Finite Fields and Their Applications 15 (2009) 497–516

Contents lists available at ScienceDirect

Finite Fields and Their Applications

www.elsevier.com/locate/ffa

Correlations between the ranks of submatrices and weights of random codes

Alexander A. Klyachko, İbrahim Özen *

Department of Mathematics, Bilkent University, 06800 Ankara, Turkey

ARTICLE INFO

Article history:
Received 18 August 2006
Revised 22 March 2008
Available online 8 April 2009
Communicated by Vera Pless

Keywords:
Random codes
Random matrices
Rank
Weight enumerator of a random code
Cumulant
Classification of subspaces

ABSTRACT

The results of our study are twofold. From the random matrix theory point of view we obtain results on the rank distribution of column submatrices. We give the moments and the covariances between the ranks ($q^{-\mathrm{rank}}$) of such submatrices. We conjecture the counterparts of these results for arbitrary submatrices. The case of higher correlations gets drastically complicated even in the case of three submatrices. We give a formula for the correlation of ranks of three submatrices and a conjecture for its closed form. From the code theoretical point of view our study yields the covariances of the coefficients of the weight enumerator of a random code. Particularly interesting is that the coefficients of the weight enumerator of a code with random parity check matrix are uncorrelated. We give a conjecture for the triple correlations between the coefficients of the weight enumerator of a random code.

1. Introduction

Random codes are closely related with random matrices over finite fields. Specifically, parameters of random codes depend on distribution and correlations between the ranks of submatrices. This paper studies ranks of random matrices for code theoretical applications.

We start with the basic definitions of linear codes. One can find this material and more on codes, for example in [9] and [12]. A *linear code* C is a linear subspace of $\mathbb{F}_q^n$ where $\mathbb{F}_q$ is a finite field of q elements. The number n is called *the length* and the dimension k of C is called the *number of information symbols* of the code. The number of nonzero coordinates of a code vector e is said to be

* Corresponding author.
E-mail addresses: klyachko@fen.bilkent.edu.tr (A.A. Klyachko), iozen@fen.bilkent.edu.tr (İ. Özen).

1071-5797/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ffa.2009.03.002

the *weight* $|e|$ of e. The minimum of the weights over nonzero code vectors is the *minimum distance* of the code and it is usually denoted by d. We call a code with parameters n, k, d over $\mathbb{F}_q$ an $[n,k,d]_q$ code.

If we consider $\mathbb{F}_q^n$ as a linear space with the standard scalar product $\langle u, v\rangle = \sum_i u_i v_i$, then the orthogonal complement $C^\perp$ of C is called the *dual code* to C.

Let C be an $[n,k]_q$ code and let G be a $k \times n$ matrix whose rows form a basis of C. Then any element $e \in C$ is a linear combination of these row vectors. We call G a *generator matrix* of the code C. Any vector $e' \in C^\perp$ is orthogonal to the basis vectors of C and the product $G \times (e')^T$ is the zero vector, this gives the criterion for lying in $C^\perp$. The matrix G is called a *parity check matrix* of the code $C^\perp$.

Let $C \subset \mathbb{F}_q^n$ be a code and let A_i be the number of code vectors with exactly i nonzero coordinates

$$A_i = \left|\left\{e \in C\colon\ |e| = i\right\}\right|.$$

The set of A_is form the *weight set* of the code C. We define the *weight enumerator* $W_C(t)$ of the code C as the polynomial

$$W_C(t) = \sum_i A_i t^{n-i}. \tag{1}$$

Let C be a code with generator matrix G. We find the following form of the weight enumerator vital for our purposes (see [11])

$$W_C(1+t) = \sum_{I \subset \{1,2,\dots,n\}} q^{k-r_I} t^{|I|}, \tag{2}$$

where r_I is the rank of the column submatrix spanned by the column set I.

All the parameters of codes can be extracted from the weight enumerator $W_C(t)$ which is given by the *rank function* q^{-r_I}, $I \subset \{1,2,\dots,n\}$ of the generator matrix. A code can be considered as a configuration of points (columns of the generator matrix) in the projective space. Then the existence of a code with given weights turns out to be equivalent to existence of a configuration with given rank function as we see in Eqs. (1) and (2). But this is a classical wild problem and it can be arbitrarily complicated [10]. So we pass to the statistical approach rather than trying to determine the explicit structure.

For a matrix G, *we take randomness in the sense that the entries are independent and they are uniformly distributed along the field* $\mathbb{F}_q$. One must keep in mind that there are two codes that are attached to a random matrix. Once we are given a $k \times n$ matrix G there is the $[n,k]_q$ code which assumes G as its generator matrix; and second, the code assuming G as its parity check matrix. Both codes will be referred to as *random codes*. We denote the random code in the first sense by C and the code in the second sense by $C^\perp$. Their weight enumerators will be referred as $W_C(t) = \sum_i A_i t^{n-i}$ and $W_{C^\perp}(t) = \sum_i A_i^\perp t^{n-i}$ respectively.

There is a major difference between the asymptotic behaviors of ranks of square and rectangular matrices. One can easily deduce from Eq. (7) that the probability of a square $n \times n$ matrix over $\mathbb{F}_q$ to be singular is positive even when $n \to \infty$. It is also clear from the same formula that as long as $R = k/n$ is kept away from 1, a $k \times n$ matrix is almost sure to have rank k. Hence we will tacitly assume maximality of the rank when codes of fixed $R < 1$ are aimed.

The main results of this paper are explicit formulas for the correlation functions between the ranks of up to three submatrices. Correlations between the weights of a random code are deduced using these results. In particular we find out that the coefficients of the weight enumerator of a random code are uncorrelated.

We list the main results as follows. The sth moment of the rank function $q^{-r_{k\times n}}$ of a $k \times n$ random matrix is given by

$$\mu_s\big(q^{-r_{k\times n}}\big)=q^{-kn}\sum_{0\leqslant r\leqslant k}q^{r(n-s)}\begin{bmatrix}k\\r\end{bmatrix}\prod_{0\leqslant i<r}\big(1-q^{-n+i}\big),$$

$$\mu_s\big(q^{-r_{k\times n}}\big)=\prod_{0\leqslant i<k}\big(q^{-n}+q^{-s}-q^{-n-s+i}\big),$$

where the product in the second formula is noncommutative in $x=q^{-n}$ and $y=q^{-s}$ with the commutation rule $yx=qxy$. The moment $\mu_s(q^{-r_{k\times n}})$ is symmetric in n and k. Furthermore if s is a nonnegative integer then it is also symmetric in s. For example the expectation of $q^{-r_{k\times n}}$ is given by

$$\mathbb{E}\big(q^{-r_{k\times n}}\big)=q^{-n}+q^{-k}-q^{-n-k}.$$

The covariance between the rank functions q^{-r_I} and q^{-r_J} of two column submatrices is given by

$$\operatorname{cov}\big(q^{-r_I},q^{-r_J}\big)=\frac{(q-1)(q^k-1)(q^{|I\cap J|}-1)}{q^{2k+|I|+|J|}}.$$

We also provide a conjecture for the covariance between the ranks of two submatrices spanned by arbitrary row and column sets.

Next we show that the covariances between the coefficients of weight enumerator of the random code $C^\perp$ are given by

$$\begin{aligned}\operatorname{cov}\big(W_{C^\perp}(x),W_{C^\perp}(y)\big)&=\sum_{i,j}\operatorname{cov}\big(A_i^\perp,A_j^\perp\big)x^{n-i}y^{n-j}\\&=\frac{(q-1)(q^{n-k}-1)}{q^{2(n-k)}}\big((xy+q-1)^n-(xy)^n\big).\end{aligned}$$

An important corollary to this theorem is that coefficients of weight enumerator of the code $C^\perp$ are uncorrelated. This result is not valid for the code C given by a random generator matrix.

We also studied the triple correlations between the ranks of three submatrices. Using the classification of triplets of subspaces we managed to deduce an extremely complicated formula. To our surprise computer experiments show beyond reasonable doubt that the formula is equivalent to a simple closed one. It is even simpler when we switch from the correlations to the cumulants. We conjecture that the closed formula of the joint cumulant of ranks of three submatrices is given by

$$\begin{aligned}\sigma\big(q^{-r_I},q^{-r_J},q^{-r_K}\big)=\frac{(q-1)^2(q^k-1)}{q^{3k+I+J+K}}&\big(q^{IJ+JK+KI-IJK}-\big(q^{IJ}+q^{JK}+q^{KI}\big)\\&+2+\big(q^k-q\big)\big(q^{IJK}-1\big)\big),\end{aligned}$$

where on the right-hand side of the equation the multiplicative notation for sets like IJ means the cardinality of their intersection $|I\cap J|$. We remind that the joint cumulant of a triplet of random variables is

$$\begin{aligned}\sigma(X,Y,Z)=\mathbb{E}(XYZ)-\mathbb{E}(X)\operatorname{cov}(Y,Z)-\mathbb{E}(Y)\operatorname{cov}(Z,X)\\-\mathbb{E}(Z)\operatorname{cov}(X,Y)-\mathbb{E}(X)\mathbb{E}(Y)\mathbb{E}(Z).\end{aligned}$$

This conjecture is equivalent to the following formula for joint cumulants of the coefficients of $W_{C^\perp}(t)$,

$$\begin{aligned}
&\sigma\big(W_{C^\perp}(x), W_{C^\perp}(y), W_{C^\perp}(z)\big)\\
&\quad=\sum_{i,j,l}\sigma(A_i^\perp, A_j^\perp, A_l^\perp)x^{n-i}y^{n-j}z^{n-l}\\
&\quad=q^{-3(n-k)}(q-1)^2(q^{n-k}-1)\big\{\big(xyz+(q-1)(x+y+z+q-2)\big)^n-(xyz)^n\\
&\qquad-\big((xy+q-1)^n-(xy)^n\big)z^n-\big((yz+q-1)^n-(yz)^n\big)x^n\\
&\qquad-\big((zx+q-1)^n-(zx)^n\big)y^n+\big(q^{n-k}-q\big)\big[(xyz+q-1)^n-(xyz)^n\big]\big\}.
\end{aligned}$$

For the random code C the corresponding equation is

$$\begin{aligned}
&\sigma\big(W_C(x), W_C(y), W_C(z)\big)\\
&\quad=(q-1)^2\big(q^k-1\big)\big\{q^{-n}(xyz+q-1)^n-q^{-2n}\big((xy+q-1)^n(z+q-1)^n+(yz+q-1)^n(x+q-1)^n\\
&\qquad+(zx+q-1)^n(y+q-1)^n\big)+2q^{-3n}(x+q-1)^n(y+q-1)^n(z+q-1)^n\\
&\qquad+\big(q^k-q\big)\big[q^{-2n}\big(xyz+(q-1)(x+y+z+q-2)\big)^n-q^{-3n}(x+q-1)^n(y+q-1)^n(z+q-1)^n\big]\big\}.
\end{aligned}$$

There exists a classification of quadruples of subspaces due to Gelfand and Ponomarev [6]. For further study of this subject, one may wonder whether it may lead to a closed formula for the quadruple correlations.

The above results on correlations imply, for example, that the roots of the weight enumerator $W_C(1+z)$ of a selfdual code $C=C^\perp$ are almost surely on the circle $|z|=\sqrt{q}$ provided $q>9$. Here by almost sure we mean that the probability to have a root off the circle tends to 0 as $n\to\infty$. Similar but less sharp results hold for all $[n,k]_q$ codes where $q>q_0(R)$, $R=k/n$. The proofs will be published elsewhere.

The paper is organized as follows. In Section 2 we evaluate the moments and covariances between the rank functions. Moments are given by Theorem 3. We derive the covariance between the ranks of two submatrices in Theorem 8. Using this, we show that the coefficients of weight enumerator of the random code $C^\perp$ are uncorrelated by Theorem 12 and Corollary 13.

In Section 3 we study the triple correlations between the ranks of three submatrices. A triplet of submatrices gives a triplet of subspaces. We review the classification of triplets of subspaces and obtain the number of triplets with given invariants. This enables us to obtain a formula for the triple correlation between the ranks of three submatrices which is given by Theorem 20. Conjecture 21 provides the closed form of the joint cumulant of three ranks. Equivalent to this conjecture we give Corollaries 23 and 24 on the joint cumulant of the coefficients of $W_C(t)$ and $W_{C^\perp}(t)$ respectively.

2. Moments of the rank function and the covariances

This section is devoted to calculation of the moments and covariance between the rank functions of column submatrices. We will derive the covariances between the coefficients of a random weight enumerator from these results.

2.1. Moments of the rank function

We will start with the moments of the random variable $q^{-r_{k\times n}}$ where $r_{k\times n}$ is the rank of a $k\times n$ matrix. The basic tool for evaluation of moments is the probability $P(n,k,r)$ of a random $k\times n$ matrix to have rank r.

We need q-analogues $[n]_q$ of positive integers n and the q-factorials to express this probability. Instead of the usual definition $[n]_q=(q^n-1)/(q-1)$ we use the following form

$$[n]_q:=q^n-1.$$

These functions extend the Binomial coefficients as follows

$$[0]_q! = 1,$$
$$[n]_q! = [n]_q[n-1]_q \dots [1]_q,$$
$$\begin{bmatrix} n \\ r \end{bmatrix}_q = \frac{[n]_q!}{[r]_q![n-r]_q!}. \tag{3}$$

The quantum Binomial coefficient in Eq. (3) gives the number of r-dimensional subspaces of an n-dimensional linear space over $\mathbb{F}_q$. The cardinality of the group $\mathrm{GL}(k,q)$ is given by the factorial

$$\bigl|\mathrm{GL}(k,q)\bigr| = q^{\frac{k(k-1)}{2}}[k]_q!. \tag{4}$$

In the following we drop the subscript q in the notation. We will need two classical formulae in the sequel:

q-Binomial theorem (*commutative*)

$$\prod_{0\leqslant j<d}\bigl(1-q^j t\bigr) = \sum_{0\leqslant i\leqslant d}(-1)^i q^{\frac{i(i-1)}{2}}\begin{bmatrix} d \\ i \end{bmatrix} t^i, \tag{5}$$

q-Binomial theorem (*noncommutative*)

$$(x+y)^d = \sum_{0\leqslant i\leqslant d}\begin{bmatrix} d \\ i \end{bmatrix} x^i y^{d-i}, \tag{6}$$

where the power on the left-hand side is noncommutative in x and y. The power is to be expanded so that every monomial is in the form $x^i y^j$ by the relation yx=qxy.

Note that the q-Binomial identities are expressed in the q-Binomial coefficients and our convention on $[n]_q$ does not change them. See [7] for the theory of the subject and proofs of the items above.

Proposition 1. *The probability $P(n,k,r)$ of a random $k \times n$ matrix to have rank r is given by*

$$P(n,k,r) = q^{-kn+\frac{r(r-1)}{2}}\frac{[n]!}{[n-r]!}\frac{[k]!}{[k-r]!}\frac{1}{[r]!}. \tag{7}$$

Proof. This is a standard formula, see for example [13]. □

In the next theorem we will use the following simple lemma.

Lemma 2.

$$\frac{[k]!}{[k-d]!} = q^{-\frac{d(d-1)}{2}}\sum_{0\leqslant i\leqslant d}(-1)^i q^{\frac{i(i-1)}{2}}\begin{bmatrix} d \\ i \end{bmatrix} q^{k(d-i)}. \tag{8}$$

Proof.

$$\frac{[k]!}{[k-d]!} = \prod_{0\leqslant i<d}\bigl(q^{k-i}-1\bigr) = q^{kd-\frac{d(d-1)}{2}}\prod_{0\leqslant i<d}\bigl(1-q^{i-k}\bigr).$$

When we replace the last product with the corresponding sum by Eq. (5) we get the result. □

The following theorem gives the moments of the rank function of a random matrix.

Theorem 3. *The sth moment $\mu_s(q^{-r_{k\times n}})$ of the rank function $q^{-r_{k\times n}}$ is given by*

$$\mu_s\bigl(q^{-r_{k\times n}}\bigr) = q^{-kn} \sum_{0\leqslant r\leqslant k} q^{r(n-s)} \begin{bmatrix} k \\ r \end{bmatrix} \prod_{0\leqslant i<r} \bigl(1-q^{-n+i}\bigr), \tag{9}$$

$$\mu_s\bigl(q^{-r_{k\times n}}\bigr) = \prod_{0\leqslant i<k} \bigl(q^{-n}+q^{-s}-q^{-n-s+i}\bigr), \tag{10}$$

where the product in the second formula is noncommutative in $x = q^{-n}$ and $y = q^{-s}$ with the commutation rule $yx = qxy$. The moment formula $\mu_s(q^{-r_{k\times n}})$ is symmetric in n and k. Furthermore if s is a nonnegative integer then it is also symmetric in s.

Proof. We use the formula in Eq. (7) for the probability.

$$\begin{aligned}
\mu_s\bigl(q^{-r_{k\times n}}\bigr) &= \sum_r q^{-rs} P(n,k,r) \\
&= \sum_r q^{-kn+\frac{r(r-1)}{2}-rs} \frac{[k]!}{[k-r]![r]!} \frac{[n]!}{[n-r]!} && (11) \\
&\underset{\text{Eq. (8)}}{=} q^{-kn} \sum_{r,i} q^{-rs} \begin{bmatrix} k \\ r \end{bmatrix} (-1)^i q^{\frac{i(i-1)}{2}} \begin{bmatrix} r \\ i \end{bmatrix} q^{n(r-i)} && (12) \\
&= q^{-kn} \sum_r q^{r(n-s)} \begin{bmatrix} k \\ r \end{bmatrix} \prod_{0\leqslant i<r} \bigl(1-q^{-n+i}\bigr). && (13)
\end{aligned}$$

To get the second formula of the moments we proceed from Eq. (12) as follows

$$\begin{aligned}
\mu_s\bigl(q^{-r_{k\times n}}\bigr) &= q^{-kn} \sum_r q^{-rs} \begin{bmatrix} k \\ r \end{bmatrix} \sum_i (-1)^i q^{\frac{i(i-1)}{2}} \begin{bmatrix} r \\ i \end{bmatrix} \bigl(q^n\bigr)^{r-i} \\
&\underset{r\to k-r}{=} q^{-kn} \sum_{r,i} q^{-s(k-r)} \begin{bmatrix} k \\ r \end{bmatrix} (-1)^i q^{\frac{i(i-1)}{2}} \begin{bmatrix} k-r \\ i \end{bmatrix} q^{n(k-r-i)} \\
&= \sum_{r,i} (-1)^i q^{\frac{i(i-1)}{2}} \begin{bmatrix} k \\ i \end{bmatrix} q^{i(-n-s)} \begin{bmatrix} k-i \\ r \end{bmatrix} \bigl(q^{-n}\bigr)^r \bigl(q^{-s}\bigr)^{k-i-r} \\
&= \sum_i (-1)^i q^{\frac{(i)(i-1)}{2}} \begin{bmatrix} k \\ i \end{bmatrix} \bigl(q^{-n-s}\bigr)^i \bigl(q^{-n}+q^{-s}\bigr)^{k-i}. && (14)
\end{aligned}$$

The last equation follows from the noncommutative q-Binomial formula (Eq. (6)). Finally by the commutative q-Binomial formula given in Eq. (5) we get

$$\mu_s\bigl(q^{-r_{k\times n}}\bigr) = \prod_{0\leqslant i<k} \bigl(q^{-n}+q^{-s}-q^{-n-s+i}\bigr),$$

where we remind that the product is noncommutative. It should be expanded in $x = q^{-n}$, $y = q^{-s}$ so that every monomial is put in the form $x^i y^j$ by the commutation relation $yx = qxy$.

The probability formula $P(n,k,r)$ is symmetric in n and k hence the same is true for the moment formulas. This can be proven by expanding $[k]!/[k-r]!$ instead of $[n]!/[n-r]!$ in Eq. (11). If s is a nonnegative integer the formula is completely symmetric in n, k and s. This follows from the symmetry of the noncommutative q-Binomial formula (6). □

Example 4. An easy calculation gives the expectation $\mathbb{E}$ and the second moment μ_2 of $q^{-r_{k\times n}}$ as follows

$$\mathbb{E}\big(q^{-r_{k\times n}}\big) = q^{-n} + q^{-k} - q^{-n-k}, \tag{15}$$

$$\mu_2\big(q^{-r_{k\times n}}\big) = q^{-2n} + q^{-2k} + q^{-2n-2k+1} + q^{-n-k}(q+1)\big(1 - q^{-n} - q^{-k}\big). \tag{16}$$

We can derive the expectations of the coefficients of a random weight enumerator from the expectation of the rank function. Expectations and the second moments of the coefficients can be found for example in [5, p. 10] and [1, p. 44].[1]

Example 5. Let G be a random $k \times n$ matrix and let C be the code with generator matrix G. Eq. (2) expresses the weight enumerator of the code C in terms of the rank functions. From there we see that the expectation of the weight enumerator of the code C is given by

$$\mathbb{E}\big(W_C(t)\big) = \sum_{I\subset\{1,2,\ldots n\}} q^k \mathbb{E}\big(q^{-r_I}\big)(t-1)^{|I|}.$$

We obtain the expectation of the weight enumerator when we substitute $\mathbb{E}(q^{-r_I}) = \mathbb{E}(q^{-r_{k\times|I|}})$ given by Eq. (15)

$$\mathbb{E}\big(W_C(t)\big) = \sum_i \mathbb{E}(A_i)t^{n-i} = t^n + \frac{(q^k-1)}{q^n}(t+q-1)^n. \tag{17}$$

Now making use of the well-known MacWilliams duality [8]

$$W_{C^\perp}(1+t) = q^{-k}t^n W_C\left(1 + \frac{q}{t}\right), \tag{18}$$

we obtain the analogous result for the random code given by a parity check matrix. Let $C^\perp$ be the random code whose parity check matrix is G. Taking expectations of both sides of Eq. (18) and using Eq. (17) we get

$$\mathbb{E}W_{C^\perp}(t) = \sum_i \mathbb{E}\big(A_i^\perp\big)t^{n-i} = q^{-k}\big((t+q-1)^n + \big(q^k-1\big)t^n\big). \tag{19}$$

2.2. The covariance between the ranks and codeweight correlations

The covariance between two rank functions is given by

$$\operatorname{cov}\big(q^{-r_I}, q^{-r_J}\big) = \mathbb{E}\big(q^{-r_I - r_J}\big) - \mathbb{E}\big(q^{-r_I}\big)\mathbb{E}\big(q^{-r_J}\big). \tag{20}$$

The new ingredient $\mathbb{E}(q^{-r_I-r_J})$ requires the joint probability $P(r_I, r_J)$ of ranks of the submatrices G_I and G_J spanned by the column sets I and J respectively. In order to express $P(r_I, r_J)$ we need an auxiliary function. Let G be a $k\times n$ matrix and let M fixed columns of G span a matrix of rank m. Denote by $P(n,k,M,m,r)$ the probability that G has rank r.

Lemma 6.

$$P(n,k,M,m,r) = P(n-M, k-m, r-m). \tag{21}$$

[1] We thank the anonymous reviewer who brought these references to our attention.

Proof. Let V_0 be the column space of the given M vectors. We have to find $n - M$ column vectors of rank $r - m$ in the space $\mathbb{F}_q^k / V_0$ and then lift them up to $\mathbb{F}_q^k$. So the number $N(n, k, M, m, r)$ of complementary matrices is given by

$$N(n, k, M, m, r) = q^{(n-M)(k-m)} P(n - M, k - m, r - m) q^{(n-M)m}.$$

To find the probability we multiply by $q^{-(n-M)k}$ and the exponents cancel. □

The following proposition gives the moments of this partial probability function.

Proposition 7. *Moments of the partial rank function are given by*

$$\sum_r q^{-rs} P(n, k, r - m) = q^{-sm} \mu_s\bigl(q^{-r_{k \times n}}\bigr) \tag{22}$$

$$= q^{-sm-kn} \sum_{0 \leqslant r \leqslant k} q^{r(n-s)} \begin{bmatrix} k \\ r \end{bmatrix} \prod_{0 \leqslant i < r} \bigl(1 - q^{-n+i}\bigr) \tag{23}$$

$$= q^{-sm} \prod_{0 \leqslant i < k} \bigl(q^{-n} + q^{-s} - q^{-n-s+i}\bigr), \tag{24}$$

where in the last formula the product is noncommutative in the variables $x = q^{-n}$ and $y = q^{-s}$. The moment formula is symmetric in n and k. Moreover when $m = 0$ and s is a nonnegative integer the formula is completely symmetric in n, k and s.

Proof. The proof of this proposition repeats the same steps as Theorem 3 so we skip the proof. □

Theorem 8. *Let G be a random $k \times n$ matrix and let $I, J \subset \{1, 2, \ldots, n\}$ be two column sets. Then the covariance between the rank functions q^{-r_I} and q^{-r_J} is given by*

$$\operatorname{cov}\bigl(q^{-r_I}, q^{-r_J}\bigr) = \frac{(q-1)(q^k-1)(q^{|I \cap J|}-1)}{q^{2k+|I|+|J|}}. \tag{25}$$

Proof. The covariance is given by

$$\operatorname{cov}\bigl(q^{-r_I}, q^{-r_J}\bigr) = \mathbb{E}\bigl(q^{-r_I - r_J}\bigr) - \mathbb{E}\bigl(q^{-r_I}\bigr)\mathbb{E}\bigl(q^{-r_J}\bigr). \tag{26}$$

So we need the expectation of $q^{-r_I - r_J}$, hence the joint probability $P(r_I, r_J)$ of the ranks r_I and r_J. This probability is given by the following sum

$$P(r_I, r_J) = \sum_r P(IJ, k, r) P(I, k, IJ, r, r_I) P(J, k, IJ, r, r_J),$$

where IJ is the shortcut notation for $|I \cap J|$. The sum runs over the column rank of the intersection. We extend the intersection to the sets I and J with the ranks r_I and r_J. Substituting for the auxiliary probabilities given by the last two factors we get

$$\mathbb{E}\bigl(q^{-r_I - r_J}\bigr) = \sum_{r, r_I, r_J} q^{-r_I - r_J} P(IJ, k, r) P(\bar{\bar{I}}, k - r, r_I - r) P(\bar{\bar{J}}, k - r, r_J - r),$$

where $\bar{\bar{I}} = |I \setminus J|$ and $\bar{\bar{J}} = |J \setminus I|$. We can sum up over r_I and r_J by Eqs. (15) and (22)

$$\mathbb{E}\big(q^{-r_I - r_J}\big) = \sum_r P(IJ, k, r)\big(q^{-k}\big(1 - q^{-\bar{\bar{I}}}\big) + q^{-r-\bar{\bar{I}}}\big)\big(q^{-k}\big(1 - q^{-\bar{\bar{J}}}\big) + q^{-r-\bar{\bar{J}}}\big).$$

Finally the summation over r gives

$$\begin{aligned}\mathbb{E}\big(q^{-r_I - r_J}\big) = {} & q^{-2k}\big(1 - q^{-\bar{\bar{I}}}\big)\big(1 - q^{-\bar{\bar{J}}}\big) + q^{-\bar{\bar{I}}-\bar{\bar{J}}}\mu_2\big(q^{-r_{k\times IJ}}\big) \\ & + \big(q^{-\bar{\bar{I}}-k}\big(1 - q^{-\bar{\bar{J}}}\big) + q^{-\bar{\bar{J}}-k}\big(1 - q^{-\bar{\bar{I}}}\big)\big)\mathbb{E}\big(q^{-r_{k\times IJ}}\big).\end{aligned}$$

We have expressed $\mathbb{E}(q^{-r_I - r_J})$ in terms of the first and second moments of the rank function and they are given by Eqs. (15) and (16). If we substitute them with the product $\mathbb{E}(q^{-r_I})\mathbb{E}(q^{-r_J})$ in Eq. (26) we get the claim. □

While the covariances between the rank functions of column submatrices are enough for the code theoretical purposes the same question about arbitrary submatrices is still an interesting problem. Let G be a random $k \times n$ matrix. Let $I \subset \{1, 2, \dots, n\}$ be a column set and $L \subset \{1, 2, \dots, k\}$ be a row set. We denote by r_{LI} the rank of the submatrix spanned by the rows in L and the columns in I.

Conjecture 9. *The covariance between the rank functions $q^{-r_{LI}}, q^{-r_{MJ}}$ of two submatrices is given by*

$$\operatorname{cov}\big(q^{-r_{LI}}, q^{-r_{MJ}}\big) = \frac{(q-1)(q^{|I\cap J|} - 1)(q^{|L\cap M|} - 1)}{q^{|I|+|J|+|L|+|M|}}.$$

Before we proceed with the covariance between the coefficients of the weight enumerators we need the following lemma.

Lemma 10. *Let*

$$S_{ij} = \sum_{\substack{I, J \subset \{1,2,\dots n\} \\ |I|=i, |J|=j}} q^{|I\cap J|}. \tag{27}$$

We have the following generating function for S_{ij}

$$\sum_{i,j} S_{ij} x^i y^j = (1 + x + y + qxy)^n. \tag{28}$$

Proof. Set the cardinality of $I \cap J$ to m. We can choose the elements in the symmetric difference and then split this set to I and J. This gives

$$S_{ij} = \sum_m \binom{n}{m,\ j-m,\ i-m,\ n+m-i-j} q^m.$$

The sum above shows that S_{ij} is the coefficient of $x^i y^j$ in $(1 + x + y + qxy)^n$. □

We have developed the necessary tools for the covariance between the coefficients of weight enumerators. The following theorem gives the covariance between the weights of a code with random generator matrix.

Theorem 11. *The covariances between the coefficients of the weight enumerator of the random code C are given by*

$$\operatorname{cov}\bigl(W_C(x), W_C(y)\bigr) = \frac{(q-1)(q^k-1)}{q^n}\left((xy+q-1)^n - \frac{(x+q-1)^n(y+q-1)^n}{q^n}\right). \tag{29}$$

Proof. Let S_i be the coefficient of z^i in $W_C(1+z)$ given by Eq. (2)

$$S_i = \sum_{|I|=i} q^{k-r_I}.$$

The covariance of coefficients S_i and S_j is obtained by Eq. (25)

$$\begin{aligned}\operatorname{cov}(S_i, S_j) &= \frac{(q-1)(q^k-1)}{q^{i+j}} \sum_{|I|=i,|J|=j} \bigl(q^{|I\cap J|}-1\bigr)\\ &= \frac{(q-1)(q^k-1)}{q^{i+j}}\left(S_{ij} - \binom{n}{i}\binom{n}{j}\right),\end{aligned}$$

where S_{ij} was defined by Eq. (27). For the covariance of weight enumerators we substitute this

$$\begin{aligned}\operatorname{cov}\bigl(W_C(1+x), W_C(1+y)\bigr) &= \sum_{i,j} \operatorname{cov}(S_i, S_j)x^i y^j\\ &= (q-1)\bigl(q^k-1\bigr)\left(\left(\sum_{i,j} S_{ij}\left(\frac{x}{q}\right)^i\left(\frac{y}{q}\right)^j\right) - \left(1+\frac{x}{q}\right)^n\left(1+\frac{y}{q}\right)^n\right)\\ &= \frac{(q-1)(q^k-1)}{q^n}\left((xy+x+y+q)^n - \frac{(x+q)^n(y+q)^n}{q^n}\right).\end{aligned}$$

When we substitute $x-1$ for x and $y-1$ for y we get the formula. □

When we transform this theorem to the case of the random code $C^\perp$ the result is more interesting. We see in the covariance formula (29) that the coefficients A_i and A_j of the weight enumerator are correlated as coefficients of $x^i y^j$ for $i \neq j$ can be nonzero. However this is not the case for $C^\perp$.

Theorem 12. *The covariances between the coefficients of the weight enumerator of the random code $C^\perp$ are given by*

$$\operatorname{cov}\bigl(W_{C^\perp}(x), W_{C^\perp}(y)\bigr) = \frac{(q-1)(q^k-1)}{q^{2k}}\bigl((xy+q-1)^n - (xy)^n\bigr). \tag{30}$$

Proof. By MacWilliams duality we have

$$\operatorname{cov}\bigl(W_{C^\perp}(1+x), W_{C^\perp}(1+y)\bigr) = \operatorname{cov}\left(q^{-k}x^n W_C\left(1+\frac{q}{x}\right), q^{-k}y^n W_C\left(1+\frac{q}{y}\right)\right).$$

When we substitute the covariance of the weight enumerators from Eq. (29) the result follows from a simple calculation. □

Note that there are only diagonal terms in the covariance of weight enumerator of a random code in this sense. The following corollary is for emphasizing the fact that the coefficients are uncorrelated.

Corollary 13. *The coefficients of the weight enumerator of the random code $C^{\perp}$ are uncorrelated*

$$\operatorname{cov}\bigl(A_i^{\perp}, A_j^{\perp}\bigr) = 0. \tag{31}$$

Proof. We see in Eq. (30) that there is no $x^i y^j$ for $i \neq j$. Hence the coefficients are uncorrelated. □

The same result is obtained by Wadayama [14] independently in the binary case.

3. The triple correlations

In this section we will study the triple correlations between the ranks of three column submatrices. A triplet of submatrices gives a triplet of subspaces. So we will start with the classification of triplets of subspaces. The classification enables us to evaluate the triple correlations between the rank functions. The formula for the rank correlations turns out to be highly complicated. We made numerous evaluations with computer and we observed that the formula simplifies to a closed form. We give the closed form of this formula as a conjecture. We will give triple correlation formulas for the coefficients of the weight enumerators as equivalent forms of this conjecture.

3.1. *Classification of triplets of subspaces*

A triplet of subspaces U, V, W in E is represented in a diagram as follows

$$\begin{tikzcd} & V \arrow[d, hook] & \\ U \arrow[r, hook] & E & W \arrow[l, hook'] \end{tikzcd} \tag{32}$$

where each arrow is an inclusion. This corresponds to a so-called quiver representation of the Dynkin diagram D_4. We will obtain the classification of subspaces from the indecomposable representations of this quiver.

We recall some of the basic definitions of quiver representations from [4]. A quiver Q is a set of vertices and a set of arrows between some of the vertices. The set of vertices of Q is denoted by Q_v and the set of arrows by Q_a. The maps $t, h : Q_a \to Q_v$ map an arrow to its tail and to its head respectively.

Definition 14. Let Q be a quiver with a finite number of vertices and let $\mathbb{F}$ be an algebraically closed field. A representation of Q is a collection of linear spaces $V(x)$ over $\mathbb{F}$ for each vertex x and $\mathbb{F}$-linear maps $V(\alpha) : V(t(\alpha)) \to V(h(\alpha))$ for each arrow α in Q.

Gabriel's theorem [3] gives the finiteness condition and the isomorphism classes of the indecomposable representations of a quiver.

Theorem 15. *The number of isomorphism classes of indecomposable representations of a quiver Q is finite if and only if Q is a disjoint union of the simply-laced Dynkin diagrams A_n, D_n, E_6, E_7 and E_8 with arbitrary orientation. In this case, there is a bijection between the set of isomorphism classes of indecomposable representations and the positive roots of the corresponding root system.*

The positive roots corresponding to D_4 are given below [2].

$$\begin{tikzcd} & 1 \arrow[d] & \\ 1 \arrow[r] & 1 & 1, \arrow[l] \end{tikzcd} \qquad \begin{tikzcd} & 0 \arrow[d] & \\ 0 \arrow[r] & 1 & 0 \arrow[l] \end{tikzcd} \tag{33}$$

$$\begin{tikzcd} & 0 \arrow[d] & & & 1 \arrow[d] & & & 0 \arrow[d] & \\ 1 \arrow[r] & 0 & 0, \arrow[l] & 0 \arrow[r] & 0 & 0, \arrow[l] & 0 \arrow[r] & 0 & 1 \arrow[l] \end{tikzcd} \tag{34}$$

$$\begin{tikzcd} & 1 \arrow[d] & & & 1 \arrow[d] & & & 0 \arrow[d] & \\ 1 \arrow[r] & 1 & 0, \arrow[l] & 0 \arrow[r] & 1 & 1, \arrow[l] & 1 \arrow[r] & 1 & 1 \arrow[l] \end{tikzcd} \tag{35}$$

$$\begin{tikzcd} & 1 \arrow[d] & & & 0 \arrow[d] & & & 0 \arrow[d] & \\ 0 \arrow[r] & 1 & 0, \arrow[l] & 1 \arrow[r] & 1 & 0, \arrow[l] & 0 \arrow[r] & 1 & 1 \arrow[l] \end{tikzcd} \tag{36}$$

$$\begin{tikzcd} & 1 \arrow[d] & \\ 1 \arrow[r] & 2 & 1 \arrow[l] \end{tikzcd} \tag{37}$$

We are interested in the indecomposables that correspond to a subspace configuration. So the indecomposable representations of a triplet of subspaces has dimension vectors coming from the roots except those given in line (34).

We denote a 1-dimensional indecomposable with the dimension vector

$$\begin{matrix} & j & \\ i & 1 & k \end{matrix}$$

by $[ijk]$ and its multiplicity by m_{ijk}. Also we fix our notation for the two-dimensional indecomposable as $[111]_2$ and for its multiplicity as h. So the triplet $U, V, W \subset E$ has the following decomposition

$$\begin{aligned}(E; U, V, W) = {} & m_{111}[111] + m_{000}[000] + m_{110}[110] + m_{011}[011] + m_{101}[101] \\ & + m_{100}[100] + m_{010}[010] + m_{001}[001] + h[111]_2. \end{aligned} \tag{38}$$

The multiplicities $m_{111}, m_{000}, m_{110}, m_{011}, m_{101}, m_{100}, m_{010}, m_{001}, h$ of the indecomposables are called the natural invariants of the triplet given in diagram (32). They have the following interpretations

$$\begin{aligned} & m_{110} = \dim \frac{U \cap V}{U \cap V \cap W}, && m_{011} = \dim \frac{V \cap W}{U \cap V \cap W}, && m_{101} = \dim \frac{W \cap U}{U \cap V \cap W}, \\ & m_{100} = \dim \frac{U}{U \cap (V + W)}, && m_{010} = \dim \frac{V}{V \cap (W + U)}, && m_{001} = \dim \frac{W}{W \cap (U + V)}, \\ & m_{111} = \dim(U \cap V \cap W), && m_{000} = \operatorname{codim}(U + V + W) \end{aligned}$$

and

$$\begin{aligned} h &= \dim \frac{U \cap (V + W)}{(U \cap V) + (W \cap U)} = \dim \frac{V \cap (W + U)}{(V \cap W) + (U \cap V)} \\ &= \dim \frac{W \cap (U + V)}{(W \cap U) + (V \cap W)} = \frac{1}{2} \dim \frac{(U + V) \cap (V + W) \cap (W + U)}{(U \cap V) + (V \cap W) + (W \cap U)}. \end{aligned} \tag{39}$$

3.2. The number of triplets with given invariants

For the triple correlation problem we need the number of triplets of subspaces with given invariants. Since GL acts transitively on the triplets with fixed invariants, the number of triplets is the index of the automorphism group. Thus as the first attempt we calculate the order of the automorphism group of a triplet.

3.2.1. Automorphism group of a triplet

A morphism M between the triplets $(E; U, V, W)$ and $(E'; U', V', W')$ is a linear map $M : E \to E'$ such that $M(U) \subset U'$, $M(V) \subset V'$ and $M(W) \subset W'$.

The triplet $U, V, W \subset E$ has a direct sum decomposition given in Eq. (38). This decomposition gives

$$\mathrm{End}(E) = \sum \mathrm{Hom}(T_i, T_j),$$

where T_i, T_j runs through the indecomposables with the correct number of occurrence. The automorphisms of the triplet are the invertible elements of the endomorphisms. It turns out that the matrices $M = \{M[i, j] = \mathrm{Hom}(T_j, T_i)\}$ of the endomorphisms are block triangular for a particular ordering of the isotypical components in the decomposition. So after finding the form of the endomorphisms it is easy to find the order of the stabilizer group and hence the number of triplets.

It is easy to see that the dimensions of the spaces of morphisms $\mathrm{Hom}(T_i, T_j)$ are as given in the following table. The $[ij]$ entry of the table is the dimension of $\mathrm{Hom}(T_j, T_i)$.

	[111]	[110]	[101]	[011]	$[111]_2$	[010]	[100]	[001]	[000]
[111]	1	1	1	1	2	1	1	1	1
[110]	0	1	0	0	1	1	1	0	1
[101]	0	0	1	0	1	0	1	1	1
[011]	0	0	0	1	1	1	0	1	1
$[111]_2$	0	0	0	0	1	1	1	1	2
[010]	0	0	0	0	0	1	0	0	1
[100]	0	0	0	0	0	0	1	0	1
[001]	0	0	0	0	0	0	0	1	1
[000]	0	0	0	0	0	0	0	0	1

(40)

Proposition 16. *Let U, V, W be a triplet in a k-dimensional space E, with given invariants $m_{111}, m_{110}, m_{011}, m_{101}, h, m_{100}, m_{010}, m_{001}, m_{000}$. Then the order of the automorphism group $\mathcal{A}$ of the triplet is given by*

$$|\mathcal{A}| = q^{EA}[m_{111}]![m_{110}]![m_{011}]![m_{101}]![h]![m_{100}]![m_{010}]![m_{001}]![m_{000}]! \tag{41}$$

where

$$\begin{aligned} EA = \binom{k-h}{2} &+ h(m_{111} + m_{000}) - m_{110}m_{011} - m_{011}m_{101} - m_{101}m_{110} - m_{110}m_{001} \\ &- m_{011}m_{100} - m_{101}m_{010} - m_{100}m_{010} - m_{010}m_{001} - m_{001}m_{100}. \end{aligned} \tag{42}$$

Proof. The space of endomorphisms of the triplet splits into the direct sum

$$\mathrm{End}(U, V, W \subset E) = \sum \mathrm{Hom}(m_i T_i, m_j T_j),$$

where T_i's are the indecomposables and m_i's are the number of T_i's in the decomposition. The nonzero entries in the table (40) form a triangular matrix. This shows that nonsingularity is obtained if we choose the elements of the diagonal entries $\mathrm{Hom}(m_i T_i, m_i T_i)$ from $\mathrm{GL}(m_i)$. The other summands are completely free. Hence we have

$$|\mathcal{A}| = q^{FS}\left|\mathrm{GL}(m_{111})\right|\left|\mathrm{GL}(m_{110})\right|\left|\mathrm{GL}(m_{011})\right|\left|\mathrm{GL}(m_{101})\right|\left|\mathrm{GL}(h)\right|\left|\mathrm{GL}(m_{100})\right|$$
$$\times \left|\mathrm{GL}(m_{010})\right|\left|\mathrm{GL}(m_{001})\right|\left|\mathrm{GL}(m_{000})\right|$$

where FS is the number of free summands and it is given by

$$FS = (m_{111} + h + m_{000})(m_{110} + m_{011} + m_{101} + m_{100} + m_{010} + m_{001}) + m_{110}(m_{100} + m_{010})$$
$$+ m_{011}(m_{010} + m_{001}) + m_{101}(m_{001} + m_{100}) + m_{111}m_{000} + 2h(m_{111} + m_{000}).$$

Substituting the cardinalities of GLs and collecting the exponents we obtain the result. □

3.2.2. The number of triplets

In this part we will obtain the number of triplets with given dimensions and the dimensions of intersections. We denote the dimension of a space U by u and the dimension of an intersection by the product notation, e.g. $\dim(U \cap V) = uv$. The correspondence between the natural invariants and the dimensions are as follows

$$m_{110} = uv - uvw, \qquad m_{100} = u - uv - wu + uvw - h, \tag{43}$$
$$m_{011} = vw - uvw, \qquad m_{010} = v - vw - uv + uvw - h, \tag{44}$$
$$m_{101} = wu - uvw, \qquad m_{001} = w - wu - vw + uvw - h, \tag{45}$$
$$m_{111} = uvw, \qquad m_{000} = k - d, \tag{46}$$
$$d = \dim(U + V + W) = u + v + w - uv - vw - wu + uvw - h. \tag{47}$$

Proposition 17. *The number $N_{U,V,W}$ of triplets U, V, W in a k-dimensional space with given dimensions u, v, w, uv, vw, wu, uvw and h is given by*

$$N_{U,V,W} = q^{EXP}\frac{[k]!}{\mathcal{D}(U,V,W)}, \tag{48}$$

where

$$\mathcal{D}(U,V,W) = [uvw]![uv - uvw]![vw - uvw]![wu - uvw]![h]![k-d]!$$
$$\times [u - uv - wu + uvw - h]![v - vw - uv + uvw - h]!$$
$$\times [w - wu - vw + uvw - h]!,$$
$$d = u + v + w - uv - vw - wu + uvw - h,$$
$$EXP = (uv + vw + wu)\big(uv + vw + wu - (u + v + w) + 2h - 3uvw\big)$$
$$+ u \times v + v \times w + w \times u + (uvw - h)(u + v + w + 3uvw)$$
$$+ \frac{h(3h-1)}{2}.$$

Proof. The group GL(k) has a transitive action on the triplets, hence the number of triplets $N_{U,V,W}$ is the index of the automorphism group. We know the order of the automorphism group by Eq. (41). An easy calculation gives the number of triplets with given natural invariants

$$N_{U,V,W} = q^{EC}\frac{[k]!}{[m_{111}]![m_{110}]![m_{011}]![m_{101}]![h]![m_{100}]![m_{010}]![m_{001}]![m_{000}]!} \tag{49}$$

where

$$\begin{aligned} EC = kh &+ m_{110}m_{011} + m_{011}m_{101} + m_{101}m_{110} + m_{110}m_{001} + m_{011}m_{100} \\ &+ m_{101}m_{010} + m_{100}m_{010} + m_{010}m_{001} + m_{001}m_{100} - \frac{h(h+1)}{2} \\ &- h(m_{111} + m_{000}). \end{aligned} \tag{50}$$

When we substitute the dimensions and the dimensions of intersections of the spaces by Eqs. (43) through (47) into the previous formula we get the number of triplets in the desired form. □

3.3. The triple correlation between the ranks

The joint cumulants are generalizations of the covariance to the case of more than two variables. Let $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ be a random vector. The joint cumulant of the random variables ξ_i is given by

$$\sigma(\xi) = \frac{\partial^n}{\partial\lambda_1\partial\lambda_2\dots\partial\lambda_n} \log\bigl(\mathbb{E}\bigl(\exp(\lambda_1\xi_1 + \lambda_2\xi_2 + \dots + \lambda_n\xi_n)\bigr)\bigr)\Big|_{\lambda=0}. \tag{51}$$

The correlations are expressed via the joint cumulants in a simple way

$$\mathbb{E}(\xi_1\xi_2\dots\xi_n) = \sum_I \prod_i \sigma(\xi_{I_i}), \tag{52}$$

where the summation is over all the disjoint partitions I of $\{1, 2, \dots, n\}$ into nonempty subsets I_i.

Example 18. The first three joint cumulants are given as follows

$$\begin{aligned} \sigma(X) &= \mathbb{E}(X), \\ \sigma(X, Y) &= \operatorname{cov}(X, Y), \\ \sigma(X, Y, Z) &= \mathbb{E}(XYZ) - \operatorname{cov}(X, Y)\mathbb{E}(Z) - \operatorname{cov}(Y, Z)\mathbb{E}(X) - \operatorname{cov}(Z, X)\mathbb{E}(Y) \\ &\quad - \mathbb{E}(X)\mathbb{E}(Y)\mathbb{E}(Z). \end{aligned}$$

In this section we give the formula of the cumulant $\sigma(q^{-r_I}, q^{-r_J}, q^{-r_K})$ of the random variables $q^{-r_I}, q^{-r_J}, q^{-r_K}$. We remind that

$$\begin{aligned} \sigma\bigl(q^{-r_I}, q^{-r_J}, q^{-r_K}\bigr) = {}& \mathbb{E}\bigl(q^{-r_I-r_J-r_K}\bigr) - \mathbb{E}\bigl(q^{-r_I}\bigr)\operatorname{cov}\bigl(q^{-r_J-r_K}\bigr) - \mathbb{E}\bigl(q^{-r_J}\bigr)\operatorname{cov}\bigl(q^{-r_K-r_I}\bigr) \\ & - \mathbb{E}\bigl(q^{-r_K}\bigr)\operatorname{cov}\bigl(q^{-r_I-r_J}\bigr) - \mathbb{E}\bigl(q^{-r_I}\bigr)\mathbb{E}\bigl(q^{-r_J}\bigr)\mathbb{E}\bigl(q^{-r_K}\bigr). \end{aligned} \tag{53}$$

Through the calculations we use the following abbreviations for the column sets and their cardinalities

$$\begin{aligned} &IJ = I \cap J, && \overline{IJ} = (I \cap J) \setminus K, && \bar{I} = I \setminus (J \cup K), \\ &JK = J \cap K, && \overline{JK} = (J \cap K) \setminus I, && \bar{J} = J \setminus (K \cup I), \\ &KI = K \cap I, && \overline{KI} = (K \cap I) \setminus J, && \bar{K} = K \setminus (I \cup J), \\ &&& IJK = I \cap J \cap K. \end{aligned}$$

The column submatrix spanned for example by $I \cap J = IJ$ will be denoted by G_{IJ}.

Recall that we made use of the joint probability distribution of the ranks of submatrices for the covariance between the rank functions. For the triple correlations we need the joint probability distribution of the ranks of three submatrices. Analogous to the previous case we will introduce an auxiliary probability which gives the joint distribution after a summation over its parameters.

Let us consider a $k \times n$ matrix G, let $I, J, K \subset \{1, 2, \ldots, n\}$ be column sets such that the column sub-matrices G_I, G_J, G_K span the matrix, i.e. $|I \cup J \cup K| = n$. Let $U, V, W \subset \mathbb{F}_q^k$ be the column spaces of the sub-matrices G_{IJ}, G_{JK} and G_{KI}. Let r_I, r_J, r_K, r_{IJK} be the ranks of the column sub-matrices G_I, G_J, G_K and G_{IJK} respectively. To keep the notation short we will refer to the invariants of the triplet as U, V, W as well. Let $P(r_I, r_J, r_K, U, V, W, r_{IJK})$ be the probability that G has the described properties.

Proposition 19. *The probability $P(r_I, r_J, r_K, U, V, W, r_{IJK})$ is given by*

$$P(r_I, r_J, r_K, U, V, W, r_{IJK}) = q^{-kn+EP} N_{U,V,W} \mathcal{P} \tag{54}$$

where $N_{U,V,W}$ is given in Eq. (48)*, EP and $\mathcal{P}$ are defined as follows*

$$EP = IJKuvw + \overline{IJ}u + \overline{JK}v + \overline{KI}w + k(\bar{I} + \bar{J} + \bar{K}), \tag{55}$$

$$\begin{aligned}\mathcal{P} = {} & P(IJK, uvw, r_{IJK}) P(\overline{IJ}, u - r_{IJK}, u - r_{IJK}) P(\overline{JK}, v - r_{IJK}, v - r_{IJK}) \\ & \times P(\overline{KI}, w - r_{IJK}, w - r_{IJK}) P(\bar{I}, k - w - u + wu, r_I - w - u + wu) \\ & \times P(\bar{J}, k - u - v + uv, r_J - u - v + uv) \\ & \times P(\bar{K}, k - v - w + vw, r_K - v - w + vw).\end{aligned} \tag{56}$$

Proof. We will find the number $N(r_I, r_J, r_K, U, V, W, r_{IJK})$ of matrices with given properties so that we will have

$$P(r_I, r_J, r_K, U, V, W, r_{IJK}) = q^{-kn} N(r_I, r_J, r_K, U, V, W, r_{IJK}). \tag{57}$$

The triplet of subspaces that G_{IJ}, G_{JK}, G_{KI} form can be fixed among $N_{U,V,W}$ triplets.

The columns in IJK must be chosen from the uvw-dimensional space with rank r_{IJK}. The number of column submatrices G_{IJK} is given by the factor $q^{IJKuvw} P(IJK, uvw, r_{IJK})$.

We will extend this construction to the column sets IJ, JK and KI. The columns in $\overline{IJ}$ must be chosen in U and they must extend the IJK column vectors to a matrix of rank u. The number of such vector sets is equal to $q^{\overline{IJ}u} P(\overline{IJ}, u - r_{IJK}, u - r_{IJK})$ in the notation of Eq. (21). Extending the column set IJK to JK and KI brings the factors $q^{\overline{JK}v} P(\overline{JK}, v - r_{IJK}, v - r_{IJK})$ and $q^{\overline{KI}w} P(\overline{KI}, w - r_{IJK}, w - r_{IJK})$ respectively.

Finally we will extend the construction to the column sets I, J and K. For the set I, we have the given $IJ \cup KI$ columns with the rank $\dim(W + U) = w + u - wu$. The generators this time will be chosen in $\mathbb{F}_q^k$, they will extend the column set $IJ \cup KI$ to I to have rank r_I. So, the number of choices for the column set $\bar{I}$ is $q^{k\bar{I}} P(\bar{I}, k - (w + u - wu), r_I - (w + u - wu))$. Similarly the number of the column sets for $\bar{J}$ and $\bar{K}$ are $q^{k\bar{J}} P(\bar{J}, k - (u + v - uv), r_J - (u + v - uv))$ and $q^{k\bar{K}} P(\bar{K}, k - (v + w - vw), r_K - (v + w - vw))$ respectively. Bringing these factors we get

$$N(r_I, r_J, r_K, U, V, W, r_{IJK}) = q^{EP} N_{U,V,W} \mathcal{P}$$

where $\mathcal{P}$ is the product of all the probabilities mentioned and EP comes from the exponents appearing before them. The probability follows from Eq. (57). □

We will obtain the correlation between the ranks of a triplet of submatrices by the auxiliary probability.

Theorem 20. *Let G be a random $k \times n$ matrix. Let $I, J, K \subset \{1, 2, \ldots, n\}$ be three column subsets with $|I \cup J \cup K| = n$. The triple correlation $\mathbb{E}(q^{-r_I - r_J - r_K})$ between the ranks is given by*

$$\mathbb{E}\big(q^{-r_I - r_J - r_K}\big) = \sum q^{-kn + EP - r_I - r_J - r_K} N_{U,V,W} \mathcal{P}, \tag{58}$$

where EP and $\mathcal{P}$ are given in Eqs. (55) *and* (56) *respectively. The sum runs over r_I, r_J, r_K, all the invariants $u, v, w, uv, vw, wu, uvw, h$ of triplets with r_{IJK} that are subject to the constraints* (61) *through* (73).

Proof. The triple correlation is given by

$$\mathbb{E}\big(q^{-r_I - r_J - r_K}\big) = \sum_{r_I, r_J, r_K} q^{-r_I - r_J - r_K} P(r_I, r_J, r_K),$$

where $P(r_I, r_J, r_K)$ is the joint probability of the ranks r_I, r_J and r_K. We get this probability by summing up over the rest of the parameters from $P(r_I, r_J, r_K, U, V, W, r_{IJK})$ given in Eq. (54). So the correlation becomes

$$\mathbb{E}\big(q^{-r_I - r_J - r_K}\big) = \sum_{r_I, r_J, r_K} q^{-r_I - r_J - r_K} \sum P(r_I, r_J, r_K, U, V, W, r_{IJK}) \tag{59}$$

with the inner sum running over the parameters $r_{IJK}, h, uvw, uv, vw, wu, u, v, w$. When we substitute $P(r_I, r_J, r_K, U, V, W, r_{IJK})$ given in (54) we obtain

$$\mathbb{E}\big(q^{-r_I - r_J - r_K}\big) = \sum q^{-kn + EP - r_I - r_J - r_K} N_{U,V,W} \mathcal{P}, \tag{60}$$

where $N_{U,V,W}$, EP and $\mathcal{P}$ are given in Eqs. (48), (55) and (56) respectively.

The constraints for the parameters are the natural bounds so that no negative q-factorial appears in the formula

$$0 \leqslant r_{IJK} \leqslant \min(IJK, uvw), \tag{61}$$

$$0 \leqslant uvw \leqslant \min(uv, vw, wu), \tag{62}$$

$$\max(u + v - r_J, 0) \leqslant uv \leqslant \min(u, v), \tag{63}$$

$$\max(v + w - r_K, 0) \leqslant vw \leqslant \min(v, w), \tag{64}$$

$$\max(w + u - r_I, 0) \leqslant wu \leqslant \min(w, u), \tag{65}$$

$$0 \leqslant u \leqslant \min(r_I, r_J, IJ, k), \tag{66}$$

$$0 \leqslant v \leqslant \min(r_J, r_K, JK, k), \tag{67}$$

$$0 \leqslant w \leqslant \min(r_K, r_I, KI, k), \tag{68}$$

$$0 \leqslant r_I \leqslant \min(I, k), \tag{69}$$

$$0 \leqslant r_J \leqslant \min(J, k), \tag{70}$$

$$0 \leqslant r_K \leqslant \min(K, k), \tag{71}$$

and finally for h we have

$$h \leqslant \min(u - uv - wu + uvw, v - vw - uv + uvw, w - wu - vw + uvw), \tag{72}$$

$$h \geqslant \max(u + v + w - uv - vw - wu + uvw - k, 0). \qquad \square \tag{73}$$

3.4. The closed formula of the cumulant

Although the triple correlation formula (58) is extremely complicated modern computers with algebra packages can easily handle it. We provide a Maple procedure for evaluation at the web address [15]. Numerous computer experiments suggest that there exists a closed formula for the triple correlations. The formula becomes simpler for the cumulant (53). We give the formula for the cumulant suggested by experiments as a conjecture. The triple correlations can be obtained using the conjecture below and Eq. (53). The formula extends the previous results. When we set for example $K = \emptyset$ in the triple correlation formula, we get the pair correlations.

Conjecture 21. *Let G be a $k \times n$ matrix and let $I, J, K \subset \{1, 2, \dots n\}$ be three column sets. Then the cumulant $\sigma(q^{-r_I}, q^{-r_J}, q^{-r_K})$ is given by*

$$\sigma\left(q^{-r_I}, q^{-r_J}, q^{-r_K}\right) = \frac{(q-1)^2(q^k-1)}{q^{3k+I+J+K}}\left(q^{IJ+JK+KI-IJK} - \left(q^{IJ} + q^{JK} + q^{KI}\right)\right.$$
$$\left. + 2 + \left(q^k - q\right)\left(q^{IJK} - 1\right)\right).$$

We give an other conjecture which is on the joint cumulant of rank functions of three arbitrary submatrices. We remind the reader that when L is a row set and I is a column set r_{LI} denotes the rank of the matrix spanned by the row set L and the column set I.

Conjecture 22. *Let G be a random $k \times n$ matrix, let $I, J, K \subset \{1, 2, \dots, n\}$ be column sets and $L, M, N \subset \{1, 2, \dots, k\}$ be row sets. Then the joint cumulant of the rank functions $q^{-r_{LI}}$, $q^{-r_{MJ}}$ and $q^{-r_{NK}}$ is*

$$\begin{aligned}
&\sigma\left(q^{-r_{LI}}, q^{-r_{MJ}}, q^{-r_{NK}}\right) \\
&= \frac{(q^2-1)}{q^{|I|+|J|+|K|+|L|+|M|+|N|}}\left(\left(q^{IJK} - 1\right)\left(q^{LM+MN+NL-LMN} - 1\right)\right. \\
&\quad + \left(q^{LMN} - 1\right)\left(q^{IJ+JK+KI-IJK} - 1\right) - \left(q^{IJ} - 1\right)\left(q^{LM} - 1\right) - \left(q^{JK} - 1\right)\left(q^{MN} - 1\right) \\
&\quad \left. - \left(q^{KI} - 1\right)\left(q^{NL} - 1\right) - (q+1)\left(q^{IJK} - 1\right)\left(q^{LMN} - 1\right)\right).
\end{aligned}$$

Using similar arguments as in the proof of Theorem 11 one gets the corresponding conjectures for the coefficients of the weight enumerators. We sketch the line as follows. Let G be a random $k \times n$ matrix and C be the code assuming G as a generator matrix. Let

$$S_i = \sum_{|I|=i} q^{k-r_I},$$

so that

$$W_C(1+x) = \sum_i S_i x^i$$

and

$$\sigma\left(W_C(1+x), W_C(1+y), W_C(1+z)\right) = \sum_{i,j,l} \sigma(S_i, S_j, S_l) x^i y^j z^l. \tag{74}$$

For $\sigma(S_i, S_j, S_l)$ we will calculate generating functions. For this we will put the conjectural formula in the following form

$$\sigma\left(q^{-r_I}, q^{-r_J}, q^{-r_K}\right) = \frac{(q-1)^2(q^k-1)}{q^{3k}}\left(q^{-(I\cup J\cup K)} - \left(q^{-(I\cup J)-K} + q^{-(J\cup K)-I} + q^{-(K\cup I)-J}\right)\right.$$
$$\left. + 2q^{-I-J-K} + \left(q^k-q\right)\left(q^{-I-J-J+IJK} - q^{-I-J-K}\right)\right).$$

This gives the cumulants of the coefficients S_i's as follows

$$\sigma(S_i, S_j, S_l) = (q-1)^2\left(q^k-1\right)\sum_{|I|=i,\,|J|=j,\,|K|=l}\left(q^{-(I\cup J\cup K)} - q^{-(I\cup J)-K} - q^{-(J\cup K)-I}\right.$$
$$\left. - q^{-(K\cup I)-J} + 2q^{-I-J-K} + \left(q^k-q\right)\left(q^{-I-J-J+IJK} - q^{-I-J-K}\right)\right). \tag{75}$$

Let us introduce more notations now and let $[x^iy^jz^l]M(x,y,z)$ be the coefficient of $x^iy^jz^l$ in a polynomial $M(x,y,z)$. It is easy to check that

$$\sum q^{-(I\cup J\cup K)} = \left[x^iy^jz^l\right]\left((x+1)(y+1)(z+1)+q-1\right)^n q^{-n},$$
$$\sum q^{-(I\cup J)-K} = \left[x^iy^jz^l\right]\left((x+1)(y+1)+q-1\right)^n (z+q)^n q^{-2n},$$
$$\sum q^{-(J\cup K)-I} = \left[x^iy^jz^l\right]\left((y+1)(z+1)+q-1\right)^n (x+q)^n q^{-2n},$$
$$\sum q^{-(K\cup I)-J} = \left[x^iy^jz^l\right]\left((z+1)(x+1)+q-1\right)^n (y+q)^n q^{-2n},$$
$$\sum q^{-I-J-K} = \left[x^iy^jz^l\right](x+q)^n(y+q)^n(z+q)^n q^{-3n},$$
$$\sum q^{-I-J-K+IJK} = \left[x^iy^jz^l\right]\left((x+1)(y+1)(z+1)+(q-1)(x+y+z+q+1)\right)^n q^{-2n},$$

where all the summations are over $I, J, K \subset \{1, 2, \dots n\}$ with $|I| = i, |J| = j$ and $|K| = l$. This set of sums once put in Eq. (75) gives the cumulant $\sigma(S_i, S_j, S_l)$. Then summation over i, j, l gives the joint cumulant of the coefficients of the weight enumerator given in Eq. (74). When we substitute $x-1$ for x, $y-1$ for y and $z-1$ for z we get the joint cumulant of triplet of the coefficients A_is

$$\sigma\left(W_C(x), W_C(y), W_C(z)\right) = \sum_{i,j,l}\sigma(A_i, A_j, A_l)x^{n-i}y^{n-j}z^{n-l}.$$

We give the joint cumulants of triplets of coefficients of the weight enumerator as a corollary to Conjecture 21.

Corollary 23. *The joint cumulants of the coefficients of $W_C(t)$ are given by*

$$\sigma\left(W_C(x), W_C(y), W_C(z)\right) = (q-1)^2\left(q^k-1\right)\left\{q^{-n}(xyz+q-1)^n\right.$$
$$- q^{-2n}\left((xy+q-1)^n(z+q-1)^n + (yz+q-1)^n(x+q-1)^n\right.$$
$$\left. + (zx+q-1)^n(y+q-1)^n\right) + 2q^{-3n}(x+q-1)^n(y+q-1)^n(z+q-1)^n$$
$$+ \left(q^k-q\right)\left[q^{-2n}\left(xyz+(q-1)(x+y+z+q-2)\right)^n\right.$$
$$\left.\left. - q^{-3n}(x+q-1)^n(y+q-1)^n(z+q-1)^n\right]\right\}.$$

Given G we can consider the code $C^{\perp}$ which assumes G as a parity check matrix with the weight enumerator $W_{C^{\perp}}(t) = \sum_i A_i^{\perp} t^{n-i}$. Once again MacWilliams duality transforms the previous corollary to this case.

Corollary 24. *The joint cumulants of the coefficients of $W_{C^{\perp}}(t)$ are given by*

$$\begin{aligned}
\sigma\bigl(W_{C^{\perp}}(x), W_{C^{\perp}}(y), W_{C^{\perp}}(z)\bigr) &= \sum_{i,j,l} \sigma\bigl(A_i^{\perp}, A_j^{\perp}, A_l^{\perp}\bigr) x^{n-i} y^{n-j} z^{n-l} \\
&= q^{-3k}(q-1)^2\bigl(q^k-1\bigr)\bigl\{\bigl(xyz+(q-1)(x+y+z+q-2)\bigr)^n-(xyz)^n \\
&\quad -\bigl((xy+q-1)^n-(xy)^n\bigr)z^n-\bigl((yz+q-1)^n-(yz)^n\bigr)x^n \\
&\quad -\bigl((zx+q-1)^n-(zx)^n\bigr)y^n+\bigl(q^k-q\bigr)\bigl[(xyz+q-1)^n-(xyz)^n\bigr]\bigr\}.
\end{aligned}$$

References

[1] V. Blinovsky, Asymptotic Combinatorial Coding Theory, Kluwer Internat. Ser. Engrg. Comput. Sci., vol. 415, Kluwer Academic Publishers, Boston, MA, 1997.

[2] N. Bourbaki, Lie Groups and Lie Algebras, Elem. Math., Springer, Berlin, 2002, Chapters 4–6.

[3] P. Gabriel, Unzerlegbare Darstellungen I, Manuscripta Math. 6 (1972) 71–103.

[4] P. Gabriel, A.V. Roiter, Representations of Finite Dimensional Algebras, 2nd edition, Springer, Berlin, 1997.

[5] R.G. Gallager, Low Density Parity-Check Codes, MIT Press, Cambridge, MA, 1963.

[6] I.M. Gelfand, V.A. Ponomarev, Problems of linear algebra and classification of quadruples of subspaces in a finite-dimensional vector space, in: Hilbert Space Operators and Operator Algebras, Proc. Internat. Conf., Tihany, 1970, in: Colloq. Math. Soc. Janos Bolyai, vol. 5, North-Holland, Amsterdam, 1972, pp. 163–237.

[7] V. Kac, P. Cheung, Quantum Calculus, Universitext, Springer, Berlin, 2001.

[8] F.J. MacWilliams, A theorem on the distribution of weights of a systematic code, Bell System Tech. J. 42 (1963) 79–94.

[9] F.J. MacWilliams, N.J.A. Sloane, The Theory of Error-correcting Codes, North-Holland Math. Library, vol. 16, North-Holland Publishing Co., Amsterdam–New York–Oxford, 1977.

[10] N.E. Mnëv, The universality theorems on the classification problem of configuration varieties and convex polytopes varieties, in: Topology and Geometry—Rohlin Seminar, in: Lecture Notes in Math., vol. 1346, Springer, Berlin, 1988, pp. 527–543.

[11] M.A. Tsfasman, S.G. Vladut, Algebraic Geometric Codes, Kluwer Academic Publishers, Dordrecht, 1991.

[12] J.H. van Lint, Introduction to Coding Theory, third ed., Grad. Texts in Math., vol. 86, Springer-Verlag, Berlin, 1999.

[13] J.H. van Lint, R.M. Wilson, A Course in Combinatorics, second ed., Cambridge Univ. Press, Cambridge, 2001.

[14] T. Wadayama, Asymptotic concentration behaviors of linear combinations of weight distributions on random linear code ensemble, preprint, arXiv:0803.1025v1.

[15] Maple procedure for joint cumulant of ranks of triplet of column submatrices, http://www.fen.bilkent.edu.tr/~iozen.