldsymbol{a}, \boldsymbol{c}) \in H$ if and only if $\boldsymbol{a}^{-1} \star \boldsymbol{c} \in (C^{\perp})_{\boldsymbol{a}^{-1}} \cap \mathbb{F}_q^n$. By Delsarte's theorem [22, Theorem 2], we have

$$(C^{\perp})_{\boldsymbol{a}^{-1}} \cap \mathbb{F}_q^n = (C_{\boldsymbol{a}})^{\perp} \cap \mathbb{F}_q^n = (\mathrm{Tr}(C_{\boldsymbol{a}}))^{\perp}. \tag{17}$$

Hence, $|(C^{\perp})_{\boldsymbol{a}^{-1}} \cap \mathbb{F}_q^n| = q^{n - \dim(\mathrm{Tr}(C_{\boldsymbol{a}}))}$. This proves

$$|H| = \sum_{\boldsymbol{a} \in (\mathbb{F}_{q^m}^*)^n} q^{n - \dim(\mathrm{Tr}(C_{\boldsymbol{a}}))}. \tag{18}$$

Let us now fix $\boldsymbol{c} \in C^{\perp}$. For each j, we have

$$\{a_j \in \mathbb{F}_{q^m}^* | a_j^{-1} c_j \in \mathbb{F}_q\} = \begin{cases} \mathbb{F}_{q^m}^*, & \text{if } c_j = 0; \\ c_j \mathbb{F}_q^*, & \text{if } c_j \neq 0. \end{cases} \tag{19}$$

Thus,

$$|\{\boldsymbol{a} \in (\mathbb{F}_{q^m}^*)^n | \boldsymbol{a}^{-1} \star \boldsymbol{c} \in \mathbb{F}_q^n\}| = (q-1)^{\mathrm{wt}(\boldsymbol{c})} (q^m - 1)^{n - \mathrm{wt}(\boldsymbol{c})}, \tag{20}$$

summing over all $\boldsymbol{c} \in C^{\perp}$, we obtain $|H| = (q^m - 1)^n \lambda(C)$. □

As $\dim(\mathrm{Tr}(C_{\boldsymbol{a}})) \le km$, each summand on the right-hand side of (15) is at least q^{n-km}. This gives a lower bound

$$\lambda(C) \ge q^{n-km}. \tag{21}$$

The upper bounds of $\lambda(C)$ can be used to find lower bounds on the proportion P_C of multiplier vectors which produce maximum trace dimension codes and on the proportion P_C' of multiplier vectors which produce trace codes with dimension at least $n - mh$.

Proposition 14.
(1) *Assume* $\lambda(C) \le q^{n-km} + E$, *where* E *is nonnegative. Then,*

$$P_C \ge 1 - \frac{E}{(q-1)q^{n-km}}. \tag{22}$$

(2) *Assume* $\lambda(C) \le q^{mh+1} - E'$. *Then,*

$$P'_C \ge \frac{E'}{q^{mh+1}-1}. \tag{23}$$

Proof. If $C_{\boldsymbol{a}}$ does not have maximum trace dimension, then the corresponding summand in (15) is at least q^{n-km+1}. Therefore, $P_C q^{n-km} + (1-P_C)q^{n-km+1} \le \lambda(C)$. The first claim follows from a straightforward computation.

In a similar manner, if $\dim(\mathrm{Tr}(C_{\boldsymbol{a}})) < n - mh$, then the corresponding summand in (15) is at least q^{mh+1}; otherwise, it is at least 1. Therefore, $P'_C + (1-P'_C)q^{mh+1} \le \lambda(C)$, hence the second claim. □

Proposition 15.

$$\lambda(C) = \left(\frac{q^m-q}{q^m-1}\right)^n \sum_{v=0}^{n} \left(\frac{q-1}{q^m-q}\right)^v \sum_{r=0}^{v} N_G(v,r) q^{m(v-r)}, \tag{24}$$

where $N_G(v,r)$ *is as defined in* (14).

Proof. Applying Proposition 3 of [14] for $C^\perp$ over $\mathbb{F}_{q^m}$, we obtain

$$\sum_{s=0}^{v} \binom{n-s}{v-s} B_s = \sum_{r=0}^{v} N_G(v,r) q^{m(v-r)}. \tag{25}$$

Multiplying with x^v and summing over $0 \le v \le n$, we obtain

$$\sum_{v=0}^{n} x^v \sum_{s=0}^{v} \binom{n-s}{v-s} B_s = \sum_{v=0}^{n} x^v \sum_{r=0}^{v} N_G(v,r) q^{m(v-r)}. \tag{26}$$

Changing the order of the summation and using the binomial theorem, the left hand side is

$$\sum_{s=0}^{n} B_s x^s \sum_{v=s}^{n} \binom{n-s}{v-s} x^{v-s} = \sum_{s=0}^{n} B_s x^s (1+x)^{n-s}. \tag{27}$$

Let us put $x = \frac{q-1}{q^m-q}$, thus $1 + x = \frac{q^m-1}{q^m-q}$. Then, by the definition,

$$\lambda(C) = \sum_{s=0}^{n} B_s \frac{(q-1)^s (q^m-1)^{n-s}}{(q^m-1)^n} = \left(\frac{q^m-q}{q^m-1}\right)^n \sum_{s=0}^{n} B_s x^s (1+x)^{n-s}. \tag{28}$$

By (26), we have

$$\lambda(C) = \left(\frac{q^m-q}{q^m-1}\right)^n \sum_{v=0}^{n} x^v \sum_{r=0}^{v} N_G(v,r) q^{m(v-r)}, \tag{29}$$

hence the proposition. □

Proof of Theorems 1 and 2. Applying Lemma 4 in [14] for $C^\perp$, all $k \times (k+h)$ submatrix of G has rank k. It follows that the rank of all $k \times v$ submatrix equals k if $v \ge k+h$ and is at least $v-h$ if $v < k+h$.

By using this observation, we can bound the inner sum on the right hand side of the previous proposition:
– For $v \ge k+h$, we have $N_G(v,r) = 0$ for $r < k$ and

$$\sum_{r=0}^{v} N_G(v,r) q^{m(v-r)} = \binom{n}{v} q^{m(v-k)}, \tag{30}$$

– for $v < k + h$, we have $N_G(v,r) = 0$ for $r < v - h$ and

$$\sum_{r=0}^{v} N_G(v,r) q^{m(v-r)} \le \binom{n}{v} q^{mh}. \tag{31}$$

In view of (30), (31), and $x = \frac{q-1}{q^m - q}$, we obtain

$$\begin{aligned}
\lambda(C) &\le \left(\frac{q^m - q}{q^m - 1}\right)^n \left(\sum_{v=0}^{k+h-1} \binom{n}{v} x^v q^{mh} + \sum_{v=k+h}^{n} \binom{n}{v} x^v q^{m(v-k)} \right) \\
&= \frac{1}{q^{mk}} \left(\frac{q^m - q}{q^m - 1}\right)^n \left(\sum_{v=0}^{n} \binom{n}{v} (xq^m)^v + \sum_{v=0}^{k+h-1} \binom{n}{v} x^v (q^{m(h+k)} - q^{mv}) \right) \\
&\le \frac{1}{q^{mk}} \left(\frac{q^m - q}{q^m - 1}\right)^n \left((1 + xq^m)^n + (q^{m(h+k)} - 1) \sum_{v=0}^{k+h} \binom{n}{v} x^v \right) \\
&\le \frac{1}{q^{mk}} \left(\frac{q^m - q}{q^m - 1}\right)^n ((1 + xq^m)^n + (q^{m(h+k)} - 1)(1 + x)^n).
\end{aligned}$$

As $1 + x = \frac{q^m - 1}{q^m - q}$ and $1 + xq^m = q \cdot \frac{q^m - 1}{q^m - q}$, we obtain

$$\lambda(C) \le q^{n-mk} + (q^{mh} - q^{-mk}). \tag{32}$$

Set $E = q^{mh} - q^{-mk}$ and

$$E' = q^{mh+1} - (q^{n-mk} + q^{mh} - q^{-mk}).$$

By using Proposition 14, we obtain the lower bounds on P_C and P_C' as in the statements. If $n \le m(k + h)$, then $E' \ge (q - 2)q^{mh} + q^{-mk} > 0$, thus $P_C' > 0$. □

6 Proof of Theorem 3

AG codes are linear error-correcting codes constructed from algebraic curves over finite fields, generalizing the Reed–Solomon code concept. They are defined by *evaluating functions* or by using *residues of differentials*. Their parameters can be derived from well-known AG theorems. Our notation and terminology on algebraic plane curves over finite fields, their function fields, divisors, and Riemann–Roch spaces are conventional (see, for example, [15]).

Let $\mathcal{X}$ be an algebraic curve, that is, an affine or projective variety of dimension one, which is absolutely irreducible and non-singular and whose defining equations are (homogeneous) polynomials with coefficients in $\mathbb{F}_q$. Let $g = g(\mathcal{X})$ be the genus of $\mathcal{X}$. $\mathbb{F}_q(\mathcal{X})$ denotes the function field of $\mathcal{X}$. A *divisor* D of $\mathcal{X}$ is a formal sum $D = n_1P_1 + \cdots + n_kP_k$, where $n_1, \ldots, n_k \in \mathbb{Z}$ and $P_1, \ldots, P_k$ are places of $\mathbb{F}_q(\mathcal{X})$. If $n_1, \ldots, n_k \ge 0$, then $D \ge 0$. If D, E are two divisors and $D - E \ge 0$, then $D \ge E$. In the case of a nonzero function f of the function field $\mathbb{F}_q(\mathcal{X})$, and a place P, $v_P(f)$ stands for the order of f at P. If $v_P(f) > 0$, then P is a zero of f, while if $v_P(f) < 0$, then P is a pole of f with multiplicity $-v_P(f)$. The *principal divisor* of a nonzero function f is $\mathrm{Div}(f) = \sum_P v_P(f)P$.

For a divisor D, the associated Riemann–Roch space $\mathcal{L}(D)$ is the vector space

$$\mathcal{L}(D) = \{f \in \mathbb{F}_q(\mathcal{X}) \backslash \{0\} | \mathrm{Div}(f) \ge -D\} \cup \{0\}.$$

The dimension $\ell(D)$ of $\mathcal{L}(D)$ is given by the Riemann–Roch Theorem [15, Theorem 1.1.15]:

$$\ell(D) = \ell(W - D) + \deg D - g + 1,$$

where W is a canonical divisor. We denote the set of *differentials* on $\mathcal{X}$ by Ω. The *differential space* of the divisor D is

$$\Omega(D) = \{dh \in \Omega | \mathrm{Div}(dh) \geqslant A\} \cup \{0\}.$$

In the following, $P_1, P_2, \ldots, P_n$ are pairwise distinct places on $\mathcal{X}$, and D is the divisor $D = P_1 + \cdots + P_n$. Let G be another divisor with support disjoint from D. We define two types of AG codes, the *functional* and the *differential codes,* respectively:

$$C_L(D, G) = \{(f(P_1), \ldots, f(P_n)) | f \in \mathcal{L}(G)\},$$
$$C_\Omega(D, G) = \{(\mathrm{res}_{P_1}(\omega), \ldots, \mathrm{res}_{P_n}(\omega)) | \omega \in \Omega(G - D)\}.$$

These codes are dual to each other, and $C_\Omega(D, G) = C_L(D, K + D - G)$ for a well-chosen canonical divisor K. The Riemann–Roch theorem enables us to estimate the dimension and the minimum distance of AG codes:

$$\dim(C_L(D, G)) \begin{cases} \geq \deg(G) - g + 1 & 0 \leq \deg(G) \leq 2g - 2, \\ = \deg(G) - g + 1 & 2g - 2 \leq \deg(G) \leq n, \\ \leq \deg(G) - g + 1 & n \leq \deg(G) \leq n + 2g - 2. \end{cases}$$

The minimum distance of a functional code is at least its *designed minimum distance*

$$\delta_L = n - \deg(G).$$

Proof of Theorem 3. Let k be the dimension, and h be the Singleton defect of the AG code $C = C_L(D, G)$. Then $h + k = n + 1 - d \leq n + 1 - \delta_L = \deg(G) + 1$. As the right hand side of (1) is monotone decreasing in $h + k$, the formula (4) follows. □

7 Conclusion

We gave a lower bound for the probability that the dimension of the trace code of a linear code with a random multiplier vector attains the obvious upper bound. This is exactly the type of question that requires solid mathematical understanding for McEliece-type cryptographic protocols. Our formula only uses the size of the underlying field, the degree of the field extension, and the three main parameters of the code: length, dimension, and minimum distance. The result provided a concise formula for the probability that an AG code has maximum trace dimension. We also proved by mathematical means that full support random alternant codes have dimension $n - mk$ with high probability. These pieces of information are useful to understand better the complexity of Monte Carlo algorithms in the key generation process of code-based cryptosystems. This provides insights into the practicality and performance of the cryptosystem in real-world applications, in particular in resource limited devices like sensor nodes or smart cards.

Our approach works for the probabilistic model of random multiplier vectors. Random Goppa codes have a different probability paradigm. Therefore, our results do not solve the dimension problem for random Goppa codes. This needs further research, but we are optimistic that our method can be extended.

Acknowledgments: This research was supported by the Ministry of Culture and Innovation and the National Research, Development, and Innovation Office within the Quantum Information National Laboratory of Hungary (Grant No. 2022-2.1.1-NL-2022-00004), and partially funded by NKFIH Grants K129335, K138596, K135885, and FK127906. This work has been accepted for presentation at CIFRIS23, the Congress of the Italian association of cryptography “De Componendis Cifris.”

Conflict of interest: Authors state no conflict of interest.

References

[1] Arute F, Arya K, Babbush R, Bacon D, Bardin JC, Barends R, et al. Quantum supremacy using a programmable superconducting processor. Nature. 2019;574(7779):505–10. doi: https://doi.org/10.1038/s41586-019-1666-5.

[2] Shor PW. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM J Comput. 1997;26(5):1484–509. doi: https://doi.org/10.1137/S0097539795293172.

[3] McEliece RJ. A public-key cryptosystem based on algebraic coding theory. DSN Progress Report, 42–44:114–116, 1978.

[4] National Institute of Standards, Technology. Post-Quantum Cryptography; Updated: March 25. 2020. http://csrc.nist.gov/projects/post-quantum-cryptography.

[5] Høholdt T, Van Lint JH, Pellikaan R. Algebraic geometry codes. Handbook of coding theory. 1998;1(Part 1):871–961.

[6] Couvreur A, Márquez-Corbella I, Pellikaan R. Cryptanalysis of public-key cryptosystems that use subcodes of algebraic geometry codes. In: Coding theory and applications. Cham: Springer; 2015. p. 133–40.

[7] Couvreur A, Márquez-Corbella I, Pellikaan R. Cryptanalysis of McEliece cryptosystem based on algebraic geometry codes and their subcodes. IEEE Trans Inform Theory. 2017;63(8):5404–18. doi: https://doi.org/10.1109/TIT.2017.2712636.

[8] Couvreur A, Otmani A, Tillich JP. Polynomial time attack on wild McEliece over quadratic extensions. IEEE Trans Inform Theory. 2016;63(1):404–27.

[9] Wieschebrink C. Cryptanalysis of the Niederreiter public key scheme based on GRS subcodes. In: International Workshop on Post-Quantum Cryptography. Springer; 2010. p. 61–72.

[10] Berger TP, Loidreau P. How to mask the structure of codes for a cryptographic use. Des Code Cryptogr. 2005;35(1):63–79.

[11] Couvreur A, Gaborit P, Gauthier-Umannna V, Otmani A, Tillich JP. Distinguisher-based attacks on public-key cryptosystems using Reed–Solomon codes. Des Code Cryptogr. 2014;73(2):641–66.

[12] Albrecht MR, Bernstein DJ, Chou T, Cid C, Gilcher J, Lange T, et al. Classic McEliece: conservative code-based cryptography; 2020. https://classic.mceliece.org/nist/mceliece-20201010.pdf.

[13] Mora R, Tillich JP. On the dimension and structure of the square of the dual of a Goppa code. Des Codes Cryptogr. 2023;91(4):1351–72. doi: https://doi.org/10.1007/s10623-022-01153-w.

[14] Meneghetti A, Pellegrini M, Sala M. A formula on the weight distribution of linear codes with applications to AMDS codes. Finite Fields Appl. 2022;77:Paper No. 101933, 15. doi: https://doi.org/10.1016/j.ffa.2021.101933.

[15] Stichtenoth H. Algebraic function fields and codes. Vol. 254 of Graduate Texts in Mathematics. 2nd edn. Berlin: Springer-Verlag; 2009.

[16] Cooper C. On the distribution of rank of a random matrix over a finite field. In: Proceedings of the Ninth International Conference “Random Structures and Algorithms” (Poznan, 1999). Vol. 17. 2000. p. 197–212. doi: https://doi.org/10.1002/1098-2418(200010/12)17:3/4<197::AID-RSA2>3.3.CO;2-B.

[17] Cooper C. On the rank of random matrices. Random Struct Algorithms. 2000;16(2):209–32. doi: https://doi.org/10.1002/(SICI)1098-2418(200003)16:2<209::AID-RSA6>3.3.CO;2-T.

[18] Salmond D, Grant A, Grivell I, Chan T. On the rank of random matrices over finite fields; 2016.

[19] Studholme C, Blake IF. Properties of random matrices and applications; 2006. http://www.cs.toronto.edu/ cvs/coding/random_report.pdf.

[20] Studholme C, Blake IF. Random matrices and codes for the erasure channel. Algorithmica. 2010;56(4):605–20. doi: https://doi.org/10.1007/s00453-008-9192-0.

[21] Wikipedia contributors. Q-Pochhammer symbol – Wikipedia, The Free Encyclopedia; 2022. [Online; accessed 27-January-2023]. https://en.wikipedia.org/w/index.php?title=Q-Pochhammer_symbololdid=1109461763.

[22] Delsartre P. On Subfield Subcodes of Modified Reed–Solomon Codes. IEEE Trans Inform Theory. 1975;21(5):575–6.