# Finite Field Matrix Channels for Network Coding

Simon R. Blackburn and Jessica Claridge
Department of Mathematics
Royal Holloway University of London
Egham, Surrey TW20 0EX, United Kingdom

**Abstract**

In 2010, Silva, Kschischang and Kötter studied certain classes of finite field matrix channels in order to model random linear network coding where exactly $t$ random errors are introduced.

In this paper we consider a generalisation of these matrix channels where the number of errors is not required to be constant, indeed the number of errors may follow any distribution. We show that a capacity-achieving input distribution can always be taken to have a very restricted form (the distribution should be uniform given the rank of the input matrix). This result complements, and is inspired by, a paper of Nobrega, Silva and Uchoa-Filho, that establishes a similar result for a class of matrix channels that model network coding with link erasures. Our result shows that the capacity of our channels can be expressed as a maximisation over probability distributions on the set of possible ranks of input matrices: a set of linear rather than exponential size.

## 1 Introduction

Network coding, first defined in [1], allows intermediate nodes of a network to compute with and modify data, as opposed to the traditional view of nodes as 'on/off' switches. This can increase the rate of information flow through a network. It is shown in [13] that linear network coding is sufficient to maximise information flow in multicast problems, that is when there is one source node and information is to be transmitted to a set of sink nodes. Moreover, in [10] it is shown that for general multisource multicast problems, random linear network coding achieves capacity with probability exponentially approaching 1 with the code length.

In random linear network coding, the source injects packets into the network; these packets can be thought of as vectors of length $m$ with entries in a finite field $\mathbb{F}_q$ (where $q$ is a fixed power of a prime). The packets flow through a network of unknown topology to a sink node. Each intermediate node forwards packets that are random linear combinations of the packets it has received. A sink node attempts to reconstruct the message from these packets. In this context, Silva, Kschischang and Kötter [18] studied a channel defined as follows. We write $\mathbb{F}_q^{n\times m}$ to denote the set of all $n \times m$ matrices over $\mathbb{F}_q$, and write $\mathrm{GL}(n,q)$ for the set of all invertible matrices in $\mathbb{F}_q^{n\times n}$.

**Definition 1.1.** The *Multiplicative Matrix Channel* (MMC) has input set $\mathcal{X}$ and output set $\mathcal{Y}$, where $\mathcal{X} = \mathcal{Y} = \mathbb{F}_q^{n\times m}$. The channel law is

$$\boldsymbol{Y} = \boldsymbol{AX}$$

where $\boldsymbol{A} \in \mathrm{GL}(n,q)$ is chosen uniformly at random.

Here the rows of $\boldsymbol{X}$ correspond to the packets transmitted by the source, the rows of $\boldsymbol{Y}$ are the packets received by the sink, and the matrix $\boldsymbol{A}$ corresponds to the linear combinations of packets computed by the intermediate nodes.

Inspired by Montanari and Urbanke [14], Silva *et al* modelled the introduction of random errors into the network by considering the following generalisation of the MMC. We write $\mathbb{F}_q^{n\times m,r}$ for the set of all $n \times m$ matrices of rank $r$.

**Definition 1.2.** The *Additive Multiplicative Matrix Channel with $t$ errors* (AMMC) has input set $\mathcal{X}$ and output set $\mathcal{Y}$, where $\mathcal{X} = \mathcal{Y} = \mathbb{F}_q^{n\times m}$. The channel law is

$$\boldsymbol{Y} = \boldsymbol{A}(\boldsymbol{X} + \boldsymbol{B})$$

where $\boldsymbol{A} \in \mathrm{GL}(n, q)$ and $\boldsymbol{B} \in \mathbb{F}_q^{n\times m,t}$ are chosen uniformly and independently at random.

So the matrix $\boldsymbol{B}$ corresponds to the assumption that exactly $t$ linearly independent random errors have been introduced. The MMC is exactly the AMMC with zero errors.

We note that the AMMC is very different from the error model studied in the well-known paper by Kötter and Kschischang [12], where the errors are assumed to be adversarial (so the worst case is studied).

In [18] the authors give upper and lower bounds on the capacity of the AMMC, which are shown to converge in certain interesting limiting cases. The exact capacity of the AMMC for fixed parameter choices is hard to determine due to the many degrees of freedom involved: the naive formula maximises over a probability distribution on the set of possible input matrices, and this set is exponentially large.

In this paper we consider a generalisation of these matrix channels that allows the modelling of channels were the number of errors is not necessarily fixed. (For example, it enables the modelling of situations when at most $t$ errors are introduced, or when the errors are not necessarily linearly independent, or both.) To define our generalisation, we need the following notation which is due to Nobrega, Silva and Uchoa-Filho [15].

**Definition 1.3.** Let $\mathcal{R}$ be a probability distribution on the set $\{0, 1, \ldots, \min\{m, n\}\}$ of possible ranks of matrices $M \in \mathbb{F}_q^{n\times m}$. We define a distribution on the set $\mathbb{F}_q^{n\times m}$ of matrices by choosing $r$ according to $\mathcal{R}$, and then once $r$ is fixed choosing a matrix $M \in \mathbb{F}_q^{n\times m,r}$ uniformly at random. We say that this distribution is *Uniform Given Rank (UGR) with rank distribution* $\mathcal{R}$. We say a distribution on $\mathbb{F}_q^{n\times m}$ is *Uniform Given Rank (UGR)* if it is UGR with rank distribution $\mathcal{R}$ for some distribution $\mathcal{R}$.

We write $\mathcal{R}(r)$ for the probability of rank $r$ under the distribution $\mathcal{R}$. So a distribution on $\mathbb{F}_q^{n\times m}$ is UGR with rank distribution $\mathcal{R}$ if and only if each $M \in \mathbb{F}_q^{n\times m}$ of rank $r$ is chosen with probability $\mathcal{R}(r)/|\mathbb{F}_q^{n\times m,r}|$.

**Definition 1.4.** Let $\mathcal{R}$ be a probability distribution on the set $\{0, 1, \ldots, \min\{m, n\}\}$ of possible ranks of matrices $M \in \mathbb{F}_q^{n\times m}$. The *Generalised Additive Multiplicative MAtrix Channel with rank error distribution $\mathcal{R}$ (the Gamma channel $\Gamma(\mathcal{R})$)* has input set $\mathcal{X}$ and output set $\mathcal{Y}$, where $\mathcal{X} = \mathcal{Y} = \mathbb{F}_q^{n\times m}$. The channel law is

$$\boldsymbol{Y} = \boldsymbol{A}(\boldsymbol{X} + \boldsymbol{B})$$

where $\boldsymbol{A} \in \mathrm{GL}(n, q)$ is chosen uniformly, where $\boldsymbol{B} \in \mathbb{F}_q^{n\times m}$ is UGR with rank distribution $\mathcal{R}$, and where $\boldsymbol{A}$ and $\boldsymbol{B}$ are chosen independently.

We see that the AMMC is the special case of $\Gamma(\mathcal{R})$ when $\mathcal{R}$ is the distribution choosing rank $t$ with probability 1. In [18, §VI-D] the authors consider a generalisation of the AMMC which is exactly the Gamma channel in the case when the error rank is bounded by $t$, that is they consider $\Gamma(\mathcal{R})$ when $\mathcal{R}$ is any distribution with $\mathcal{R}(r) = 0$ for all $r > t$. For $t = n$ this channel is identical to the Gamma channel. However, the authors consider $t$ as a maximum value for the error rank (thus will be taken to be less than $n$), whereas we consider any full rank distribution, with high rank having low probability.

Our model covers several very natural situations (some of which are also covered by the generalised AMMC with $t < n$). For example, we may drop the assumption that the $t$ errors are linearly independent by defining $\mathcal{R}(r)$ to be the probability that $t$ vectors span a subspace of dimension $r$, when the vectors are chosen uniformly and independently. We can also extend this to model situations when the number $t$ of vectors varies according to natural distributions such as the binomial distribution (which arises when, for example, a packet is corrupted with some fixed non-zero probability). In practice, given a particular network, one may run tests on the network to see the actual error patterns produced and define an empirical distribution on ranks. One could also define an appropriate distribution by considering some combination of the situations described.

We are interested in the capacity of the Gamma channel. In the generalised AMMC [18, §VI.D] the authors establish a lower bound on the capacity that is at most $\log_q(t+1)$ lower than the capacity of the AMMC with the same value of $t$. Therefore in the limiting cases considered their generalised channel performs at least as well as the AMMC. This is a very useful result when $t$ is significantly smaller than $n$. However, if we take $t = n$ then the expressions [18, Eq. 19 & 20] for the capacity in the limiting cases considered evaluate to zero.

Throughout this paper, we assume that $q$, $n$, $m$ and $\mathcal{R}$ are fixed by the application. We will refer to these values as the *channel parameters.*

We note that the Gamma channel assumes that the *transfer matrix* $\boldsymbol{A}$ is always invertible. This is a realistic assumption in random linear network coding in standard situations: the field size $q$ is normally large, which means linear dependencies between random vectors are less likely.

In both [15] and [16] the authors consider (different) generalisations of the MMC channel that do not necessarily have a square full rank transfer matrix. Such channels allow modelling of network coding when no erroneous packets are injected into the network, but there may be link erasures. In [15], Nobrega, Silva and Uchoa-Filho define the transfer matrix to be picked from a UGR distribution. One result of [15] is that a capacity-achieving input distribution for their class of MMC channels can always be taken to be UGR.

A main result of this paper (Theorem 5.5) is that a capacity-achieving input distribution for Gamma channels can always be taken to be UGR. Theorem 5.5 is a significant extension of the result of [15] to a new class of channels; the extension requires new technical ideas. This result is in contrast to the coding schemes proposed in [18], which restrict input matrices to have a specific form, and achieves capacity in the limiting cases considered. Their restrictions on the input allows for straightforward decoding, whereas it is not immediately obvious of how to construct an efficient UGR coding scheme which achieves capacity for any given parameters, indeed this is a problem of interest for future research.

Corollary 6.2 to the main result of the paper provides an explicit method for computing the capacity of Gamma channels which maximises over a probability distribution over the set of possible ranks of input matrices, rather than the set of all input matrices itself. Thus we have reduced the problem of computing the capacity of a Gamma channel to an explicit maximisation over a set of variables of linear rather than exponential size. As examples of the results of this approach, the table below gives the computed capacity $C$ of the AMMC channel with 2 errors for $n \times 2n$ matrices over $\mathbb{F}_2$; the capacity $C'$ of the Gamma channel for $n \times 2n$ matrices over $\mathbb{F}_2$ when the number of errors is binomially distributed with expected number of errors equal to 2; and the capacity $C''$ when the number of errors is 0, 1 or 2 with

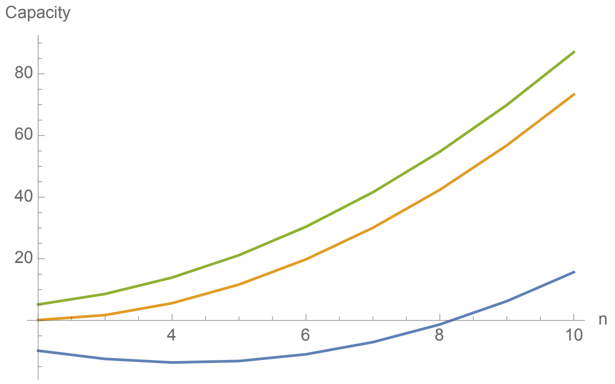

Figure 1: Capacity (in bits) of the AMMC channel with 2 errors for $n \times 2n$ matrices over $\mathbb{F}_2$. The three curves are: the upper bound from [18, Theorem 6]; the capacity computed using methods in this paper; the lower bound from [18, Theorem 7].

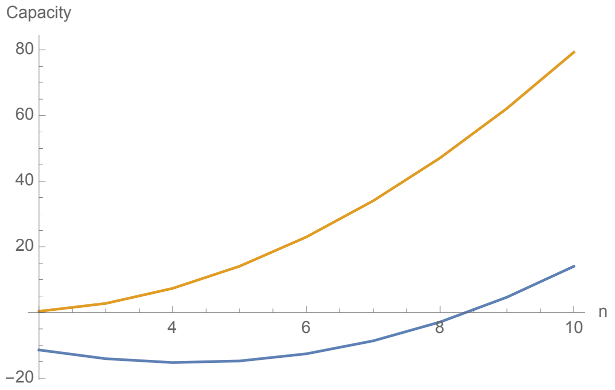

Figure 2: Capacity $C''$ (in bits) of the matrix channel with 0, 1 and 2 errors for $n \times 2n$ matrices over $\mathbb{F}_2$ specified in the text, together with the lower bound on the capacity from [18, § VI.D].

probabilities 1/7, 2/7 and 4/7 respectively.

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|
| $C$ | 1.731 | 5.586 | 11.644 | 18.807 | 30.050 | 42.381 | 56.798 | 73.290 |
| $C'$ | 1.117 | 4.633 | 10.422 | 18.368 | 28.395 | 40.491 | 54.676 | 70.996 |
| $C''$ | 2.765 | 7.354 | 14.090 | 23.003 | 34.032 | 47.094 | 62.177 | 79.274 |

Figure 1 plots the capacity $C$ of the AMMC channel together with general upper and lower bounds on the capacity. (These bounds are due to Silva et al. [18, Theorem 6 and 7]. We comment that an improved upper bound due to Claridge [5, Equation 6.6.2] is very similar to [18, Theorem 6] for these parameters.) Similarly, Figure 2 plots the capacity $C''$ of the third example channel together with the lower bound on the capacity due to Silva et al. [18, § VI.D].

The remainder of the paper is organised as follows. Section 2 proves some preliminary results needed in what follows. In Section 3 we state results from matrix theory that we use. Section 4 establishes a relationship between the distributions of the ranks of input and output matrices for a Gamma channel. Section 5 proves Theorem 5.5, and Section 6 proves Corollary 6.2, giving an exact expression for the capacity of the Gamma channels. In Section 7 we prove the results from matrix theory that we use in earlier sections. Finally, Section 8 contains some concluding remarks.

## 2 Preliminaries on finite-dimensional vector spaces

In this section we discuss finite-dimensional vector spaces and consider several counting problems involving subspaces and quotient spaces.

The *Gaussian binomial coefficient*, denoted $\left[ {m \atop d} \right]_q$, is defined to be the number of $d$-dimensional subspaces of an $m$-dimensional space over $\mathbb{F}_q$. It is given by (e.g. [4, §9.2])

$$\begin{bmatrix} m \\ d \end{bmatrix}_q = \begin{cases} \displaystyle\prod_{i=0}^{d-1} \frac{(q^m - q^i)}{(q^d - q^i)}, & \text{for } d \leq m \\ 0, & \text{for } d > m. \end{cases} \tag{1}$$

Let $V_1$ be a subspace of $V$. The following lemma gives the number of subspaces $U$ of $V$ where the intersection of $U$ and $V_1$, and the image of $U$ in the quotient space $V/V_1$, are both fixed.

**Lemma 2.1.** *Let $V$ be a $d_V$-dimensional vector space. Let $V_1$, $V_2$ be subspaces of $V$, of dimensions $d_{V_1}$ and $d_{V_2}$ respectively, such that $V_2 \subseteq V_1$. The number of $d_U$-dimensional subspaces $U \subseteq V$ such that $U \cap V_1 = V_2$ and the image of $U$ in the quotient space $V/V_1$ is the fixed $d_U - d_{V_2}$ dimensional space $U'$, is given by*

$$q^{(d_U - d_{V_2})(d_{V_1} - d_{V_2})}.$$

*Proof.* Fix a basis for $V_2$, say $\{b_{1,1}, \ldots, b_{1,d_{V_2}}\}$. Let $\pi : V \to V/V_1$ be the map which takes vectors in $V$ to their image in $V/V_1$. For $d_{U'} = d_U - d_{V_2}$, let $\{y_1, \ldots y_{d_{U'}}\}$ be a basis for $U'$, and let $\{b_{2,1}, \ldots, b_{2,d_{U'}}\}$ be some vectors in $V$ such that $\pi(b_{2,i}) = y_i$, for $i = 1, \ldots, d_{U'}$.

It is easy to check that every subspace $U$ of the form we are counting has a basis

$$B = \{b_{1,1}, \ldots, b_{1,d_{V_2}}, v_1 + b_{2,1}, \ldots, v_{d_{U'}} + b_{2,d_{U'}}\}$$

where $v_1, \ldots, v_{d_{U'}} \in \ker \pi = V_1$. Moreover, all bases of this form span a subspace $U$ of the form we are counting, and a basis

$$B' = \{b_{1,1}, \ldots, b_{1,d_{V_2}}, v'_1 + b_{2,1}, \ldots, v'_{d_{U'}} + b_{2,d_{U'}}\}$$

spans the same subspace as $B$ if and only if $[v_i] = [v'_i]$ for $i = 1, \ldots, d_{U'}$, where $[v]$ denotes the image of a vector $v$ in the quotient space $V_1/V_2$. Therefore there is a bijection between spaces $U$ of the required form and ordered sets $\{[v_1], \ldots, [v_{d_{U'}}]\}$ of elements in the quotient space $V_1/V_2$.

For $i = 1, \ldots d_{U'}$, there are $q^{d_{V_1} - d_{V_2}}$ choices for $[v_i] \in V_1/V_2$, thus there are

$$q^{d_{U'}(d_{V_1} - d_{V_2})} = q^{(d_U - d_{V_2})(d_{V_1} - d_{V_2})}. \tag{2}$$

choices for the ordered set $\{[v_1], \ldots, [v_{d_{U'}}]\}$. The result follows. □

Given a vector space $V$ and a subspace $V_1 \subseteq V$, Lemma 2.1 can be used to count subspaces $U$ of $V$ when either $U \cap V_1$ is fixed, or the image of $U$ in $V/V_1$ is fixed, or when only the dimensions of these spaces are fixed. These results are given in the following three corollaries.

**Corollary 2.2.** *Let $V$ be a $d_V$-dimensional vector space. Let $V_1$, $V_2$ be subspaces of $V$, of dimensions $d_{V_1}$ and $d_{V_2}$ respectively, such that $V_2 \subseteq V_1$. The number of $d_U$-dimensional subspaces $U \subseteq V$ such that $U \cap V_1 = V_2$, is given by*

$$q^{(d_U - d_{V_2})(d_{V_1} - d_{V_2})} \begin{bmatrix} d_V - d_{V_1} \\ d_U - d_{V_2} \end{bmatrix}_q.$$

*Proof.* The quotient space $V/V_1$ is a space of dimension $d_V - d_{V_1}$. Let $U'$ be a $(d_U - d_{V_2})$-dimensional subspace of $V/V_1$. There are

$$\begin{bmatrix} d_V - d_{V_1} \\ d_U - d_{V_2} \end{bmatrix}_q$$

possible choices for $U'$. For each such space $U'$, there are $q^{(d_U-d_{V_2})(d_{V_1}-d_{V_2})}$ possibilities for a space $U$ with whose image in the quotient $V/V_1$ is $U'$, by Lemma 2.1. □

**Corollary 2.3.** *Let $V$ be a $d_V$-dimensional vector space. Let $V_1$ be a $d_{V_1}$-dimensional subspace of $V$. The number of $d_U$-dimensional subspaces $U \subseteq V$ such that the image of $U$ in the quotient space $V/V_1$ is some fixed $d_{U'}$ dimensional space $U'$, is given by*

$$q^{d_{U'}(d_{V_1}-(d_U-d_{U'}))} \begin{bmatrix} d_{V_1} \\ d_U - d_{U'} \end{bmatrix}_q. \tag{3}$$

*Proof.* There are

$$\begin{bmatrix} d_{V_1} \\ d_U - d_{U'} \end{bmatrix}_q$$

possible choices for a $(d_U - d_{U'})$-dimensional subspace $V_2$ of $V_1$. For each choice of $V_1$, there are $q^{(d_U-(d_U-d_{U'}))(d_{V_1}-(d_U-d_{U'}))} = q^{d_{U'}(d_{V_1}-(d_U-d_{U'}))}$ possibilities for the space $U$ whose intersection with $V_1$ is the fixed space $V_2$, by Lemma 2.1. □

**Corollary 2.4.** *Let $V$ be a $d_V$-dimensional vector space. Let $V_1$ be a $d_{V_1}$-dimensional subspace of $V$. For a subspace $U$ of $V$, let $[U]$ denote the image of $U$ in the quotient space $V/V_1$. The number of $d_U$-dimensional subspaces $U \subseteq V$ such that $\dim(U \cap V_1) = d_{UV_1}$ is equal to the number of $d_U$-dimensional subspaces $U$ such that $\dim([U]) = d_U - d_{UV_1}$. This number is equal to*

$$q^{(d_U-d_{UV_1})(d_{V_1}-d_{UV_1})} \begin{bmatrix} d_V - d_{V_1} \\ d_U - d_{UV_1} \end{bmatrix}_q \begin{bmatrix} d_{V_1} \\ d_{UV_1} \end{bmatrix}_q.$$

*Proof.* Note that $\dim(U \cap V_1) = d_{UV_1}$ if and only if $\dim([U]) = d_U - d_{UV_1}$ hence the first statement of the lemma holds. Let $V_2$ be a $(d_{UV_1})$-dimensional subspace of $V_1$. There are

$$\begin{bmatrix} d_{V_1} \\ d_{UV_1} \end{bmatrix}_q$$

possible choices for $V_2$, and

$$\begin{bmatrix} d_V - d_{V_1} \\ d_U - d_{UV_1} \end{bmatrix}_q$$

possible choices for a $(d_U - d_{UV_1})$-dimensional subspace $U'$ of $V/V_1$. For each choice of $V_2$ and $U'$, Lemma 2.1 implies there are $q^{(d_U-d_{UV_1})(d_{V_1}-d_{UV_1})}$ possibilities for the space $U$ whose intersection with $V_1$ is the fixed space $V_2$, and image in the quotient $V/V_1$ is the fixed space $U'$. □

## 3 Matrices over finite fields

This short section describes the notation and results we use from the theory of matrices over finite fields.

Let $q$ be a non-trivial prime power, that is $q = p^n$ for some prime $p$ and integer $n \geq 1$. Let $\mathbb{F}_q$ be the finite field of order $q$. In the introduction we defined $\mathbb{F}_q^{n\times m}$ to be the set of $n \times m$ matrices with entries in $\mathbb{F}_q$, we defined $\mathbb{F}_q^{n\times m,r}$ to be the matrices in $\mathbb{F}_q^{n\times m}$ of rank $r$, and we defined $\mathrm{GL}(n, q)$ to be the set of invertible matrices in $\mathbb{F}_q^{n\times n}$.

For a matrix $M$, we write $\mathrm{rk}(M)$ for the rank of $M$ and we write $\mathrm{Row}(M)$ for the row space of $M$.

**Lemma 3.1.** *Let $U$ be a subspace of $\mathbb{F}_q^m$ of dimension $u$. The number $f_0(u)$ of matrices $M \in \mathbb{F}_q^{n\times m}$ such that* $\mathrm{Row}(M) = U$ *can be efficiently computed; it depends only on $q$, $n$, $m$ and $u$. For $0 \leq u \leq \min\{n, m\}$,*

$$f_0(u) = \prod_{i=0}^{u-1} q^n - q^i \tag{4}$$

$$= \sum_{v=0}^{u} (-1)^{u-v} q^{nv+\binom{u-v}{2}} \begin{bmatrix} u \\ v \end{bmatrix}_q . \tag{5}$$

By an efficient computation, we mean a polynomial (in $\max\{n, m\}$) number of arithmetic operations. Gabidulin [7, Theorem 4] establishes (4), and (5) follows from [7, Equation 13]. Therefore Lemma 3.1 immediately follows.

The following results will be proved in Section 7.

**Lemma 3.2.** *Let $U$ and $V$ be subspaces of $\mathbb{F}_q^m$ of dimensions $u$ and $v$ respectively. Let $h = \dim(U \cap V)$. Let $M \in \mathbb{F}_q^{n\times m}$ be a fixed matrix such that* $\mathrm{Row}(M) = U$. *Let $r$ be a non-negative integer. The number of matrices $B \in \mathbb{F}_q^{n\times m,r}$ such that* $\mathrm{Row}(B + M) = V$ *can be efficiently computed; it depends only on $q$, $n$, $m$, $r$, $u$, $v$ and $h$. We write $f_1(u, v, h; r)$ for the number of matrices $B$ of this form.*

**Lemma 3.3.** *Let $r$, $r_B$ and $r_X$ be non-negative integers. Let $X$ be a fixed matrix such that* $\mathrm{rk}(X) = r_X$. *The number of matrices $B \in \mathbb{F}_q^{n\times m,r_B}$ such that* $\mathrm{rk}(X + B) = r$ *can be efficiently computed; it depends only on $q$, $n$, $m$, $r$, $r_B$ and $r_X$. We write $f_2(r, r_X, r_B)$ for the number of matrices $B$ of this form.*

In Section 7, Theorems 7.3 and 7.4 we give exact expressions for the functions $f_1$ and $f_2$ respectively in terms of their inputs and the values $q$, $n$ and $m$, from which Lemmas 3.2 and 3.3 follow immediately.

We comment that the function $f_2$ has connections with rank metric codes (see e.g. [8], [17] for example). For a fixed matrix $X$ of rank $r_X$, the function $f_2(r_X, r_B, r)$ gives the number of matrices $B$ of rank $r_B$ such that $\mathrm{rk}(X+B) = r$. This is equal to the number of matrices $B'$ of rank $r_B$ such that $\mathrm{rk}(X - B') = r$ (setting $B' = -B$). The *rank distance* is a metric defined for two matrices $M_1, M_2 \in \mathbb{F}_q^{n\times m}$ to be

$$d_R(M_1, M_2) = \mathrm{rk}(M_1 - M_2).$$

Therefore, the value $f_2(r_X, r_B, r)$ gives the number of matrices of rank $r_B$, that have rank distance $r$ from some fixed matrix of rank $r_X$. Or equivalently, considering the space of all $n \times m$ matrices over $\mathbb{F}_q$, $f_2(r_X, r_B, r)$ is the volume of intersection of two spheres with rank radii $r_X$ and $r_W$ with centres at rank distance $r$. The analysis of the volume of intersection of spheres in the rank metric space can lead to the development of covering properties for rank metric codes, as explored by Gadouleau and Yan [9]. In [9, Lemma 1], the authors give an expression for the function $f_2$, showing that indeed it is efficiently computable. The expression they give was developed using the theory of association schemes. In Section 7 we give an expression for $f_2$ that avoids this theory, using direct counting arguments. Thus our new formula and proof give extra insight.

# 4 Input and output rank distributions

A distribution $\mathcal{P}_{\boldsymbol{X}}$ on the input set $\mathcal{X}$ of the Gamma channel induces a distribution (the *input rank distribution*) $\mathcal{R}_{\boldsymbol{X}}$ on the set of possible ranks of input matrices. Let $\mathcal{R}_{\boldsymbol{Y}}$ be the corresponding *output rank distribution*, induced from the distribution on the output set of the Gamma channel. A key result (Lemma 4.2) is that $\mathcal{R}_{\boldsymbol{Y}}$ depends on only the channel parameters and $\mathcal{R}_{\boldsymbol{X}}$ (rather than on $\mathcal{P}_{\boldsymbol{X}}$ itself). This section aims to prove this result: it will play a vital role in the proof of Theorem 5.5 below.

**Definition 4.1.** Let $r, r_X, r_B \in \{0, \ldots, \min\{n, m\}\}$. Define

$$\rho(r; r_X, r_B) = \frac{f_2(r, r_X, r_B)}{|\mathbb{F}_q^{n\times m, r_B}|},$$

where $f_2$ is as defined in Lemma 3.3. For any fixed matrix $X \in \mathbb{F}_q^{n\times m, r_X}$, we see that $\rho(r; r_X, r_B)$ gives the proportion of matrices $B \in \mathbb{F}_q^{n\times m, r_B}$ with $\operatorname{rk}(X + B) = r$. Let $\mathcal{R}$ be a probability distribution on the set $\{0, 1, \ldots, \min\{n, m\}\}$ of possible ranks of $n \times m$ matrices. Define

$$\rho(r; r_X) = \sum_{r_B=0}^{\min\{n,m\}} \mathcal{R}(r_B)\rho(r; r_X, r_B),$$

so that $\rho(r; r_X)$ gives the weighted average of this proportion over the possible ranks of matrices $B$.

**Lemma 4.1.** *Let $\boldsymbol{X}$ be an $n\times m$ matrix sampled from some distribution $\mathcal{P}_{\boldsymbol{X}}$ on $\mathbb{F}_q^{n\times m}$. Let $\boldsymbol{B}$ be an $n \times m$ matrix sampled from a UGR distribution with rank distribution $\mathcal{R}$, where $\boldsymbol{X}$ and $\boldsymbol{B}$ are chosen independently. Let $r, r_X, r_B \in \{0, \ldots, \min\{n, m\}\}$. Then*

$$\rho(r; r_X, r_B) = \Pr(\operatorname{rk}(\boldsymbol{X} + \boldsymbol{B}) = r | \operatorname{rk}(\boldsymbol{X}) = r_X \text{ and } \operatorname{rk}(\boldsymbol{B}) = r_B), \tag{6}$$

*and*

$$\rho(r; r_X) = \Pr(\operatorname{rk}(\boldsymbol{X} + \boldsymbol{B}) = r | \operatorname{rk}(\boldsymbol{X}) = r_X). \tag{7}$$

*Proof.* Let $X$ be a fixed $n \times m$ matrix of rank $r_X$. Then, since $\boldsymbol{B}$ has a UGR distribution,

$$\begin{aligned}
&\Pr(\operatorname{rk}(X + \boldsymbol{B}) = r | \operatorname{rk}(\boldsymbol{B}) = r_B)\\
&\quad = \frac{|\{B \in \mathbb{F}_q^{n\times m, r_B} : \operatorname{rk}(X + B) = r\}|}{|\mathbb{F}_q^{n\times m, r_B}|}\\
&\quad = \frac{f_2(r, r_X, r_B)}{|\mathbb{F}_q^{n\times m, r_B}|}\\
&\quad = \rho(r; r_X, r_B).
\end{aligned} \tag{8}$$

Note that (8) only depends on $\operatorname{rk}(X)$, not $X$ itself. Hence

$$\begin{aligned}
&\Pr(\operatorname{rk}(\boldsymbol{X} + \boldsymbol{B}) = r | \operatorname{rk}(\boldsymbol{X}) = r_X, \operatorname{rk}(\boldsymbol{B}) = r_B)\\
&= \sum_X \Pr(\boldsymbol{X} = X) \Pr(\operatorname{rk}(X + \boldsymbol{B}) = r | \operatorname{rk}(\boldsymbol{B}) = r_B)\\
&= \sum_X \Pr(\boldsymbol{X} = X)\rho(r; r_X, r_B)\\
&= \rho(r; r_X, r_B),
\end{aligned}$$

where the sums are over matrices $X \in \mathbb{F}_q^{n\times m, r_X}$. Thus (6) holds. Also

$$\begin{aligned}
&\Pr(\operatorname{rk}(\boldsymbol{X} + \boldsymbol{B}) = r | \operatorname{rk}(\boldsymbol{X}) = r_X)\\
&\quad = \sum_{r_B=0}^{\min\{n,m\}} \mathcal{R}(r_B)\rho(r; r_X, r_B) \text{ (by (6))}\\
&\quad = \rho(r; r_X).
\end{aligned}$$

Thus (7) holds, and so the lemma follows. □

**Lemma 4.2.** *For the Gamma channel* $\Gamma(\mathcal{R})$ *with input rank distribution* $\mathcal{R}_{\boldsymbol{X}}$*, the output rank distribution is given by*

$$\mathcal{R}_{\boldsymbol{Y}}(r) = \sum_{r_X, r_B = 0}^{\min\{n,m\}} \mathcal{R}_{\boldsymbol{X}}(r_X)\mathcal{R}(r_B)\frac{f_2(r, r_X, r_B)}{|\mathbb{F}_q^{n\times m, r_B}|}$$

*for* $r = 1, \ldots, \min\{n, m\}$*. In particular,* $\mathcal{R}_{\boldsymbol{Y}}$ *depends only on the input rank distribution (and the channel parameters), not on the input distribution itself.*

*Proof.* We have that $\Pr(\mathrm{rk}(\boldsymbol{X}) = r_X) = \mathcal{R}_{\boldsymbol{X}}(r_X)$ and $\Pr(\mathrm{rk}(\boldsymbol{B}) = r_B) = \mathcal{R}(r_B)$. Hence, by (6),

$$\begin{aligned}\mathcal{R}_{\boldsymbol{Y}}(r) &= \Pr(\mathrm{rk}(\boldsymbol{Y}) = r)\\ &= \sum_{r_X, r_B = 0}^{\min\{n,m\}} \mathcal{R}_{\boldsymbol{X}}(r_X)\mathcal{R}(r_B)\rho\left(r_Y; r_X, r_B\right)\\ &= \sum_{r_X, r_B = 0}^{\min\{n,m\}} \mathcal{R}_{\boldsymbol{X}}(r_X)\mathcal{R}(r_B)\frac{f_2(r, r_X, r_B)}{|\mathbb{F}_q^{n\times m, r_B}|}.\end{aligned}$$

□

# 5 A UGR input distribution achieves capacity

This section shows (Theorem 5.5) that there exists a UGR input distribution to the Gamma channel that achieves capacity.

**Lemma 5.1.** *Let* $M$ *and* $M'$ *be fixed* $n \times m$ *matrices of the same rank. Let* $\boldsymbol{B}$ *be an* $n \times m$ *matrix picked from a UGR distribution, and let* $\boldsymbol{A}$ *be an* $n \times n$ *matrix picked uniformly from* $\mathrm{GL}(n, q)$*, with* $\boldsymbol{B}$ *and* $\boldsymbol{A}$ *picked independently. Let* $\boldsymbol{Y} = \boldsymbol{A}(M + \boldsymbol{B})$ *and let* $\boldsymbol{Y'} = \boldsymbol{A}(M' + \boldsymbol{B})$*. Then*

$$H(\boldsymbol{Y}) = H(\boldsymbol{Y'}).$$

*Proof.* Let $A$ be a fixed $n\times n$ invertible matrix. Since the matrices $AM$ and $AM'$ have the same rank, there exist invertible matrices $R$ and $C$ such that $AM' = RAMC$. Consider the linear transformation $\varphi : \mathbb{F}_q^{n\times m} \to \mathbb{F}_q^{n\times m}$ defined by $\varphi(\boldsymbol{Y}) = R\boldsymbol{Y}C$. It is simple to check that $\varphi$ is well defined and a bijection. Note that

$$\begin{aligned}\varphi(A(M + \boldsymbol{B})) &= RAMC + RA\boldsymbol{B}C\\ &= A(M' + A^{-1}RA\boldsymbol{B}C).\end{aligned}$$

Since $\boldsymbol{B}$ is picked uniformly once its rank is determined, pre- and post-multiplying $\boldsymbol{B}$ by fixed invertible matrices gives a uniform matrix of the same rank, therefore $\boldsymbol{B}$ and $A^{-1}RA\boldsymbol{B}C$ have the same distribution. Now

$$\begin{aligned}&\Pr\left(\boldsymbol{Y} = Y|\boldsymbol{A} = A\right)\\ &\quad= \Pr\left(A(M + \boldsymbol{B}) = Y\right)\\ &\quad= \Pr\left(\varphi(A(M + \boldsymbol{B})) = \varphi(Y)\right)\\ &\quad= \Pr\left(A(M' + A^{-1}RA\boldsymbol{B}C) = \varphi(Y)\right)\\ &\quad= \Pr\left(A(M' + \boldsymbol{B}) = \varphi(Y)\right) &(9)\\ &\quad= \Pr\left(\boldsymbol{Y'} = \varphi(Y)|\boldsymbol{A} = A\right), &(10)\end{aligned}$$

where (9) holds since the distributions of $\boldsymbol{B}$ and $A^{-1}RA\boldsymbol{B}C$ are the same. Since

(10) is true for any fixed matrix $A$, it follows that

$$\begin{aligned}\Pr(\boldsymbol{Y} = Y) &= \sum_{A\in \mathrm{GL}(n,q)} \Pr(\boldsymbol{A} = A) \Pr(\boldsymbol{Y} = Y | \boldsymbol{A} = A) \\ &= \sum_{A\in \mathrm{GL}(n,q)} \Pr(\boldsymbol{A} = A) \Pr(\boldsymbol{Y'} = \varphi(Y) | \boldsymbol{A} = A) \\ &= \Pr(\boldsymbol{Y'} = \varphi(Y)). \end{aligned} \tag{11}$$

Thus $\boldsymbol{Y}$ and $\boldsymbol{Y'}$ have the same distribution, up to relabeling by $\varphi$. In particular, we find that $H(\boldsymbol{Y}) = H(\boldsymbol{Y'})$. □

**Definition 5.1.** Let $M$ be any $n \times m$ matrix of rank $r$. Let $\boldsymbol{A}$ be an $n \times n$ invertible matrix chosen uniformly from $\mathrm{GL}(n, q)$. Let $\boldsymbol{B}$ be an $n \times m$ matrix chosen from a UGR distribution with rank distribution $\mathcal{R}$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are picked independently. We define

$$h_r = H\left(\boldsymbol{A}(M + \boldsymbol{B})\right).$$

Lemma 5.1 implies that the value $h_r$ does not depend on $M$, only on the rank $r$ and the channel parameters $q, n, m$ and $\mathcal{R}$. The exact value of $h_r$ will be calculated later, in Theorem 6.1.

**Lemma 5.2.** *Consider the Gamma channel $\Gamma(\mathcal{R})$. Let the input matrix $\boldsymbol{X}$ be sampled from a distribution $\mathcal{P}_{\boldsymbol{X}}$ with associated rank distribution $\mathcal{R}_{\boldsymbol{X}}$, and let $\boldsymbol{Y}$ be the corresponding output matrix. Then*

$$H(\boldsymbol{Y}|\boldsymbol{X}) = \sum_{r=0}^{\min\{n,m\}} \mathcal{R}_{\boldsymbol{X}}(r) h_r.$$

*In particular, $H(\boldsymbol{Y}|\boldsymbol{X})$ depends only on the associated input rank distribution $\mathcal{R}_{\boldsymbol{X}}$ and the channel parameters.*

*Proof.* Choosing $\boldsymbol{A}$ and $\boldsymbol{B}$ as in the definition of the Gamma channel, we see that

$$\begin{aligned} H(\boldsymbol{Y}|\boldsymbol{X}) &= \sum_{X\in\mathcal{X}} P(\boldsymbol{X} = X) H(\boldsymbol{A}(X + \boldsymbol{B})) \\ &= \sum_{X\in\mathcal{X}} P(\boldsymbol{X} = X) h_{\mathrm{rk}(X)} \\ &= \sum_{r=0}^{\min\{n,m\}} \mathcal{R}_{\boldsymbol{X}}(r) h_r, \end{aligned}$$

which establishes the first assertion of the lemma. The second assertion follows since $h_r$ depends only on $r$ and the channel parameters. □

The following lemma is a well known result, see for example [6, Ex. 2.28].

**Lemma 5.3.** *Let $\boldsymbol{Y_1}$ and $\boldsymbol{Y_2}$ be two random $n \times m$ matrices, sampled from distributions with the same associated rank distribution $\mathcal{R}_{\boldsymbol{Y}}$. If the distribution of $\boldsymbol{Y_2}$ is UGR then $H(\boldsymbol{Y_2}) \geq H(\boldsymbol{Y_1})$.*

**Lemma 5.4.** *Consider the Gamma channel $\Gamma(\mathcal{R})$. If the input distribution $\mathcal{P}_{\boldsymbol{X}}$ is UGR then the induced output distribution $\mathcal{P}_{\boldsymbol{Y}}$ is also UGR.*

*Proof.* Suppose the input distribution is UGR, with rank distribution $\mathcal{R}_{\boldsymbol{X}}$. We start by showing that the distribution of $\boldsymbol{X} + \boldsymbol{B}$ is UGR. Let $D$ be any $n \times m$ matrix.

Then

$$\begin{aligned}
&\Pr(\boldsymbol{X}+\boldsymbol{B}=D)\\
&=\sum_{X\in\mathbb{F}_q^{n\times m}}\Pr(\boldsymbol{X}=X)\Pr(\boldsymbol{X}+\boldsymbol{B}=D|\boldsymbol{X}=X)\\
&=\sum_{X\in\mathbb{F}_q^{n\times m}}\frac{\mathcal{R}_{\boldsymbol{X}}(\mathrm{rk}(X))}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(X)}|}\Pr(X+\boldsymbol{B}=D),
\end{aligned}$$

since $\boldsymbol{X}$ is sampled from a UGR distribution. Hence

$$\begin{aligned}
&\Pr(\boldsymbol{X}+\boldsymbol{B}=D)\\
&=\sum_{r=0}^{\min\{n,m\}}\frac{\mathcal{R}_{\boldsymbol{X}}(r)}{|\mathbb{F}_q^{n\times m,r}|}\sum_{X\in\mathbb{F}_q^{n\times m,r}}\Pr(\boldsymbol{B}=D-X)\\
&=\sum_{r=0}^{\min\{n,m\}}\frac{\mathcal{R}_{\boldsymbol{X}}(r)}{|\mathbb{F}_q^{n\times m,r}|}\sum_{X\in\mathbb{F}_q^{n\times m,r}}\frac{\mathcal{R}(\mathrm{rk}(D-X))}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(D-X)}|},
\end{aligned}$$

since $\boldsymbol{X}$ and $\boldsymbol{B}$ are independent, and since $\boldsymbol{B}$ has a UGR distribution with rank distribution $\mathcal{R}$. Now

$$\begin{aligned}
&\sum_{X\in\mathbb{F}_q^{n\times m,r}}\frac{\mathcal{R}(\mathrm{rk}(D-X))}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(D-X)}|}\\
&=\sum_{r_B=0}^{\min\{n,m\}}|\{X\in\mathbb{F}_q^{n\times m,r}:\mathrm{rk}(D-X)=r_B\}|\frac{\mathcal{R}(r_B)}{|\mathbb{F}_q^{n\times m,r_B}|}\\
&=\sum_{r_B=0}^{\min\{n,m\}}f_2(r_B,\mathrm{rk}(D),r)\frac{\mathcal{R}(r_B)}{|\mathbb{F}_q^{n\times m,r_B}|}
\end{aligned}$$

and so

$$\begin{aligned}
&\Pr(\boldsymbol{X}+\boldsymbol{B}=D)\\
&=\sum_{r=0}^{\min\{n,m\}}\frac{\mathcal{R}_{\boldsymbol{X}}(r)}{|\mathbb{F}_q^{n\times m,r}|}\sum_{r_B=0}^{\min\{n,m\}}f_2(r_B,\mathrm{rk}(D),r)\frac{\mathcal{R}(r_B)}{|\mathbb{F}_q^{n\times m,r_B}|}.
\end{aligned}$$

So $\Pr(\boldsymbol{X}+\boldsymbol{B}=D)$ does not depend on the specific matrix $D$, only its rank. Therefore, given any two $n\times m$ matrices $D_1,D_2$ of the same rank,

$$\Pr(\boldsymbol{X}+\boldsymbol{B}=D_1)=\Pr(\boldsymbol{X}+\boldsymbol{B}=D_2).$$

Hence $\boldsymbol{X}+\boldsymbol{B}$ has a UGR distribution.

Let $A$ be a fixed $n\times n$ invertible matrix. Since $\boldsymbol{X}+\boldsymbol{B}$ is picked uniformly once its rank is determined, multiplying $\boldsymbol{X}+\boldsymbol{B}$ by the invertible matrix $A$ will give a uniform matrix of the same rank, therefore $A(\boldsymbol{X}+\boldsymbol{B})$ has a UGR distribution. So, defining $\boldsymbol{Y}=\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{B})$ to be the output of the Gamma channel, we see that for any $n\times m$ matrix $Y$

$$\begin{aligned}
\Pr(\boldsymbol{Y}=Y|\boldsymbol{A}=A)&=\Pr(A(\boldsymbol{X}+\boldsymbol{B})=Y)\\
&=\frac{\Pr\left(\mathrm{rk}(A(\boldsymbol{X}+\boldsymbol{B}))=\mathrm{rk}(Y)\right)}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(Y)}|}\\
&=\frac{\Pr\left(\mathrm{rk}(\boldsymbol{Y})=\mathrm{rk}(Y)|\boldsymbol{A}=A\right)}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(Y)}|},
\end{aligned}$$

where the second equality follows since $A(\boldsymbol{X}+\boldsymbol{B})$ has a UGR distribution. Thus

$$\begin{aligned}
\Pr(\boldsymbol{Y}=Y) &= \sum_{A\in\mathrm{GL}(n,q)} \Pr(\boldsymbol{A}=A)\Pr(\boldsymbol{Y}=Y|\boldsymbol{A}=A)\\
&= \sum_{A\in\mathrm{GL}(n,q)} \Pr(\boldsymbol{A}=A)\frac{\Pr(\mathrm{rk}(\boldsymbol{Y})=\mathrm{rk}(Y)|\boldsymbol{A}=A)}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(Y)}|}\\
&= \frac{1}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(Y)}|}\Pr(\mathrm{rk}(\boldsymbol{Y})=\mathrm{rk}(Y)). \qquad (12)
\end{aligned}$$

Since (12) holds for all $Y\in\mathbb{F}_q^{n\times m}$ it follows that $\boldsymbol{Y}$ has a UGR distribution. □

**Theorem 5.5.** *For the Gamma channel $\Gamma(\mathcal{R})$, there exists a UGR input distribution that achieves channel capacity. Moreover, given any input distribution $\mathcal{P}_{\boldsymbol{X}}$ with associated rank distribution $\mathcal{R}_{\boldsymbol{X}}$, if $\mathcal{P}_{\boldsymbol{X}}$ achieves capacity then the UGR distribution with rank distribution $\mathcal{R}_{\boldsymbol{X}}$ achieves capacity.*

*Proof.* Let $\boldsymbol{X_1}$ be a channel input, with output $\boldsymbol{Y_1}$ such that $\mathcal{P}_{\boldsymbol{X_1}}$ is a capacity achieving input distribution. That is $\max_{\mathcal{P}_{\boldsymbol{X}}}\{I(\boldsymbol{X},\boldsymbol{Y})\}=I(\boldsymbol{X_1},\boldsymbol{Y_1})$. Then define the input $\boldsymbol{X_2}$ with output $\boldsymbol{Y_2}$ to be distributed such that $\mathcal{P}_{\boldsymbol{X_2}}$ is the UGR distribution with $\mathcal{R}_{\boldsymbol{X_2}}=\mathcal{R}_{\boldsymbol{X_1}}$. To prove the theorem it suffices to show $I(\boldsymbol{X_2},\boldsymbol{Y_2})\geq I(\boldsymbol{X_1},\boldsymbol{Y_1})$.

By Lemma 4.2, $\mathcal{R}_{\boldsymbol{Y_2}}=\mathcal{R}_{\boldsymbol{Y_1}}$ and by Lemma 5.4, $\boldsymbol{Y_2}$ has a UGR distribution. Therefore, by Lemma 5.3,

$$H(\boldsymbol{Y_2})\geq H(\boldsymbol{Y_1}). \qquad (13)$$

Also, since $\mathcal{R}_{\boldsymbol{X_2}}=\mathcal{R}_{\boldsymbol{X_1}}$, Lemma 5.2 implies that

$$H(\boldsymbol{Y_2}|\boldsymbol{X_2})=H(\boldsymbol{Y_1}|\boldsymbol{X_1}). \qquad (14)$$

Using (13) and (14), it follows that

$$\begin{aligned}
I(\boldsymbol{X_2},\boldsymbol{Y_2}) &= H(\boldsymbol{Y_2})-H(\boldsymbol{Y_2}|\boldsymbol{X_2})\\
&\geq H(\boldsymbol{Y_1})-H(\boldsymbol{Y_2}|\boldsymbol{X_2})\\
&= H(\boldsymbol{Y_1})-H(\boldsymbol{Y_1}|\boldsymbol{X_1})\\
&= I(\boldsymbol{X_1},\boldsymbol{Y_1}).
\end{aligned}$$

□

# 6 Optimal input distributions and channel capacity

Theorem 5.5 reduces the problem of computing the Gamma channel capacity to a maximisation over a set of variables of linear rather than exponential size, since the UGR distribution is determined by the distribution $\mathcal{R}_{\boldsymbol{X}}$ on a set of size $\min\{n,m\}+1$. In this section we give an expression for this maximisation problem in terms of the channel parameters and the efficiently computable functions $f_0$, $f_1$ and $f_2$ defined in Section 3. Since the mutual information is concave when considered as a function over possible input distributions (see e.g. [6, Theorem 2.7.4]), this is a concave maximisation problem and hence efficiently computable (see e.g. [3]). Therefore the expression obtained provides a means for efficiently computing the exact channel capacity, and determining an optimal input rank distribution.

We begin by computing the value of $h_r$, as defined in Definition 5.1. This is needed to compute the maximisation problem in Corollary 6.2 that gives rise to the channel capacity.

**Theorem 6.1.** *The value $h_r$, as defined in Definition 5.1, is given by*

$$h_r = \sum_{v=0}^{\min\{n,m\}} \sum_{h=0}^{\min\{r,v\}} q^{(v-h)(r-h)} \begin{bmatrix} r \\ h \end{bmatrix}_q \begin{bmatrix} m-r \\ v-h \end{bmatrix}_q$$
$$\cdot \left( \sum_{r_B=0}^{\min\{n,m\}} \mathcal{R}(r_B) \frac{f_1(r,v,h;r_B)}{|\mathbb{F}_q^{n\times m,r_B}|} \right) \log \left( \frac{f_0(v)}{\sum_{r_B=h}^{\min\{n,m,v+h\}} \mathcal{R}(r_B) \frac{f_1(r,v,h;r_B)}{|\mathbb{F}_q^{n\times m,r_B}|}} \right).$$

*where $f_0$ is as defined in Lemma 3.1 and $f_1$ is as defined in Lemma 3.2.*

*Proof.* Let $M$ be a fixed $n \times m$ matrix of rank $r$. Let $\boldsymbol{Y} = \boldsymbol{A}(M + \boldsymbol{B})$, where $\boldsymbol{A}$ is picked uniformly from $\mathrm{GL}(n,q)$ and $\boldsymbol{B}$ has a UGR distribution with rank distribution $\mathcal{R}$. Then

$$h_r = H(\boldsymbol{A}(M+\boldsymbol{B}) |\, \mathrm{rk}(M) = r) = H(\boldsymbol{Y}).$$

Since $\mathrm{Row}(\boldsymbol{Y})$ is fully determined by $\boldsymbol{Y}$, it follows that $H(\boldsymbol{Y}, \mathrm{Row}(\boldsymbol{Y})) = H(\boldsymbol{Y})$. Therefore, using the chain rule for entropy (e.g. [6, Thm. 2.2.1]), we have

$$\begin{aligned} H(\boldsymbol{Y}) &= H(\boldsymbol{Y}, \mathrm{Row}(\boldsymbol{Y})) \\ &= H(\boldsymbol{Y} |\, \mathrm{Row}(\boldsymbol{Y})) + H(\mathrm{Row}(\boldsymbol{Y})). \end{aligned} \tag{15}$$

Now, multiplying $(M + \boldsymbol{B})$ by a uniformly picked invertible matrix will result in a uniform matrix of the same rowspace as $(M + \boldsymbol{B})$. That is, the distribution of $\boldsymbol{Y}$ is uniform given the rowspace of $\boldsymbol{Y}$. Thus (see [6, Thm. 2.6.4])

$$\begin{aligned} H(\boldsymbol{Y} |\, \mathrm{Row}(\boldsymbol{Y}) = V) &= \log\left(|\{Y' : Y' \in \mathbb{F}_q^{n\times m}, \mathrm{Row}(Y') = V\}|\right) \\ &= \log\left(f_0(\dim(V))\right), \end{aligned} \tag{16}$$

where $f_0$ is as defined in Lemma 3.1. Therefore

$$\begin{aligned} H(\boldsymbol{Y} |\, \mathrm{Row}(\boldsymbol{Y})) &= \sum_{V \subseteq \mathbb{F}_q^m} \Pr(\mathrm{Row}(\boldsymbol{Y}) = V) H(\boldsymbol{Y} |\, \mathrm{Row}(\boldsymbol{Y}) = V) \\ &= \sum_{V \subseteq \mathbb{F}_q^m} \Pr(\mathrm{Row}(\boldsymbol{Y}) = V) \log\left(f_0(\dim(V))\right). \end{aligned} \tag{17}$$

Hence

$$\begin{aligned} h_r &= H(\boldsymbol{Y}) \\ &= H(\boldsymbol{Y} |\, \mathrm{Row}(\boldsymbol{Y})) + H(\mathrm{Row}(\boldsymbol{Y})) \\ &= \sum_{V \subseteq \mathbb{F}_q^m} \Pr(\mathrm{Row}(\boldsymbol{Y}) = V) \log\left(f_0(\dim(V))\right) \\ &\qquad\qquad - \sum_{V \subseteq \mathbb{F}_q^m} \Pr(\mathrm{Row}(\boldsymbol{Y}) = V) \log\left(\Pr(\mathrm{Row}(\boldsymbol{Y}) = V)\right) \\ &= \sum_{V \subseteq \mathbb{F}_q^m} \Pr(\mathrm{Row}(\boldsymbol{Y}) = V) \log\left(\frac{f_0(\dim(V))}{\Pr(\mathrm{Row}(\boldsymbol{Y}) = V)}\right). \end{aligned} \tag{18}$$

Now, we calculate the probability of $\boldsymbol{Y}$ having a particular rowspace $V$. Set $U = \mathrm{Row}(M)$, so $\dim U = r$. For $V \subseteq \mathbb{F}_q^m$, let $d_{UV} = \dim(U \cap V)$. Using the function $f_1$

defined in Lemma 3.2, we obtain the following result.

$$\begin{aligned}
&\Pr(\mathrm{Row}(\boldsymbol{Y}) = V) \\
&\quad = \Pr(\mathrm{Row}(M + \boldsymbol{B}) = V) \\
&\quad = \sum_{r_B=0}^{\min\{n,m\}} \Pr(\mathrm{rk}(\boldsymbol{B}) = r_B) \Pr(\mathrm{Row}(M + \boldsymbol{B}) = V | \,\mathrm{rk}(\boldsymbol{B}) = r_B) \\
&\quad = \sum_{r_B=0}^{\min\{n,m\}} \mathcal{R}(r_B) \frac{|\{B : \mathrm{rk}(B) = r_B, \mathrm{Row}(M + B) = V\}|}{|\mathbb{F}_q^{n\times m, r_B}|} \qquad (19) \\
&\quad = \sum_{r_B=0}^{\min\{n,m\}} \mathcal{R}(r_B) \frac{f_1(r, \dim(V), d_{UV}; r_B)}{|\mathbb{F}_q^{n\times m, r_B}|}, \qquad (20)
\end{aligned}$$

where (19) follows since $\boldsymbol{B}$ has a UGR distribution.

Substituting (20) into (18) we get

$$h_r = \sum_{V\subseteq \mathbb{F}_q^m} \left( \sum_{r_B=0}^{\min\{n,m\}} \mathcal{R}(r_B) \frac{f_1(r, \dim(V), d_{UV}; r_B)}{|\mathbb{F}_q^{n\times m, r_B}|} \right) \cdot \log \left( \frac{f_0(\dim(V))}{\sum_{r_B=0}^{\min\{n,m\}} \mathcal{R}(r_B) \frac{f_1(r, \dim(V), d_{UV}; r_B)}{|\mathbb{F}_q^{n\times m, r_B}|}} \right) \qquad (21)$$

In (21), for a given subspace $V \subseteq \mathbb{F}_q^m$, the corresponding term in the sum depends only on $\dim(V)$ and $d_{UV} = \dim(\mathrm{Row}(M) \cap V)$. Clearly $0 \le d_{UV} \le \min\{\dim U, \dim V\}$. Moreover, Corollary 2.4 implies that the number of spaces $V$ with $\dim(V) = v$ and $\dim(\mathrm{Row}(M) \cap V) = h$ for fixed integers $v$ and $h$ is

$$q^{(v-h)(r-h)} \begin{bmatrix} r \\ h \end{bmatrix}_q \begin{bmatrix} m - r \\ v - h \end{bmatrix}_q .$$

Combining this with (21) proves the theorem. □

Now we give the result of this section: an efficiently computable expression for the Gamma channel capacity as a maximisation over the set of possible input rank distributions.

**Corollary 6.2.** *The capacity of the Gamma channel $\Gamma(\mathcal{R})$ is given by*

$$C = \max_{\mathcal{R}_X} \left\{ \left( \sum_{r_Y=0}^{\min\{n,m\}} \mathcal{R}_{\boldsymbol{Y}}(r_Y) \log \left( \frac{|\mathbb{F}_q^{n\times m, r_Y}|}{\mathcal{R}_{\boldsymbol{Y}}(r_Y)} \right) \right) - \sum_{r_X=0}^{\min\{n,m\}} \mathcal{R}_{\boldsymbol{X}}(r_X) h_{r_X} \right\},$$

*where $h_{r_X}$ may be computed using Theorem 6.1, and $\mathcal{R}_{\boldsymbol{Y}}(r_Y)$ may be computed using Lemma 4.2.*

*Proof.* The capacity $C$ of the channel is defined to be

$$C = \max_{\mathcal{P}_{\boldsymbol{X}}} \{I(\boldsymbol{X}; \boldsymbol{Y})\} = \max_{\mathcal{P}_{\boldsymbol{X}}} \{H(\boldsymbol{Y}) - H(\boldsymbol{Y}|\boldsymbol{X})\}. \qquad (22)$$

By Theorem 5.5, to achieve capacity we can chose the input distribution $\mathcal{P}_{\boldsymbol{X}}$ to be UGR. By Lemma 5.4, the output distribution will also be UGR. Therefore the output distribution is given by

$$\mathcal{P}_{\boldsymbol{Y}}(Y) = \Pr(\boldsymbol{Y} = Y) = \frac{1}{|\mathbb{F}_q^{n\times m, \mathrm{rk}(Y)}|} \mathcal{R}_{\boldsymbol{Y}}(\mathrm{rk}(Y)) \qquad (23)$$

for any $Y \in \mathbb{F}_q^{n\times m}$. Thus the entropy of $\boldsymbol{Y}$ is given by

$$\begin{aligned} H(\boldsymbol{Y}) &= -\sum_{Y\in\mathbb{F}_q^{n\times m}} \Pr(\boldsymbol{Y}=Y)\log\Pr(\boldsymbol{Y}=Y) \\ &= -\sum_{Y\in\mathbb{F}_q^{n\times m}} \left(\frac{1}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(Y)}|}\mathcal{R}_{\boldsymbol{Y}}(\mathrm{rk}(Y))\right)\log\left(\frac{1}{|\mathbb{F}_q^{n\times m,\mathrm{rk}(Y)}|}\mathcal{R}_{\boldsymbol{Y}}(\mathrm{rk}(Y))\right) \\ &= -\sum_{r_Y=0}^{\min\{n,m\}} \mathcal{R}_{\boldsymbol{Y}}(r_Y)\log\left(\frac{1}{|\mathbb{F}_q^{n\times m,r_Y}|}\mathcal{R}_{\boldsymbol{Y}}(r_Y)\right). \end{aligned}$$

Since $H(\boldsymbol{Y}|\boldsymbol{X}) = \sum_{r_X=0}^{\min\{n,m\}}\mathcal{R}_{\boldsymbol{X}}(r_X)h_{r_X}$ by Lemma 5.2, the result follows from (22).
□

# 7 Matrix function proofs

The aim of this section is to derive efficiently computable expressions for the functions $f_1$ and $f_2$, thus proving Lemmas 3.2 and 3.3 respectively and providing a method for computing the capacity formula given in Corollary 6.2.

We approach this problem by first exploring several combinatorial results. In Subsection 7.1 we establish a counting result we need later, using Möbius theory. In Subsection 7.2, we use this result to derive expressions for the functions $f_1$ and $f_2$.

## 7.1 A counting lemma

In this subsection, we prove an 'inversion' lemma, Lemma 7.1, that we need in the following subsection. We use Möbius theory (a generalisation of inclusion–exclusion) to establish this lemma: see Bender and Goldman [2], for example, for a nice introduction to this theory and an exposition of all the results we use here.

Let $\mathrm{Po}(\mathbb{F}_q^m)$ denote the poset of all subspaces of $\mathbb{F}_q^m$ ordered by containment. Let $P$ and $Q$ be two posets. Recall that the direct product $P\times Q$ is the poset where $(p_1,q_1)\leq(p_2,q_2)$ if and only if $p_1\leq p_2$ and $q_1\leq q_2$, where $p_1,p_2\in P$ and $q_1,q_2\in Q$.

**Lemma 7.1.** *Let $f((U,V))$ be a real valued function defined for all pairs $(U,V)\in \mathrm{Po}(\mathbb{F}_q^m)\times\mathrm{Po}(\mathbb{F}_q^m)$. If*

$$g((U,V)) = \sum_{(U',V')\leq(U,V)} f((U',V'))$$

*then*

$$f((U,V)) = \sum_{(U',V')\leq(U,V)} (-1)^{u-u'+v-v'}q^{\binom{u-u'}{2}+\binom{v-v'}{2}}g((U',V')),$$

*where* $\dim(U)=u, \dim(U')=u', \dim(V)=v$ *and* $\dim(V')=v'$.

*Proof.* By the Möbius inversion formula (see [2, Theorem 1], for example)

$$f((U,V)) = \sum_{(U',V')\leq(U,V)} \mu((U',V'),(U,V))g((U',V')), \tag{24}$$

where $\mu$ is the Möbius function of $\mathrm{Po}(\mathbb{F}_q^m)\times\mathrm{Po}(\mathbb{F}_q^m)$. But (see [2, Theorem 3], for example),

$$\mu((U',V'),(U,V)) = \mu'(U',U)\mu'(V',V)$$

where $\mu'$ is the Möbius function of $\mathrm{Po}(\mathbb{F}_q^m)$. Moreover (see, for example, [2, §5]) the Möbius function of $\mathrm{Po}(\mathbb{F}_q^m)$ may be written explicitly as

$$\mu'(X,Y) = (-1)^{\dim(Y)-\dim(X)}q^{\binom{\dim(Y)-\dim(X)}{2}}$$

for any $X,Y\in\mathrm{Po}(\mathbb{F}_q^m)$ with $X\subseteq Y$. So the lemma follows. □

## 7.2 Computing $f_1$ and $f_2$

By 'basic dimension properties', we mean that all specified dimensions are non-negative integers, and if dimensions $d_U$ and $d_V$ of subspaces $U \subseteq V$ are specified, then $d_U \leq d_V$.

**Lemma 7.2.** *Let $z$ be a non-negative integer. Let $X$ and $Y$ be subspaces of $\mathbb{F}_q^z$ of dimensions $d_X$ and $d_Y$ respectively such that $X \cap Y = \{0\}$. Let $c(d_X, d_Y, z, d_R, d_{RX}, d_{RY})$ be the number of $d_R$-dimensional subspaces $R$ of $\mathbb{F}_q^z$ such that $\dim(R \cap X) = d_{RX}$, $\dim(R \cap Y) = d_{RY}$ and such that $X \subseteq Y + R$. If the basic dimension properties are satisfied then $c(d_X, d_Y, z, d_R, d_{RX}, d_{RY})$ is given by the formula*

$$\begin{bmatrix} d_X \\ d_{RX} \end{bmatrix}_q \begin{bmatrix} d_Y \\ d_{RY} \end{bmatrix}_q \begin{bmatrix} z - d_X - d_Y \\ d_R - d_{RY} - d_X \end{bmatrix}_q q^{(d_Y - d_{RY})(d_R - d_X - d_{RY})} \prod_{i=0}^{d_X - d_{RX} - 1} (q^{d_Y - d_{RY}} - q^i),$$

*otherwise $c(d_X, d_Y, z, d_R, d_{RX}, d_{RY}) = 0$.*

We remark that, in this case, the basic dimension properties are that $0 \leq d_{RX} \leq \min\{d_R, d_X\}$, $0 \leq d_{RY} \leq \min\{d_R, d_Y\}$ and $d_X + d_Y \leq z$.

*Proof.* Suppose that $d_X > d_R - d_{RY}$. The condition that $X \subseteq Y + R$ is equivalent to the condition that the subspace $(X+Y)/Y$ is contained in the subspace $(R+Y)/Y$. The dimensions of these subspaces are $d_X$ and $d_R - d_{RY}$ respectively, and so our count is zero in this case. But $\left[ {z - d_X - d_Y \atop d_R - d_{RY} - d_X} \right]_q = 0$ when $d_R - d_{RY} - d_X < 0$ and so the lemma follows in this case. So we may assume that $d_X \leq d_R - d_{RY}$.

There are $\left[ {d_X \atop d_{RX}} \right]_q$ choices for the subspace $R \cap X$ and $\left[ {d_Y \atop d_{RY}} \right]_q$ choices for the subspace $R \cap Y$. Assume that these subspaces are fixed. The quotient $Q = (R+Y)/Y$ of $R$ in $\mathbb{F}_q^z/Y$ has dimension $(d_R - d_{RY})$. The condition that $X \subseteq Y + R$ implies that $(X+Y)/Y \subseteq Q$. Since $(X+Y)/Y$ has dimension $d_X$, the number of choices for $Q$ is therefore $\left[ {z - d_Y - d_X \atop d_R - d_{RY} - d_X} \right]_q$. Assume that $Q$ is now also fixed.

Fix $u_1, u_2, \ldots, u_{d_R - d_{RY} - d_X} \in \mathbb{F}_q^m$ with the property that $\{u_i + Y : 1 \leq i \leq d_R - d_{RY} - d_X\}$ spans a complement to $(X+Y)/Y$ in $Q$. Fix a basis $x_1, \ldots, x_{d_{RX}}$ of $R \cap X$, and extend this basis to a basis $x_1, x_2, \ldots, x_{d_X}$ of $X$. Fix a basis $y_1, y_2, \ldots, y_{d_{RY}}$ of $R \cap Y$. Every subspace $R$ we are counting has a basis of the form

$$\{y_i : 1 \leq i \leq d_{RY}\} \cup \{x_i : 1 \leq i \leq d_{RX}\} \cup \{x_i + \epsilon_i : d_{RX} + 1 \leq i \leq d_X\} \\ \cup \{u_i + \delta_i : 1 \leq i \leq d_R - d_{RY} - d_X\}$$

for some $\epsilon_i, \delta_i \in Y$. Note that all subspaces with a basis of this form intersect $Y$ in precisely the space spanned by $\{y_i : 1 \leq i \leq d_{RY}\}$, and all subspaces are equal to $Q$ after taking a quotient by $Y$. Moreover, a subspace of this form intersects $X$ in precisely the subspace spanned by $\{x_i : 1 \leq i \leq d_{RX}\}$ if and only if the vectors $\epsilon_i + (R \cap Y)$ are linearly independent in $Y/(R \cap Y)$. Finally, two subspaces of this form are distinct if and only if the ordered set of vectors $\epsilon_i$ and $\delta_i$ are different modulo $R \cap Y$. There are $q^{(d_Y - d_{RY})(d_R - d_X - d_{RY})}$ choices for vectors $\delta_i + (R \cap Y) \in Y/(R \cap Y)$, and there are $\prod_{i=0}^{d_X - d_{RX} - 1} (q^{d_Y - d_{RY}} - q^i)$ choices for linearly independent vectors $\epsilon_i + (R \cap Y) \in Y/(R \cap Y)$. So the lemma follows. □

We define a function $f_1(d_U, d_V, d_{UV}; r)$ as follows. When $r < (d_U - d_{UV})$, we define $f_1(d_U, d_V, d_{UV}; r) = 0$. Otherwise we proceed as follows. For integers $d_{V'}$, $d_{W'}$ and $d_{V'W'}$, define

$$\kappa_1(d_{V'}, d_{W'}, d_{V'W'}) = (-1)^{(r - d_{W'})} q^{\binom{r - d_{W'}}{2}} (-1)^{(d_V - d_{V'})} q^{\binom{d_V - d_{V'}}{2}} q^{n d_{V'W'}}.$$

For integers $d_{V'}$, $d_{W'}$, $d_{V'W'}$, $d_{UW'}$, $d_{VW'}$ and $d_{UVW'}$, define

$$\kappa_2(d_{V'}, d_{W'}, d_{V'W'}, d_{UW'}, d_{VW'}, d_{UVW'}) = \nu_1 \nu_2 \nu_3$$

where

$$\nu_1 = c(d_U - d_{UV}, d_V - d_{UV}, m - d_{UV}, d_{W'} - d_{UVW'}, d_{UW'} - d_{UVW'}, d_{VW'} - d_{UVW'})$$
$$\nu_2 = \begin{bmatrix} d_{UV} \\ d_{UVW'} \end{bmatrix}_q \begin{bmatrix} d_{VW'} \\ d_{V'W'} \end{bmatrix}_q \begin{bmatrix} (d_V - d_{VW'}) - (d_U - d_{UW'}) \\ (d_{V'} - d_{V'W'}) - (d_U - d_{UW'}) \end{bmatrix}_q \begin{bmatrix} m - d_{W'} \\ r - d_{W'} \end{bmatrix}_q \text{ and}$$
$$\nu_3 = q^{(d_{W'} - d_{UVW'})(d_{UV} - d_{UVW'})} q^{(d_{VW'} - d_{V'W'})(d_{V'} - d_{V'W'})},$$

where $c$ is the function defined in Lemma 7.2. Then $f_1(d_U, d_V, d_{UV}; r)$ is equal to

$$\sum_{d_{W'}=0}^{r} \sum_{d_{V'}=0}^{d_V} \sum_{d_{V'W'}=0}^{\min\{d_{V'}, d_{W'}\}} \sum_{d_{UW'}=0}^{\min\{d_U, d_{W'}\}} \sum_{d_{VW'}=0}^{\min\{d_V, d_{W'}\}} \sum_{d_{UVW'}=0}^{\min\{d_{UV}, d_{VW'}, d_{UW'}\}} \kappa_1 \kappa_2,$$

where $\kappa_1 = \kappa_1(d_{V'}, d_{W'}, d_{V'W'})$ and $\kappa_2 = \kappa_2(d_{V'}, d_{W'}, d_{V'W'}, d_{UW'}, d_{VW'}, d_{UVW'})$.

**Theorem 7.3.** *Let $f_1$ be as defined in Lemma 3.2. That is, if $U$ and $V$ are subspaces of $\mathbb{F}_q^m$ of dimensions $d_U$ and $d_V$ respectively, with $d_{UV} = \dim(U \cap V)$ and $M \in \mathbb{F}_q^{n\times m}$ is a fixed matrix such that $\mathrm{Row}(M) = U$; then $f_1(d_U, d_V, d_{UV}; r)$ gives the number of matrices $B \in \mathbb{F}_q^{n\times m, r}$ such that $\mathrm{Row}(M+B) = V$. Then the value $f_1(d_U, d_V, d_{UV}; r)$ is as given above.*

*Proof.* We begin the proof with a simpler counting problem, and then use this result to establish the formula we are aiming for.

For a subspace $W$ of $\mathbb{F}_q^m$, let $g(V, W)$ be the number of $n \times m$ matrices $B$ with $\mathrm{Row}(B) \subseteq W$ and $\mathrm{Row}(M + B) \subseteq V$. We claim that

$$g(V, W) = \begin{cases} q^{n\, d_{VW}} & \text{if } U \subseteq V + W \\ 0 & \text{otherwise.} \end{cases}$$

To see this, we proceed as follows. Let $x_1, x_2, \ldots, x_n \in \mathbb{F}_q^m$ be the rows of $M$. Suppose that $U \not\subseteq V + W$. Then $(x_i + W) \cap V = \emptyset$ for some $i$, and so we must have $g(V, W) = 0$, since there is no valid choice for the $i$th row of $B$ in this case. Now suppose that $U \subseteq V + W$, so for all $i$ we have $(x_i + W) \cap V \neq \emptyset$ and therefore there exist $w_1, w_2, \ldots, w_n \in W$ such that $x_i + w_i \in V$. It is not hard to check that a matrix $B$ with rows $b_i$ has the property that $\mathrm{Row}(B) \subseteq W$ and $\mathrm{Row}(M + B) \subseteq V$ if and only if $b_i - w_i \in V \cap W$. Hence there are $q^{d_{VW}}$ choices for each row $b_i$ of $B$. Since $B$ has $n$ rows, the claim follows.

Let $f(V, W)$ be the number of $n \times m$ matrices $B$ with $\mathrm{Row}(B) = W$ and $\mathrm{Row}(M + B) = V$. Now $g(V, W) = \sum_{(V',W')} f(V, W)$, where the sum is over all pairs of subspaces $(V', W')$ with $V' \subseteq V$ and $W' \subseteq W$. So, by Lemma 7.1,

$$\begin{aligned} f(V, W) &= \sum_{(V',W')} (-1)^{(d_W - d_{W'}) + (d_V - d_{V'})} q^{\binom{d_W - d_{W'}}{2} + \binom{d_V - d_{V'}}{2}} g(V', W') \\ &= \sum_{\substack{(V',W') \\ U \subseteq V' + W'}} (-1)^{(d_W - d_{W'}) + (d_V - d_{V'})} q^{\binom{d_W - d_{W'}}{2} + \binom{d_V - d_{V'}}{2}} q^{n d_{V'W'}} \\ &= \sum_{\substack{(V',W') \\ U \subseteq V' + W'}} \kappa_1(d_{V'}, d_{W'}, d_{V'W'}), \end{aligned}$$

where again $V' \subseteq V$ and $W' \subseteq W$ in our sums.

The number of matrices $B$ of rank $r$ such that $\mathrm{Row}(M + B) = V$ is

$$\sum_{\substack{W \subseteq \mathbb{F}_q^m \\ \dim W = r}} f(V, W).$$

So we can express this count as

$$\sum_{d_{W'}=0}^{r}\sum_{d_{V'}=0}^{d_V}\sum_{d_{V'W'}=0}^{\min\{d_{V'},d_{W'}\}}\sum_{d_{UW'}=0}^{\min\{d_U,d_{W'}\}}\sum_{d_{VW'}=0}^{\min\{d_V,d_{W'}\}}\sum_{d_{UVW'}=0}^{\substack{\min\{d_{UV},d_{VW'},\\ d_{UW'}\}}}\sum_{V',W',W}\kappa_1(d_{V'},d_{W'},d_{V'W'}), \tag{25}$$

where the last sum is over all triples $(V',W',W)$ of subspaces of $\mathbb{F}_q^m$ with $V'\subseteq V$, $W'\subseteq W$, $U\subseteq V'+W'$, $\dim(W')=d_{W'}$, $\dim(W)=r$, $\dim(V')=d_{V'}$, $\dim(V'\cap W')=d_{V'W'}$, $\dim(U\cap W')=d_{UW'}$, $\dim(V\cap W')=d_{VW'}$ and $\dim U\cap V\cap W'=d_{UVW'}$.

We aim to count the number of possibilities for a subspace $W'$ such that $\dim(W')=d_{W'}$, $\dim(U\cap W')=d_{UW'}$, $\dim(V\cap W')=d_{VW'}$ and $\dim U\cap V\cap W'=d_{UVW'}$ and that satisfy the weaker condition that $U\subseteq V+W'$. We will show (see below) that the number of such subspaces $W'$ is

$$c\begin{bmatrix} d_{UV} \\ d_{UVW'} \end{bmatrix}_q q^{(d_{W'}-d_{UVW'})(d_{UV}-d_{UVW'})}, \tag{26}$$

where $c=c(d_U-d_{UV},d_V-d_{UV},m-d_{UV},d_{W'}-d_{UVW'},d_{UW'}-d_{UVW'},d_{VW'}-d_{UVW'})$ is defined in Lemma 7.2.

Once we have fixed such a subspace $W'$, we choose $V'$ and $W$ as follows. We first choose the subspace $V'\cap W'$. There are $\left[\begin{smallmatrix} d_{VW'} \\ d_{V'W'} \end{smallmatrix}\right]_q$ choices for this subspace. The quotient space $(V'+W')/W'$ of $V'$ by $W'$ is a space of dimension $d_{V'}-d_{V'W'}$. It is contained in the $(d_V-d_{VW'})$-dimensional space $(V+W')/W'$ and contains the $(d_U-d_{UW'})$-dimensional space $(U+W')/W'$. So the number of choices for $(V'+W')/W'$ is

$$\begin{bmatrix} (d_V-d_{VW'})-(d_U-d_{UW'}) \\ (d_{V'}-d_{V'W'})-(d_U-d_{UW'}) \end{bmatrix}_q.$$

Once this quotient space is also fixed, there are $q^{(d_{VW'}-d_{V'W'})(d_{V'}-d_{V'W'})}$ choices for $V'$. Finally we choose the $r$-dimensional subspace $W$ containing $W'$: there are $\left[\begin{smallmatrix} m-d_{W'} \\ r-d_{W'} \end{smallmatrix}\right]_q$ choices for $W$.

Combining the formula (26) with (25) and the counting argument of the previous paragraph, the theorem follows. So it remains to establish (26).

The number of choices (26) for $W'$ may be found as follows. There are $\left[\begin{smallmatrix} d_{UV} \\ d_{UVW'} \end{smallmatrix}\right]_q$ choices for the subspace $T=(U\cap V)\cap W'$. Suppose that $T$ is now fixed. We now consider the images $X$, $Y$ and $R$ of $U$, $V$ and $W'$ respectively in the quotient by $U\cap V$. So $X=(U+(U\cap V))/(U\cap V)$ has dimension $d_U-d_{UV}$ and $Y=(V+(U\cap V))/(U\cap V)$ has dimension $d_V-d_{UV}$. Moreover $R$ is a subspace of dimension $d_{W'}-d_{UVW'}$ which intersects $X$ and $Y$ in subspaces of dimension $d_{UW'}-d_{UVW'}$ and $d_{VW'}-d_{UVW'}$ respectively. The subspaces $X$ and $Y$ intersect trivially. Since $U\subseteq V+W'$, we see that $X\subseteq Y+R$. Hence, by Lemma 7.2, the number of choices for the subspace $R$ is $c(d_U-d_{UV},d_V-d_{UV},m-d_{UV},d_{W'}-d_{UVW'},d_{UW'}-d_{UVW'},d_{VW'}-d_{UVW'})$. Suppose now that $R$ is fixed. There are $q^{(d_{W'}-d_{UVW'})d_{UVW'}}$ subspaces $W'$ with $(W'+(U\cap V))/(U\cap V)=R$ and $(U\cap V)\cap W'=T$. Since all of these subspace have the property that $U\subseteq W'+V$, the formula (26) follows, and so the theorem is proved. □

**Theorem 7.4.** *Let $f_2$ be as defined in Lemma 3.3. That is, for a fixed matrix $X$ with $\mathrm{rk}(X)=r_X$, $f_2(r,r_X,r_B)$ gives the number of matrices $B\in\mathbb{F}_q^{n\times m,r_B}$ such that $\mathrm{rk}(X+B)=r$. Then,*

$$f_2(r,r_X,r_B)=\sum_{h=0}^{\min\{r,r_X\}} q^{(r-h)(r_X-h)}\begin{bmatrix} m-r_X \\ r-h \end{bmatrix}_q\begin{bmatrix} r_X \\ h \end{bmatrix}_q f_1(r_X,r,h;r_B).$$

*Proof.* Using the definition of $f_1$ given above Theorem 7.3, we see that

$$\begin{aligned} & f_2(r, r_X, r_B) \\ &= \sum_{V \subseteq \mathbb{F}_q^m : \dim(V) = r} f_1(r_X, r, \dim(V \cap \mathrm{Row}(X)); r_B) \qquad (27) \\ &= \sum_{h=0}^{\min\{r, r_X\}} |\{V \subseteq \mathbb{F}_q^m : \dim(V) = r, \dim(V \cap \mathrm{Row}(X)) = h\}| f_1(r_X, r, h; r_B) \quad (28) \end{aligned}$$

where (27) follows since the number of matrices $B$ with $\mathrm{rk}(X + B) = r$ is equal to the number of matrices $B$ with $\mathrm{Row}(X + B) = V$, summed over all spaces $V \subseteq \mathbb{F}_q^m$ with $\dim(V) = r$.

By Corollary 2.4, the number of $r$-dimensional subspaces $V \subseteq \mathbb{F}_q^m$, with $\dim(V \cap \mathrm{Row}(X)) = h$ is

$$q^{(r-h)(r_X-h)} \begin{bmatrix} m - r_X \\ r - h \end{bmatrix}_q \begin{bmatrix} r_X \\ h \end{bmatrix}_q. \qquad (29)$$

Substituting (29) into (28) gives the result. □

# 8 Conclusion

In this paper we have considered a class of matrix channels (Gamma channels) suitable for modelling random linear network coding when random errors are introduced during transmission. The Gamma channels are a generalisation of the AMMC channel considered in [18]. Random errors are modelled by a matrix whose rank represents the number of linearly independent errors. The error matrix is chosen by first picking its rank according to a rank distribution $\mathcal{R}$ dependent on the application, and then choosing uniformly from all matrices of this rank (a UGR distribution). We show that in this model there always exists a capacity achieving input distribution that is UGR. This key result allows us to compute the capacity of the channel as a maximisation problem over possible (input) rank distributions, a set of linear rather than exponential size. We presented sample capacity computations in the introduction: all computations used a simple hill-climbing algorithm to perform the maximisation, and were implemented in Mathematica 10.4 [11].

**Open Problem 1.** Can bounds for the AMMC capacity be improved, to give good asymptotic results in more situations?

We ran simulations to show that for the AMMC channel with two errors, the true capacity of the channel closely follows the trend of the previously known upper bound for the capacity. It might be possible to improve the lower bound on the capacity by using simulation results as a guide.

**Open Problem 2.** Can good asymptotic bounds on the capacity of the Gamma channel be established?

We believe it will be hard to find good capacity bounds that hold in complete generality. But it would be very interesting to investigate the binomial rank distribution for errors, or the distribution arising for errors that are not linearly independent mentioned in the introduction. Many natural error rank distributions, such as those just mentioned, cluster around a mean value $\mu$. For such distributions, and when $n \leq m$, we believe that the capacity of the Gamma channel (in $q$-ary units) should be approximately $\tilde{c}$ where

$$\tilde{c} = \begin{cases} (m - n)(n - \mu) & \text{when } \mu \geq 2n - m, \\ (m - \mu)^2/4 & \text{otherwise.} \end{cases}$$

Moreover, in these cases both experiments and heuristic arguments lead us to believe that a capacity-achieving input rank distribution peaks at one or two values close to $\tilde{k}$, where

$$\tilde{k} = \begin{cases} n - \mu & \text{when } \mu \geq 2n - m, \\ (m-\mu)/2 & \text{otherwise.} \end{cases}$$

It might be possible to use Corollary 6.2 to prove asymptotic capacity results, but this corollary is designed for precision rather than asymptotics. A more promising approach would be to further develop the theory in Silva et al [18]. (We note that work is needed to establish asymptotically tight bounds even in the case when the error rank is bounded by $t$. For example, the lower bound on the capacity in [18, VI.D] will not be tight when the error rank is normally significantly smaller than $t$.)

**Open Problem 3.** Can explicit good coding schemes for the Gamma channel be constructed?

Theorem 5.5 shows that there are UGR input distributions that achieve capacity. It would be interesting to see explicit good coding schemes that use UGR input distributions. (We are not aware of such schemes, even in special cases such as the AMMC channel.)

## Acknowledgements

Jessica Claridge would like to acknowledge the support of an EPSRC PhD studentship. Both authors would like to acknowledge the support of the EU COST Action IC1104, and would like thanks reviewers of an earlier version of this paper for their comments.

## References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network information flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, Jul 2000.

[2] E. A. Bender and J. R. Goldman. On the applications of Möbius inversion in combinatorial analysis. *The American Mathematical Monthly*, 82(8):789–803, October 1975.

[3] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge university press, 2004.

[4] P. J. Cameron. *Combinatorics: Topics, Techniques, Algorithms.* Cambridge University Press, 1994.

[5] J. Claridge. *On Matrix Models for Network Coding.* PhD Thesis, Royal Holloway, University of London, 2017.

[6] T. M. Cover and J. A. Thomas. *Elements of information theory.* John Wiley & Sons, 2012.

[7] È. M. Gabidulin. Theory of codes with maximum rank distance. *Problems of Information Transmission*, 21(1):1–12, January 1985.

[8] M. Gadouleau and Z. Yan. Packing and covering properties of rank metric codes. *IEEE Transactions on Information Theory*, 54(9):3873–3883, 2008.

[9] M. Gadouleau and Z. Yan. Bounds on covering codes with the rank metric. *IEEE Communications Letters*, 13(9):691–693, 2009.

[10] T. Ho, M. Médard, R. Kötter, D. R. Karger, M. Effros, J. Shi, and B. Leong. A random linear network coding approach to multicast. *IEEE Transactions on Information Theory*, 52(10):4413–4430, October 2006.

[11] Wolfram Research, Inc. Mathematica, Version 10.4. Champaign, IL, 2016.

[12] R. Kötter and F. R. Kschischang. Coding for errors and erasures in random network coding. *IEEE Transactions on Information Theory*, 54(8):3579–3591, Aug 2008.

[13] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear network coding. *IEEE Transactions on Information Theory*, 49(2):371–381, 2003.

[14] A. Montanari and R. L. Urbanke. Iterative coding for network coding. *IEEE Transactions on Information Theory*, 59(3):1563–1572, 2013.

[15] R. W. Nobrega, D. Silva, and B. F. Uchoa-Filho. On the capacity of multiplicative finite-field matrix channels. *IEEE Transactions on Information Theory*, 59(8):4949–4960, 2013.

[16] M. J. Siavoshani, S. Mohajer, C. Fragouli, and S. N. Diggavi. On the capacity of noncoherent network coding. *IEEE Transactions on Information Theory*, 57(2):1046–1066, 2011.

[17] D. Silva, F. R. Kschischang, and R. Kötter. A rank-metric approach to error control in random network coding. *IEEE Transactions on Information Theory*, 54(9):3951–3967, 2008.

[18] D. Silva, F. R. Kschischang, and R. Kötter. Communication over finite-field matrix channels. *IEEE Transactions on Information Theory*, 56(3):1296–1305, March 2010.