On the Capacity of Multiplicative 
Finite-Field Matrix Channels

Roberto W. Nóbrega, Bartolomeu F. Uchôa-Filho, Danilo Silva 
Department of Electrical Engineering
Federal University of Santa Catarina, Brazil
{rwnobrega, uchoa, danilo}@eel.ufsc.br

Abstract—This paper deals with the multiplicative finite-field matrix channel, a discrete memoryless channel whose input and output are matrices (over a finite field) related by a multiplicative transfer matrix. Our model allows this transfer matrix to have any rank, while assuming that all transfer matrices with the same rank are equiprobable. While in general the capacity cannot be obtained in closed form, we provide a simplification of the problem (from $q^{n^m}$ to $O(n)$ variables) which allows for easy numerical computation. A tight upper bound on the capacity is also derived, and for the special case of constant-rank input, we obtain an exact formula. Several existing results can be obtained as special cases of our approach. In addition, we prove that the well-known approach of treating inputs and outputs as subspaces is information-lossless even in this more general case.

I. INTRODUCTION

Finite-field matrix channels are communication channels where both the input and the output are matrices over a finite field $\mathbb{F}_q$. The interest in such channels has been rising since the seminal work of Koetter and Kschischang [1], which connects finite-field matrix channels to the problem of error control in non-coherent linear network coding. In contrast with the combinatorial framework of [1], the present paper follows the work of Silva et al. [2] and adopts a probabilistic approach.

The object of study of this work is the multiplicative matrix channel, in which the input and output matrices (both of dimension $n \times n$) are related by a random multiplicative matrix (of dimension $n \times n$), called the transfer matrix. We consider the input matrix and the transfer matrix to be statistically independent, so that the probability distribution of the transfer matrix completely characterizes the channel. This model turns out to be well-suited for random linear network coding systems [3] in the absence of malicious nodes, but subject to link erasures. The probability distribution for the transfer matrix is then dictated by the network topology, the random choices of coding coefficients, and the link erasure probabilities.

The capacity of the multiplicative finite-field matrix channel was derived exactly in [2] under the assumption that the transfer matrix is chosen uniformly at random among all full-rank matrices. In addition, assuming that the transfer matrix has i.i.d. entries selected uniformly at random, Jafari et al. [4] obtained the channel capacity as a solution of a convex optimization problem over $O(n)$ variables; they also found an exact expression for the capacity when the field size is greater than a certain threshold.

On the other hand, Yang et al. [5] (also [6], [7]) consider a completely general scenario, making no assumptions on the distribution of the transfer matrix. For general input, they obtained upper and lower bounds on the channel capacity, while for subspace-coded input, they were able to reduce the number of optimization variables of the original capacity problem—although to a number of variables that is still exponential in the matrix size.

We herein consider a narrower situation than that addressed in [5], but of which both [2] and [4] are special cases. Specifically, we allow the probability distribution of the rank of the transfer matrix to be arbitrary; nevertheless we consider that all matrices with the same rank are equiprobable. Under this assumption, the probability distribution of the rank of the transfer matrix completely determines the distribution of the transfer matrix itself and, therefore, also completely determines the channel. We believe that our model represents a good trade-off between the previously considered ones: it allows a broader applicability in non-coherent network coding when compared to [2] and [4] (especially for small network parameters), but still requires $O(n)$ parameters to describe the channel, as opposed to [5], which requires $O(q^{n^2})$.

The present work—which is based on some of our earlier ideas in [8]—concentrates on the problem of finding the capacity and mutual information of the multiplicative finite-field matrix channel under the aforementioned conditions. The mutual information is obtained under the assumption that the input matrix (similarly to the transfer matrix) is uniformly distributed conditioned on its rank. We show that this restriction is still enough to achieve the capacity. As a consequence, we are able to greatly reduce the complexity of the convex optimization problem involved in obtaining the channel capacity and the associated optimal input. We then specialize our results to match the models adopted in [2] and [4]. We also consider the special case of constant-rank input, which is of great theoretical and practical significance; in this case, we are able to obtain a closed-form expression for the channel capacity.

As a final contribution, this work presents yet another justification for the idea of communication via subspaces, first suggested in [1]. Our results in this line generalize similar conclusions previously obtained in [2] and [4], by proving that the transmitter should encode information in the subspace of the transmitted matrix, while the receiver needs only to consider...
the subspace of the received matrix. This clearly simplifies the encoding and decoding operations. It is worth noting that, akin to [2] and [4], our result is valid for any choice of channel parameters. This contrasts with the optimality result derived in [6], which holds only asymptotically in the packet size.

Proofs are omitted throughout the text and can be found in [9].

II. BACKGROUND AND NOTATION

We denote by \( T_{q}^{n\times m,k} \) the set of all \( n \times m \) matrices over \( \mathbb{F}_{q} \) with rank \( k \). For notational convenience, we sometimes set \( T_{k} = T_{q}^{n \times m,k} \) when the matrix dimension \( n \times m \) and the field size \( q \) are implied by the context. Also, \( T_{q}^{n\times m} \) is the set of all \( n \times m \) full-rank matrices. Let

\[
\left[ \begin{array}{c} \frac{m}{k} \end{array} \right]_{q} \triangleq \prod_{i=0}^{k-1} \frac{q^{m} - q^{i}}{q^{k} - q^{i}},
\]

for \( 0 \leq k \leq m \), denote the Gaussian binomial coefficient. Finally, \( \langle A \rangle \) denotes the row space of \( A \).

A. Some Enumeration Results

We start by presenting some matrix enumeration results that will prove to be useful in the rest of the paper. For a proof of the following fact see, e.g., [10].

**Fact 1:** The number of \( n \times m \) matrices over \( \mathbb{F}_{q} \) with rank \( k \) is given by

\[
|T_{q}^{n \times m,k}| = \frac{|T_{q}^{n \times k}| |T_{q}^{m \times k}|}{|T_{q}^{k \times k}|} = \left[ \begin{array}{c} m \\ k \end{array} \right]_{q}^{k-1} \prod_{i=0}^{k-1} \left( q^{m} - q^{i} \right),
\]

where \( |T_{q}^{n \times m}| = \prod_{i=0}^{n-1} (q^{m} - q^{i}) \) is the number of full-rank \( n \times m \) matrices.

The next fact is a combinatorial result by Brawley and Carlitz [11].

**Fact 2:** Let \( A_{0} \) be an arbitrary \( s \times n_{0} \) matrix of rank \( k_{0} \) over the finite field \( \mathbb{F}_{q} \). Then, the number of \( s \times n \) matrices of rank \( k \) over \( \mathbb{F}_{q} \) whose first \( n_{0} \) columns are the columns of \( A_{0} \) is given by

\[
\phi_{q}(s; n_{0}, n, k_{0}, k) = \left[ \begin{array}{c} N \\ K \end{array} \right]_{q}^{k_{0}(N - K)} \prod_{i=0}^{K-1} (q^{s} - q^{k_{0}+i}),
\]

where \( N = n - n_{0} \) and \( K = k - k_{0} \).

B. Discrete Memoryless Channels

A **discrete memoryless channel** (DMC) [12] with input \( X \) and output \( Y \) is defined by a triplet \( (\mathcal{X}, p_{Y\mid X}, \mathcal{Y}) \), where \( \mathcal{X} \) and \( \mathcal{Y} \) are the channel inputs and output alphabets, respectively, and \( p_{Y\mid X}(Y\mid X) \) is the conditional probability that \( Y \in \mathcal{Y} \) is received given that \( X \in \mathcal{X} \) is sent.\(^{1}\) The channel is memoryless in the sense that what happens to the transmitted symbol at one time is independent of what happens to the transmitted symbol at any other time. The **capacity** of the DMC is then given by

\[
C = \max_{p_{X}} I(X; Y),
\]

where \( I(X; Y) \) is the mutual information between \( X \) and \( Y \). The maximization is over all possible input distributions \( p_{X}(X) \).

III. PROBLEM STATEMENT

The **multiplicative finite-field matrix channel** is modeled by the channel law

\[
Y = GX,
\]

where \( X \) is the \( n_{1} \times m \) channel input matrix, \( Y \) is the \( n_{o} \times m \) channel output matrix, \( G \) is the \( n_{o} \times n_{1} \) transfer matrix, and all matrices are over a given finite field \( \mathbb{F}_{q} \). Being memoryless, the channel is fully described by the conditional probability distribution \( p_{Y\mid X}(Y\mid X) \), which in turn, in view of (1), is induced by the conditional probability distribution \( p_{G\mid X}(G\mid X) \).

It is customary to consider \( G \) and \( X \) independent, in which case \( p_{G\mid X}(G\mid X) = p_{G}(G) \). Following [2], we assume for simplicity \( n_{o} = n_{1} = n \) and \( m \leq m \).

This work deals with a special class of this channel, in which the transfer matrix \( G \) is said to be “uniform given rank,” a concept defined next.

**Definition 1:** A random matrix is called **uniform given rank** (u.g.r., for short) if any two matrices with the same rank are equiprobable.

Let \( A \) be any random matrix over \( \mathbb{F}_{q}^{n \times m} \). Also, let \( k = \text{rank } A \); this is a random variable taking values on \( \{0, \ldots, n\} \) according to a probability distribution \( p_{k}(k) \) given by

\[
p_{k}(k) = \sum_{A \in T_{k}} p_{A}(A).
\]

Clearly, if \( A \) is u.g.r., then

\[
p_{A}(A) = \frac{p_{k}(k)}{|T_{q}^{n \times m,k}|},
\]

where \( k = \text{rank } A \). In this way, the rank probability distribution \( p_{k}(k) \) completely characterizes \( p_{A}(A) \) for \( A \) u.g.r.

The next sections are devoted to the problem of finding the capacity and the mutual information of the multiplicative finite-field matrix channel just described. We are also particularly interested in the mutual information when the input is restricted to constant-rank matrices.

We remark that for non u.g.r. transfer matrices, our results lend themselves as lower bounds. Indeed, it can be shown that, regardless of the distribution for \( G \), the matrix \( G' = T_{2}GT_{1} \) is always u.g.r. (here \( T_{1} \) and \( T_{2} \) are full-rank uniformly distributed \( n \times n \) random matrices, independent of \( G \) and of each other). This adjustment can be accomplished by premultiplying the transmitted matrix by \( T_{1} \), and the received matrix by \( T_{2} \).

IV. CHANNEL TRANSITION PROBABILITY

From now on we define the random variables

\[
u \triangleq \text{rank } X, \quad v \triangleq \text{rank } Y, \quad r \triangleq \text{rank } G,
\]

distributed according to \( p_{u}(u), p_{v}(v), \) and \( p_{r}(r) \), respectively.

In this section we derive the channel transition probability induced by the channel law and by the probability distribution

\[
249
\]
of $G$ (or, since $G$ is u.g.r., by the probability distribution of its rank $r$). We start with another basic enumeration result which is closely related to the multiplicative matrix channel.

**Lemma 3:** Let $X \in T_q^{n \times m,u}$ and $Y \in T_q^{n \times m,v}$. The number of matrices $G \in T_q^{n \times n,r}$ such that $GX = Y$ is

$$|\{G \in T_r : GX = Y\}| = \begin{cases} \phi_q(n; u, n, v, r), & \text{if } \langle Y \rangle \subseteq \langle X \rangle; \\ 0, & \text{else.} \end{cases}$$

Now, let $X \in T_q^{n \times m,u}$ and $v$ be an integer. The number of matrices $G \in T_q^{n \times n,r}$ such that rank $GX = v$ is given by

$$|\{G \in T_r : \text{rank } GX = v\}| = |T_q^{n \times u,v}| \phi_q(n; u, n, v, r).$$

The “rank transition probability” plays an important role in this work.

**Lemma 4:** The probability of $v$ conditioned on $u$, called the rank transition probability, is given by

$$p_{v|u}(v|u) = \sum_{r=0}^{n} p_r(v) \frac{|T_q^{n \times u,v}|}{|T_q^{n \times n,r}|} \phi_q(n; u, n, v, r). \quad (2)$$

**Remark:** We readily identify in (2) the rank transition probability conditioned on $r$,

$$p_{v|u,r}(v|u) = \frac{|T_q^{n \times u,v}|}{|T_q^{n \times n,r}|} \phi_q(n; u, n, v, r).$$

It is interesting to note that $p_{v|u,r}(v|u) \neq 0$ if and only if $\max\{u + r - n, 0\} \leq v \leq \min\{u, r\}$, as expected (the lower bound on $v$ agrees with Sylvester’s rank inequality). Of course, we also have $p_{v|u}(v|u) = 0$ for $v > u$.

We are now in a position to derive the channel transition probability.

**Lemma 5:** Let $X$ and $Y$ be related by $Y = GX$, where $G \in T_q^{n \times n}$ is u.g.r. and independent of $X$. Then, for $X \in T_q^{n \times m,u}$ and $Y \in T_q^{n \times m,v}$, we have

$$p_{Y|X}(Y|X) = \begin{cases} \frac{p_{v|u}(v|u)}{|T_q^{n \times u,v}|}, & \text{if } \langle Y \rangle \subseteq \langle X \rangle; \\ 0, & \text{else,} \end{cases}$$

where $p_{v|u}(v|u)$ is the rank transition probability as given by Lemma 4.

V. CHANNEL MUTUAL INFORMATION AND CAPACITY

This section derives the channel mutual information for u.g.r. input. We will see that this kind of input suffices to achieve the capacity, so that there is no need to consider more general inputs. We start by calculating the conditional entropy of the output.

**Lemma 6:** The entropy of the output $Y$ conditioned on the input $X$ is given by

$$H(Y|X) = \sum_{u=0}^{n} b_u p_u(u),$$

where

$$b_u = \sum_{v=0}^{n} p_{v|u}(v|u) \log_q \frac{|T_q^{n \times u,v}|}{p_{v|u}(v|u)}. \quad (3)$$

Note that, no matter how $X$ is distributed, $H(Y|X)$ depends on $X$ only through $u$.

**Lemma 7:** If the input $X$ is u.g.r., then

$$p_Y(Y) = \frac{p_v(v)}{|T_q^{n \times m,v}|},$$

that is, the output $Y$ is also u.g.r.

Next, we have a result regarding the unconditional entropy of the output.

**Lemma 8:** The entropy of $Y$ is upper-bounded by

$$H(Y) \leq \sum_{v=0}^{n} p_v(v) \log_q \frac{|T_q^{n \times m,v}|}{p_v(v)},$$

with equality when $X$ is u.g.r.

Now, let

$$I^*(p_u) = \max_{p_X \in P_u} I(X; Y),$$

where the maximum is over the collection of all matrix probability distributions $p_X$ with associated rank probability distribution equal to $p_u$, that is, over the set

$$\{p_X: \sum_{X \in T_u} p_X(X) = p_u(u), u = 0, \ldots, n\}.$$

We shall now state the main result of this work.

**Theorem 9:** Consider a DMC with input $X \in \mathbb{F}_q^{n \times m}$, output $Y \in \mathbb{F}_q^{n \times m}$, and channel law given by $Y = GX$, where $G$ is u.g.r. and independent of $X$. Then, $I^*(p_u)$ as defined in (4) is achieved by u.g.r. input, and is given by

$$I^*(p_u) = \sum_{v=0}^{n} p_v(v) \log_q \frac{|T_q^{n \times m,v}|}{p_v(v)} - \sum_{u=0}^{n} b_u p_u(u), \quad (5)$$

where

$$p_v(v) = \sum_{u=0}^{n} p_u(u) p_{v|u}(v|u), \quad (6)$$

the parameter $b_u$ is defined by (3), and $p_{v|u}(v|u)$ is the rank transition probability as given by (2).

Since any $p_X$ must be associated with some $p_u$, we have the following corollary.

**Corollary 10:** The capacity of the channel is

$$C = \max_{p_u} I^*(p_u).$$

Thus, the problem of finding the capacity and the corresponding optimal input for the multiplicative matrix channel, which was originally a convex optimization problem over $q^{nm}$ variables (namely, $p_X(X)$ for $X \in \mathbb{F}_q^{n \times m}$), is simplified to another convex optimization problem, this time involving only $n + 1$ variables (namely, $p_u(u)$, for $u = 0, \ldots, n$). This can be accomplished by standards methods.
VI. AN UPPER BOUND ON THE CAPACITY

The aforementioned convex optimization problem has two constraints, namely, \( \sum_{u=0}^{n} p_u(u) = 1 \) and \( p_u(u) \geq 0 \). If we drop the latter constraint on non-negativity, the problem allows for a closed-form analytic solution, which stands as an upper bound on the capacity.

The idea is to solve this new optimization problem in terms of \( p_v(v) \) instead of \( p_u(u) \). We assume \( p_v(n) > 0 \), for this assures that \( p_v[u](v|u) > 0 \) for \( v = u \) (a condition which will prove to be useful later).

With this in mind, we rewrite the mutual information (5) as

\[
I^*(p_v) = \sum_{v=0}^{n} p_v(v) \log_q \frac{|T_{q \times m, v}|}{p_v(v)} - \sum_{v=0}^{n} c_v p_v(v),
\]

where each \( c_v \) may be obtained from \( b_u \) and \( p_v[u](v|u) \) by solving the triangular (recall \( p_v[u](v|u) = 0 \) for \( v > u \)) linear system of equations

\[
\sum_{v=0}^{n} p_v[u](v|u)c_v = b_u,
\]

for \( u = 0, \ldots, n \). Furthermore, the constraint \( \sum_{u=0}^{n} p_u(u) = 1 \) is mapped to \( \sum_{v=0}^{n} p_v(v) = 1 \). By using the method of Lagrange multipliers, one may show that optimal values for \( p_v(v) \) are

\[
\tilde{p}_v(v) = \frac{|T_{q \times m, v}|^{q-c_v}}{\sum_{j=0}^{n} |T_{q \times m, j}|^{q-c_j}},
\]

and the maximum for \( I^*(p_v) \) is

\[
\tilde{C} = \log_q \sum_{v=0}^{n} |T_{q \times m, v}|^{q-c_v}.
\]

We then come back to \( p_u(u) \) by using (6), that is by solving

\[
\sum_{v=0}^{n} p_v[u](v|u)\tilde{p}_u(u) = \tilde{p}_v(v),
\]

for \( v = 0, \ldots, n \), which is again a triangular linear system of equations. Both systems (7) and (10) are guaranteed to have unique solutions, because \( p_v[u](v|u) > 0 \) for \( v = u \).

Note that \( \tilde{p}_v(v) \) as given by (8) is always non-negative. The same may not be true about \( \tilde{p}_u(u) \) obtained by solving (10); nevertheless, if \( \tilde{p}_u(u) \) happens to be non-negative, then it is indeed the solution of the original optimization problem (that is, the one with the non-negativity constraint), and the corresponding maximum mutual information (9) is the true channel capacity.

We thus have the following.

**Theorem 11:** Consider \( p_v(n) > 0 \). The capacity \( C \) of the multiplicative channel is bounded by

\[
C \leq \log_q \sum_{v=0}^{n} |T_{q \times m, v}|^{q-c_v},
\]

in which \( c_v \) may be computed by solving (7). Furthermore, let \( \tilde{p}_u(u) \) be obtained by solving (10), where \( \tilde{p}_v(v) \) is given by (8). If \( \tilde{p}_u(u) \geq 0 \) for \( u = 0, \ldots, n \), then equality holds in (11) and \( \tilde{p}_u(u) \) defines a u.g.r. capacity-achieving input.

VII. PARTICULAR CASES

This section is devoted to some important particular cases.

A. Constant-Rank Input

If the input is restricted to rank-\( u \) matrices, then \( u = u \) is a constant, and therefore \( p_v(v) = p_v[u](v|u) \). The channel mutual information given by Theorem 9 simplifies to

\[
I^*(p_u) = \sum_{v=0}^{n} p_v[u](v|u) \log_q \frac{|T_{q \times m, v}|}{|T_{q \times m, u}|}.
\]

After applying Fact 1, and bearing in mind that this constant-rank input defines a new matrix channel, we get the following.

**Corollary 12:** If the input is restricted to rank-\( u \) matrices, then the capacity of the (new) multiplicative channel is achieved by the uniform (over \( T_{q \times m, u} \)) input distribution, and is given by

\[
C_u = \sum_{v=0}^{n} p_v[u](v|u) \log_q \frac{[m]_q}{[u]_q},
\]

with \( p_v[u](v|u) \) given by (2).

Note that this naturally yields a lower bound for the general input problem, that is, \( C \geq \max_u C_u \).

Also, Corollary 12 agrees with our previous result in [8], where \( C_u \) was obtained.

B. Full-Rank Transfer Matrix

We now recover a result from Silva et al. [2], which follows from Theorem 11.

**Corollary 13:** If \( G \) is uniformly distributed over \( T_q^{n \times m} \), then

\[
C = \log_q \sum_{i=0}^{n} \frac{[m]_i}{[i]_q},
\]

achieved with

\[
p_u(u) = \frac{[m]_u}{\sum_{i=0}^{n} [m]_i}. \]

C. Uniform Transfer Matrix

This case is studied in details by Jafari et al. [4]. First note that \( G \) uniformly distributed over \( F_q^{m \times n} \) is a particular case of \( G \) u.g.r. with

\[
p_v(r) = \frac{|T_{q \times n, r}|}{q^n}.
\]

This allows for simplifications on the rank transition probability \( p_v[u](v|u) \), namely,

\[
p_v[u](v|u) = \begin{cases} |T_{q \times n, u}| & \text{if } u \leq v; \\ 0 & \text{else,} \end{cases}
\]

as well as in the coefficients \( b_u \), namely, \( b_u = un \). (Note that \( \sum_{r=0}^{n} \phi_q(n; u, n, v, r) = q^{(n-u)n} \) one may then apply Theorems 9 and 11.
VIII. COMMUNICATION VIA SUBSPACES

In this section we show that the multiplicative matrix channel can be converted into an equivalent subspace channel, without loss of information. We start by presenting a general approach to the problem. Let $\mathcal{X}, \mathcal{Y}$ be a DMC. Let $f : \mathcal{X} \to \mathcal{U}$ and $g : \mathcal{Y} \to \mathcal{V}$ be surjective functions. We will interpret $f$ as a grouping of input letters and likewise $g$ as a grouping of output letters. We are interested in certain conditions under which these groupings are information-lossless.

**Lemma 14:** Suppose that $p_{Y|X}(Y|X') = p_{Y|X}(Y'|X')$ for all $Y \in \mathcal{Y}$ and all $X, X' \in \mathcal{X}$ such that $f(X) = f(X')$. Then

$$I(X; Y) = I(f(X); Y)$$

for all $p_X$. Moreover, for all $U \in \mathcal{U}$ and all $Y \in \mathcal{Y}$, we have

$$p_{Y|U}(Y|U) = p_{Y|X}(Y|X)$$

for any $X$ such that $f(X) = U$.

**Lemma 15:** Suppose that $p_{Y|X}(Y|X) = p_{Y|X}(Y'|X)$ for all $X \in \mathcal{X}$ and all $Y, Y' \in \mathcal{Y}$ such that $g(Y) = g(Y')$. Then

$$I(X; Y) = I(f(X); g(Y))$$

for all $p_X$. Moreover, for all $X \in \mathcal{X}$ and all $Y \in \mathcal{Y}$, we have

$$p_{V|X}(V|X) = |\{Y' : g(Y') = V\}| \cdot p_{Y|X}(Y|X)$$

for any $Y$ such that $g(Y) = V$.

We now apply these results to the matrix channel considered in this paper. Let $\mathcal{P}(\mathbb{F}_q^n, n)$ denote the set of all subspaces of $\mathbb{F}_q^n$ with dimension $n$ or less.

**Theorem 16:** Let $\mathcal{X} = \mathcal{Y} = \mathbb{F}_q^{n \times m}$ and $\mathcal{U} = \mathcal{V} = \mathcal{P}(\mathbb{F}_q^n, n)$. Consider the matrix channel $(\mathcal{X}, p_{Y|X}(Y|X))$ with input $X$, output $Y$, and transition probabilities $p_{Y|X}(Y|X)$ induced by the channel law $Y = GX$, where $G$ is u.g.r. and independent of $X$. Consider also the subspace channel $(\mathcal{U}, p_{V|U}(V|U))$ obtained by grouping input (and output) matrices that span the same subspace, and let $U \triangleq \langle X \rangle$ and $V \triangleq \langle Y \rangle$. Then,

$$p_{V|U}(V|U) = |T_q^{n \times \dim V}| \cdot p_{Y|X}(Y|X),$$

where $X$ and $Y$ are any matrices such that $\langle X \rangle = U$ and $\langle Y \rangle = V$. Furthermore,

$$I(X; Y) = I(U; V)$$

for every input distribution $p_X(X)$.

Theorem 16 shows that, with respect to the multiplicative matrix channel with u.g.r. transfer matrix, the groupings $U = f(X) = \langle X \rangle$ and $V = g(Y) = \langle Y \rangle$ are information-lossless. Thus, the matrix channel $(\mathcal{X}, p_{Y|X}(Y|X))$ can be transformed into a (simpler) subspace channel $(\mathcal{U}, p_{V|U}(V|U), V)$ with transition probabilities given by (12). Concretely, the new channel is obtained by concatenating the original channel at the input with a device that takes a subspace $U$ to any matrix $X$ such that $\langle X \rangle = U$, and at the output with a device that computes $V = \langle Y \rangle$. Due to (13), any coding scheme for the matrix channel has a counterpart in the subspace channel achieving exactly the same mutual information, and vice-versa. In particular, one may focus solely on $(\mathcal{U}, p_{V|U}(V|U), V)$ when designing and analyzing capacity-achieving schemes.

IX. CONCLUSIONS

This paper has focused on computing the capacity of multiplicative finite-field matrix channels under the assumption of a u.g.r. transfer matrix. Our main contribution is to show that, for a given input rank distribution, the mutual information is maximized by a u.g.r. input. As a consequence, the problem of finding the capacity reduces to a convex optimization problem on $n + 1$ variables (rather than $q^{nm}$), allowing for easy numerical computation by standard techniques. An upper bound on capacity is then obtained by relaxing this optimization problem.

Applications of our results include the special cases of constant-rank input, full-rank transfer matrix, and uniform transfer matrix, from which we easily recover some previous results. In particular, we prove that, even in our more general setup, subspace coding is still sufficient to achieve capacity. To obtain this result, we have introduced the concept of grouping input or output letters in a general DMC, which may have applications in other scenarios.

Several questions arise from the results of this paper. Work is in progress to approximate the channel capacity in the limit of large field size or packet length, as well as to obtain numerical results for interesting special cases. In addition, the design of low-complexity capacity-achieving schemes for this channel (recently addressed by Yang et al. in [5]) is an important and still largely open problem.

REFERENCES