Communication over Finite-Ring Matrix Channels

Chen Feng*, Roberto W. Nóbrega†, Frank R. Kschischang*, Danilo Silva†
*Department of Electrical and Computer Engineering, University of Toronto, Canada
†Department of Electrical Engineering, Federal University of Santa Catarina, Brazil
cfeng@eecg.utoronto.ca, rwnobrega@eel.ufsc.br, frank@comm.utoronto.ca, danilo@eel.ufsc.br

***Abstract*—Though network coding is traditionally performed over finite fields, recent work on nested-lattice-based network coding suggests that, by allowing network coding over finite rings, more efficient physical-layer network coding schemes can be constructed. This paper considers the problem of communication over a finite-chain-ring matrix channel $Y = AX + BZ$, where X is the channel input, Y is the channel output, Z is random noise, and A and B are random transfer matrices. Tight capacity results are obtained and simple polynomial-complexity capacity-achieving coding schemes are provided under certain distributions of A, B, and Z, extending the work of Silva, Kschischang and Kötter (2010), who handled the case of finite fields. This extension is based on several new results that generalize concepts and methods from matrices over finite fields to matrices over finite chain rings.**

I. Introduction

Motivated by nested-lattice-based physical-layer network coding [1]–[3], this paper extends the work of [4] (which considered finite-field matrix channels) to the case of finite chain rings. As shown in [3], if the ambient message space is allowed to be over some finite ring, then *any* pair of nested complex lattices can be used to design *linear* physical-layer network coding (PNC) (also called compute-and-forward). This not only leads to a much richer design space for linear PNC, but also enables the use of powerful lattice construction methods, such as Construction D and complex Construction D (see, e.g., [5]). For these constructions, the message space Ω takes the form

$$\Omega = T/\langle p^{t_1}\rangle \times \cdots \times T/\langle p^{t_m}\rangle, \tag{1}$$

where T is some principal ideal domain, p is a prime in T, and $t_1, \ldots, t_m$ are positive integers satisfying $t_1 \leq \ldots \leq t_m$. Algebraically, the above message space Ω is isomorphic to a finite module over the finite chain ring $T/\langle p^{t_m}\rangle$. Operationally, the above message space Ω implies that each packet in the system can be viewed as an m-tuple over $T/\langle p^{t_m}\rangle$.

This message space Ω naturally induces a matrix channel over the finite chain ring $T/\langle p^{t_m}\rangle$. Transmitted and received packets, drawn from the message space Ω, can be gathered into the rows of a transmitted matrix X and a received matrix Y. Error packets injected into the network can be described by the rows of a noise matrix Z. Due to the nature of linear PNC, the linear transformation of transmitted packets X and the linear propagation of error packets Z can be modelled by writing

$$Y = AX + BZ \tag{2}$$

for some transfer matrices A, B. One typically assumes that A, B, and Z are random matrices (drawn according to some distributions) and independent of X. This type of stochastic model is appropriate in situations where the noise Z arises due to decoding errors, rather than from the malicious actions of an adversary.

In finite-field matrix channels, the noise matrix Z is typically assumed to have rank bounded by some parameter t, which may be regarded as a measure of "noise level" in the message space. For matrices over rings, the concept of "rank" is more subtle, and must be suitably generalized. In this paper we study the situation when Z spans a module with some bounded "shape", the appropriate chain-ring-theoretic generalization of dimension. We generalize the results of [4], deriving tight capacity bounds and simple, polynomial-complexity, asymptotically capacity-achieving coding schemes for matrix channel models related to (2). As in [4], we gather insight by first studying two variations: the noise-free multiplicative matrix channel $Y = AX$, and the multiplication-free additive matrix channel $Y = X + BZ$. Capacity-approaching schemes for the general case are obtained by combining the transmission strategies from these two special cases.

The results we obtain parallel those of [4]; however they are based on several new results that generalize concepts and methods from matrices over finite fields to matrices over finite chain rings. These results may be of independent interest.

The remainder of this paper is organized as follows. In Section II we review some basic facts about rings, modules and matrices. Section III presents our new results on matrices over finite chain rings that serve as building blocks to the subsequent sections. Section IV formalizes the basic channel that is studied in this paper. The three variations of this channel are addressed in Sections V, VI and VII, where capacity and coding results are presented. Finally, Section VIII presents our conclusions. Space constraints have forced us to omit proofs in most cases; however, see [6].

II. Preliminaries

In this section, we present some basic results for finite chain rings and modules and matrices over finite chain rings. More details can be found in, *e.g.*, [7], [8].

The work of R. W. Nóbrega and D. Silva was partly supported by CNPq-Brazil.

A. Finite Chain Rings

All rings in this paper will be commutative with identity $1 \neq 0$. A finite ring R is called a *finite chain ring* if the ideals of R satisfy the chain condition: for any two ideals I, J of R, either $I \subseteq J$ or $J \subseteq I$. Clearly, a finite chain ring has a unique maximal ideal. Moreover, it is also known that such a ring is a principal ideal ring, i.e., all its ideals have a single generator. Let $\pi \in R$ be any generator of the maximal ideal of R, and let s be the *nilpotency index* of π, *i.e.*, the smallest positive integer such that $\pi^s = 0$. Then it is known that R has exactly $s + 1$ ideals, namely,

$$R = \langle \pi^0 \rangle \supset \langle \pi^1 \rangle \supset \cdots \supset \langle \pi^{s-1} \rangle \supset \langle \pi^s \rangle = \{0\},$$

where $\langle x \rangle$ denotes the ideal generated by x. It is also known that the quotient $R/\langle \pi \rangle$ is a field, called the residue field of R. We say R is a (q, s) chain ring, if q is the size of the residue field and s is the nilpotency index. For example, $\mathbb{Z}_8$ is a $(2, 3)$ chain ring with the maximal ideal $\langle 2 \rangle = \{0, 2, 4, 6\}$.

Let $\mathcal{R}(R, \pi) \subseteq R$ be a complete set of residues with respect to π and, without loss of generality, assume that $0 \in \mathcal{R}(R, \pi)$. Every element $a \in R$ then has a unique representation in the form

$$a = a_0 + a_1\pi + \cdots + a_{s-1}\pi^{s-1} \tag{3}$$

where $a_0, \ldots, a_{s-1} \in \mathcal{R}(R, \pi)$. We call (3) a *$\pi$-adic decomposition* of the element a. The *degree* of a nonzero element a is defined as the least index j for which $a_j \neq 0$.

B. Modules over Finite Chain Rings

Let R be a (q, s) chain ring. An *s-shape* $\mu = (\mu_1, \ldots, \mu_s)$ is a sequence of non-decreasing non-negative integers, *i.e.*, $0 \leq \mu_1 \leq \cdots \leq \mu_s$. We denote by $|\mu|$ the sum of its components, *i.e.*, $|\mu| = \sum_{i=1}^{s} \mu_i$. An s-shape κ is said to be a *subshape* of μ, written $\kappa \preceq \mu$, if $\kappa_i \leq \mu_i$ for all $i = 1, \ldots, s$. Thus, for example, $(1, 1, 3) \preceq (2, 4, 4)$. For convenience, we will sometimes identify an integer t with the s-shape $(t, \ldots, t)$. Thus, for example, $\mu \preceq t$ means $\mu_i \leq t$ for all i, $\tau = t$ means $\tau_i = t$ for all i, and $\mu - t = (\mu_1 - t, \ldots, \mu_s - t)$.

For any s-shape μ, we define an R-module R^μ as

$$\underbrace{R \times \cdots \times R}_{\mu_1} \times \underbrace{\pi R \times \cdots \times \pi R}_{\mu_2 - \mu_1} \times \cdots \times \underbrace{\pi^{s-1} R \times \cdots \times \pi^{s-1} R}_{\mu_s - \mu_{s-1}}.$$

The module R^μ can be viewed as a collection of μ_s-tuples whose components are subject to some constraints imposed by μ: the first μ_1 components can be any elements of R, the next $\mu_2 - \mu_1$ components must be multiples of π, and so on. It is easy to check that the size of R^μ is $|R^\mu| = q^{|\mu|}$.

For any finite R-module M, there is a unique s-shape μ such that $M \cong R^\mu$. We call the unique shape μ the shape of M, and write $\mu = \operatorname{shape}(M)$. If M' is a submodule of M, then $\operatorname{shape}(M') \preceq \operatorname{shape}(M)$. It is known [7], [8] that the number of submodules of R^μ whose shape is κ is given by

$$\begin{bmatrix}\!\begin{bmatrix} \mu \\ \kappa \end{bmatrix}\!\end{bmatrix}_q = \prod_{i=1}^{s} q^{(\mu_i - \kappa_i)\kappa_{i-1}} \begin{bmatrix} \mu_i - \kappa_{i-1} \\ \kappa_i - \kappa_{i-1} \end{bmatrix}_q,$$

where $\begin{bmatrix} m \\ k \end{bmatrix}_q \triangleq \prod_{i=0}^{k-1} (q^m - q^i)/(q^k - q^i)$ is the Gaussian coefficient.

C. Matrices over Finite Chain Rings

The set of all $n \times m$ matrices with entries from R will be denoted by $R^{n \times m}$. If $A \in R^{n \times m}$, we will let $A[i, j]$ denote the entry of A in the ith row and jth column; we will let $A[i_1 \colon i_2, j_1 \colon j_2]$ denote a submatrix of A formed by rows i_1 to i_2 and by columns j_1 to j_2. We say $A[i, j]$ is *above* $A[i', j']$ if $i < i'$ and $A[i, j]$ is *earlier than* $A[i', j']$ if $j < j'$.

A square matrix $U \in R^{n \times n}$ is *invertible* if $UV = VU = I_n$ for some $V \in R^{n \times n}$, where I_n denotes the $n \times n$ identity matrix. Two matrices $A, B \in R^{n \times m}$ are *row-equivalent* if there exists an invertible matrix U such that $UA = B$.

For any $A \in R^{n \times m}$, the *row module* of A, denoted by $\operatorname{row}(A)$, is the set of all R-linear combinations of the rows of A. Clearly, row-equivalent matrices have the same row module. The *shape* of a matrix A is defined as the shape of the row module of A, *i.e.*, $\operatorname{shape}(A) = \operatorname{shape}(\operatorname{row}(A))$.

D. Row Canonical Form

The row canonical form presented here is essentially the same as the reduced row echelon form defined in Kiermaier's thesis [9, Definition 2.2.2], which is a variant of the row canonical form in [7, p. 329, Exercise XVI.7]. It appears that the key idea behind these forms was proposed by Fuller [10] based on an earlier result of Birkhoff [11].

To introduce the row canonical form, we first need a few definitions. The *pivot* of a nonzero row of a matrix $A \in R^{n \times m}$ is the first entry among the entries having least degree in that row. For example, 6 is the pivot of the row $[0\ 4\ 6\ 2]$. For $0 < \ell < s$, let

$$\mathcal{R}(R, \pi^\ell) = \left\{ \sum_{i=0}^{\ell-1} a_i \pi^i : a_0, \ldots, a_{\ell-1} \in \mathcal{R}(R, \pi) \right\}$$

and let $\mathcal{R}(R, \pi^0) = \{0\}$.

A matrix A is in *row canonical form* if it satisfies the following four conditions:

1) All nonzero rows are above any zero rows.
2) If A has two pivots of the same degree, the one that occurs earlier is above the one that occurs later. If A has two pivots of different degree, the one with smaller degree is above the one with larger degree.
3) Every pivot is of the form π^ℓ for some $\ell \in \{0, \ldots, s-1\}$.
4) For every pivot (say π^ℓ), all entries below and in the same column as the pivot are zero, and all entries above and in the same column as the pivot are elements of $\mathcal{R}(R, \pi^\ell)$.

For instance, the following matrix over $\mathbb{Z}_8$,

$$A = \begin{bmatrix} 0 & 2 & \bar{1} \\ 0 & \bar{4} & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

in which the pivots have been identified with an overline, is in row canonical form.

For any $A \in R^{n \times m}$, we say a matrix $B \in R^{n \times m}$ is a *row canonical form of* A, if (i) B is in row canonical form, and (ii) B is row-equivalent to A. We can show that such

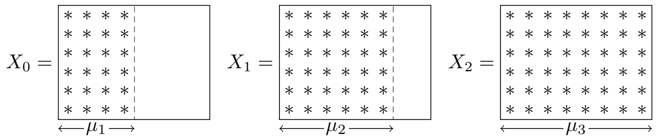

Fig. 1. Illustration of a π-adic decomposition for $s = 3$ and $\mu = (4, 6, 8)$.

a row canonical form is unique, and can be computed with $\mathcal{O}(nm \min\{n, m\})$ basic operations over R.

III. Matrices under Row Constraints

In this section, we study a class of matrices in $R^{n\times m}$ whose rows are constrained to be elements of R^{μ}, which is of primary importance to our study of capacities and coding schemes.

Let $R^{n\times\mu}$ denote the set of matrices in $R^{n\times m}$ whose rows are elements of R^{μ}. Then the size of $R^{n\times\mu}$ is $|R^{n\times\mu}| = q^{n|\mu|}$, since each row has $|R^{\mu}| = q^{|\mu|}$ choices.

Every matrix $X \in R^{n\times\mu}$ can be constructed based on its π-adic decomposition

$$X = X_0 + \pi X_1 + \cdots + \pi^{s-1} X_{s-1},$$

with each auxiliary matrix X_i ($i = 0, \ldots, s-1$) satisfying:

- $X_i[1\colon n, 1\colon \mu_{i+1}]$ can be any matrix over $\mathcal{R}(R, \pi)$.
- All other entries in X_i are zero.

This construction (illustrated in Fig. 1) will be used in Sec. VI, and can be viewed as a (one-to-one) mapping from sequences of $n|\mu|$ q-ary symbols to matrices in $R^{n\times\mu}$.

Let $\mathcal{T}_\kappa(R^{n\times\mu})$ denote the set of matrices in $R^{n\times\mu}$ whose shape is κ. A central result in this section is the following:

Theorem 1: When $\kappa \preceq n, \mu$, the size of $\mathcal{T}_\kappa(R^{n\times\mu})$ is

$$|\mathcal{T}_\kappa(R^{n\times\mu})| = \begin{bmatrix}\!\begin{bmatrix}\mu \\ \kappa\end{bmatrix}\!\end{bmatrix}_q |R^{n\times\kappa}| \prod_{i=0}^{\kappa_s - 1} (1 - q^{i-n}). \tag{4}$$

This result allows us to derive tight capacity bounds.

A row canonical form in $\mathcal{T}_\kappa(R^{n\times\mu})$ is called *principal* if its diagonal entries $d_1, d_2, \ldots, d_k$ ($k = \min\{n, m\}$) have the following form

$$d_1, \ldots, d_k = \underbrace{1, \ldots, 1}_{\kappa_1}, \underbrace{\pi, \ldots, \pi}_{\kappa_2 - \kappa_1}, \ldots, \underbrace{\pi^{s-1}, \ldots, \pi^{s-1}}_{\kappa_s - \kappa_{s-1}}, \underbrace{0, \ldots, 0}_{k - \kappa_s}.$$

Every principal row canonical form $X \in \mathcal{T}_\kappa(R^{n\times\mu})$ can be constructed based on its π-adic decomposition

$$X = X_0 + \pi X_1 + \cdots + \pi^{s-1} X_{s-1},$$

with each auxiliary matrix X_i following the rules below:

- $X_i[1\colon \kappa_{i+1}, 1\colon \kappa_{i+1}] = \operatorname{diag}(\underbrace{0, \ldots, 0}_{\kappa_i}, \underbrace{1, \ldots, 1}_{\kappa_{i+1} - \kappa_i})$.
- $X_i[1\colon \kappa_{i+1}, \kappa_{i+1} + 1\colon \mu_{i+1}]$ can be any matrix over $\mathcal{R}(R, \pi)$.
- All other entries in X_i are zero.

This construction (as illustrated in Fig. 2) is crucial to our coding schemes, and can be viewed as a (one-to-one) mapping from sequences of $\sum_{i=1}^{s} \kappa_i(\mu_i - \kappa_i)$ q-ary symbols to principal row canonical forms in $\mathcal{T}_\kappa(R^{n\times\mu})$.

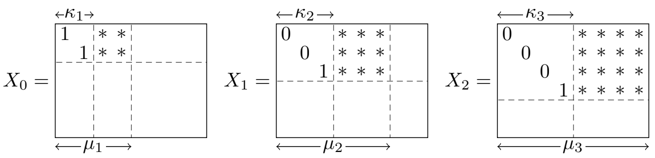

Fig. 2. Illustration of the construction of principal row canonical forms for $\mathcal{T}_\kappa(R^{n\times\mu})$ with $s = 3$, $n = 6$, $\mu = (4, 6, 8)$, and $\kappa = (2, 3, 4)$.

IV. Matrix Channels over Finite Chain Rings

Let R be a (q, s) chain ring, and let μ be an s-shape. Consider a matrix channel given by

$$Y = AX + BZ \tag{5}$$

where $X \in R^{n\times\mu}$ and $Y \in R^{N\times\mu}$ are the input and output matrices, respectively. The error matrix $Z \in R^{t\times\mu}$ and the transfer matrices $A \in R^{N\times n}$ and $B \in R^{N\times t}$ are random matrices with some joint distribution conditioned on X. Clearly, (5) is an instance of the discrete memoryless channel $(\mathcal{X}, p_{Y|X}, \mathcal{Y})$ with input alphabet $\mathcal{X} = R^{n\times\mu}$, output alphabet $\mathcal{Y} = R^{N\times\mu}$ and channel transition probability $p_{Y|X}$. The capacity of this channel is given by $C = \max_{p_X} I(X;Y)$, where p_X is the input distribution. We assume that logarithms are taken to the base q, so the capacity is given in q-ary units per channel use.

Note that we may think of the rows of X, Y and Z as packets over an ambient space R^{μ}. To support this ambient space, the length of a packet, denoted by m, is equal to μ_s. In many situations, it is useful to understand the capacity scaling as the packet length grows. For that reason, we introduce a notion of asymptotic capacity

$$\bar{C} = \lim_{m\to\infty} \frac{1}{n|\mu|} C = \lim_{m\to\infty} \frac{1}{\bar{n}|\bar{\mu}|m^2} C,$$

where we assume that $\bar{n} = n/m$ and $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_s) = \mu/m$ are fixed. Note that $\bar{C}$ is normalized such that $\bar{C} = 1$ if the channel is noiseless.

Let τ be an s-shape such that $\tau \preceq t, \mu$. Similarly to [4], we are primarily interested in the case where the transfer matrix A is invertible ($A \in \mathrm{GL}_n(R)$, in particular, $N = n$), B is uniform over $\mathcal{T}_t(R^{n\times t})$, Z is uniform over $\mathcal{T}_\tau(R^{t\times\mu})$, and X, A, B and Z are statistically independent. (The model in [4] is recovered when $R = \mathbb{F}_q$, $\mu = m$, and $\tau = t$.) In this case, we can rewrite (5) as

$$Y = A(X + A^{-1}BZ) = A(X + B'Z). \tag{6}$$

Note that if B is uniform over $\mathcal{T}_t(R^{n\times t})$ and independent from A, so is B'. Since both B' and Z are uniform and independent from A, it follows that $W = B'Z$ is uniform over $\mathcal{T}_\tau(R^{n\times\mu})$ and independent from A. Hence, the channel law (6) is equivalent to

$$Y = A(X + W) \tag{7}$$

where $A \in \mathrm{GL}_n(R)$ and $W \in \mathcal{T}_\tau(R^{n\times\mu})$ are chosen uniformly at random and independently from any other variables. We will focus on the channel model (7) in the rest of this paper.

V. The Multiplicative Matrix Channel

As a first special case, following [4], we consider the *multiplicative matrix channel (MMC)* defined by the law $Y = AX$, where A is uniform over $\mathrm{GL}_n(R)$ and independent from X. This model is a special case of the channel model (7) with $W = 0$.

A. Capacity

Theorem 2: The capacity of the MMC, in q-ary units per channel use, is given by

$$C_{\text{MMC}} = \log_q \sum_{\lambda \preceq n, \mu} \left[\!\!\left[\begin{matrix} \mu \\ \lambda \end{matrix} \right]\!\!\right]_q ,$$

and is bounded by

$$\sum_{i=1}^{s} \kappa_i(\mu_i - \kappa_i) \leq C_{\text{MMC}} \leq \sum_{i=1}^{s} \kappa_i(\mu_i - \kappa_i) + \log_q 4^s \tbinom{n+s}{s} \quad (8)$$

where $\kappa_i = \min\{n, \lfloor \mu_i/2 \rfloor\}$ for all i.

Theorem 3: The asymptotic capacity $\bar{C}_{\text{MMC}}$ is given by

$$\bar{C}_{\text{MMC}} = \frac{\sum_{i=1}^{s} \bar{\kappa}_i(\bar{\mu}_i - \bar{\kappa}_i)}{\bar{n}|\bar{\mu}|}, \quad (9)$$

where $\bar{\kappa} = \kappa/m$ with $\kappa_i = \min\{n, \lfloor \mu_i/2 \rfloor\}$ for all i.

B. A Simple Coding Scheme

In this section, we present a simple coding scheme that achieves the asymptotic capacity in Theorem 3.

1) Encoding: The input matrix X is chosen from the set of principal row canonical forms for $\mathcal{T}_\kappa(R^{n \times \mu})$ by using the construction presented in Section III. Clearly, the encoding rate of the scheme is $R = \sum_{i=1}^{s} \kappa_i(\mu_i - \kappa_i)$.

2) Decoding: Upon receiving $Y = AX$, the decoder simply computes the row canonical form of Y. The decoding is always correct by the uniqueness of the row canonical form. By comparing the encoding rate with the asymptotic capacity, we have the following theorem.

Theorem 4: The coding scheme described above achieves the asymptotic capacity (9).

VI. The Additive Matrix Channel

In this section, we consider the *additive matrix channel (AMC)* defined by the law $Y = X + W$, where W is uniform over $\mathcal{T}_\tau(R^{n \times \mu})$ and independent from X. This model is a special case of the channel model (7) with $A = I$.

A. Capacity

Theorem 5: The capacity of the AMC, in q-ary units per channel use, is given by

$$C_{\text{AMC}} = \log_q |R^{n \times \mu}| - \log_q |\mathcal{T}_\tau(R^{n \times \mu})|,$$

and is bounded by

$$C_{\text{AMC}} > \sum_{i=1}^{s} (n - \tau_i)(\mu_i - \tau_i) - \log_q 4^s \prod_{i=0}^{\tau_s - 1} (1 - q^{i-n}),$$

$$C_{\text{AMC}} < \sum_{i=1}^{s} (n - \tau_i)(\mu_i - \tau_i) - \log_q \prod_{i=0}^{\tau_s - 1} (1 - q^{i-n}).$$

Theorem 6: The asymptotic capacity $\bar{C}_{\text{AMC}}$ is given by

$$\bar{C}_{\text{AMC}} = \frac{\sum_{i=1}^{s} (\bar{n} - \bar{\tau}_i)(\bar{\mu}_i - \bar{\tau}_i)}{\bar{n}|\bar{\mu}|}. \quad (10)$$

B. Coding Scheme

We focus on a special case when $\tau = t$, and present a coding scheme based on the idea of error-trapping in [4]. This scheme achieves the asymptotic capacity.

1) Encoding: Set $v \geq t$. The input matrix X is constructed as

$$X = \begin{bmatrix} 0 & 0 \\ 0 & U \end{bmatrix},$$

where the size of U is $(n - v) \times (m - v)$, and the sizes of other zero matrices are readily available. Here, U is chosen from the set $R^{(n-v) \times (\mu - v)}$ by using the construction in Section III. Clearly, the encoding rate is $R = \sum_{i=1}^{s} (n - v)(\mu_i - v)$.

2) Decoding: The decoder receives

$$\begin{aligned} Y &= X + W \\ &= \begin{bmatrix} 0 & 0 \\ 0 & U \end{bmatrix} + \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & U + W_{22} \end{bmatrix}, \end{aligned}$$

where the size of W_{11} is $v \times v$, and the sizes of other submatrices of W are readily obtained.

The decoder can observe W_{11}, W_{12} and W_{21} from Y thanks to the error traps. Thus, the goal of the decoder is to recover W_{22} from W_{11}, W_{12} and W_{21}. The following lemma says that if $\text{shape}(W_{11}) = t$, then W_{22} can be recovered correctly from W_{11}, W_{12} and W_{21}.

Lemma 1: If $\text{shape}(W_{11}) = t$, then there exists some matrix G such that $W_{21} = GW_{11}$. For any such matrix G, $W_{22} = GW_{12}$.

The decoding rule is summarized as follows. First, the decoder observes W_{11} and checks the condition $\text{shape}(W_{11}) = t$. If the condition does not hold, the decoder simply declares a failure. Otherwise, the decoder can always find a matrix G such that $W_{21} = GW_{11}$. Then, the decoder recovers W_{22} by using the formula $W_{22} = GW_{12}$, and thus obtain U.

Clearly, the error probability of the scheme is zero. The failure probability of the scheme is $P_f = \Pr[\text{shape}(W_{11}) \neq t]$.

Lemma 2: The failure probability P_f of the above scheme is upper-bounded by $P_f < \frac{2t}{q^{1+v-t}}$.

Recall that the encoding rate of the scheme is $R = \sum_{i=1}^{s} (n - v)(\mu_i - v)$. Thus, if we set v such that $v - t \to \infty$, and $\frac{v-t}{m} \to 0$, as $m \to \infty$, then we have $P_f \to 0$ and $\bar{R} = \frac{R}{n|\mu|} \to \bar{C}_{\text{AMC}}$. Therefore, we have the following theorem.

Theorem 7: When $\tau = t$, the coding scheme described above can achieve the capacity expression (10).

VII. The Additive-Multiplicative Matrix Channel

In this section, we consider the *additive-multiplicative matrix channel (AMMC)* defined by the law $Y = A(X + W)$, where $A \in \mathrm{GL}_n(R)$ and $W \in \mathcal{T}_\tau(R^{n \times \mu})$ are uniformly distributed and independent from any other variables.

A. Capacity Bounds

Theorem 8: The capacity of the AMMC, in q-ary units per channel use, is upper-bounded by

$$C_{\text{AMMC}} \leq \sum_{i=1}^{s}(\mu_i - \xi_i)\xi_i + \sum_{i=1}^{s}(n-\mu_i)\tau_i + 2s\log_q 4 \\ + \log_q \tbinom{n+s}{s} + \log_q \tbinom{\tau_s+s}{s} - \log_q \textstyle\prod_{i=0}^{\tau_s-1}(1-q^{i-n}),$$

where $\xi_i = \min\{n, \lfloor \mu_i/2 \rfloor\}$ for all i.

Theorem 9: When $\mu \succeq 2n$, the asymptotic capacity $\bar{C}_{\text{AMMC}}$ is upper-bounded by

$$\bar{C}_{\text{AMMC}} \leq \frac{\sum_{i=1}^{s}(\bar{n}-\bar{\tau}_i)(\bar{\mu}_i - \bar{n})}{\bar{n}|\bar{\mu}|}. \tag{11}$$

B. A Coding Scheme

We focus on a special case when $\tau = t$. We describe a coding scheme that achieves the asymptotic capacity in Theorem 9 when $\mu \succeq 2n$.

1) Encoding: The encoding is a direct combination of the encoding strategies for the MMC and the AMC. Specifically, with $v \geq t$, the input matrix X is constructed as

$$X = \begin{bmatrix} 0 & 0 \\ 0 & \bar{X} \end{bmatrix},$$

where the size of $\bar{X}$ is $(n-v)\times(m-v)$, and the sizes of other zero matrices are readily available. Here, $\bar{X}$ is chosen from the set of principal row canonical forms for $\mathcal{T}_\kappa(R^{(n-v)\times(\mu-v)})$ by using the construction in Section III, where $\kappa_i = \min\{n-v, \lfloor(\mu_i - v)/2\rfloor\}$ for all i. Clearly, the encoding rate of the scheme is $R = \sum_{i=1}^{s}\kappa_i(\mu_i - v - \kappa_i)$. In particular, when $\mu \succeq 2n$, we have $\lfloor(\mu_i - v)/2\rfloor \geq n - v$ for all i. Thus, $\kappa_i = n - v$ for all i, and the encoding rate is $R = \sum_{i=1}^{s}(n-v)(\mu_i - n)$.

2) Decoding: The decoder receives

$$\begin{aligned} Y &= A(X+W) \\ &= A\left(\begin{bmatrix} 0 & 0 \\ 0 & \bar{X} \end{bmatrix} + \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}\right) = A\begin{bmatrix} W_{11} & W_{12} \\ W_{21} & \bar{X} + W_{22} \end{bmatrix} \end{aligned}$$

where the size of W_{11} is $v \times v$, and the sizes of other submatrices of W are readily obtained.

If $\text{shape}(W_{11}) = t$, then $\bar{X}$ can be recovered correctly from the row canonical form of Y, as shown in the following lemma.

Lemma 3: Let $\tilde{Y}$ be the row canonical form of Y. If $\text{shape}(W_{11}) = t$, then $\bar{X} = \tilde{Y}[t+1\colon n+t-v, v+1\colon m]$.

The decoding rule can be summarized as follows. First, the decoder computes the row canonical form $\tilde{Y}$. Second, the decoder checks the condition $\text{shape}(W_{11}) = t$. (This can be achieved by looking at $\tilde{Y}$.) If the condition does not hold, the decoder declares a failure. Otherwise, the decoder outputs $\bar{X} = \tilde{Y}[t+1\colon n+t-v, v+1\colon m]$.

Clearly, the error probability of the scheme is zero; the failure probability of the scheme is $P_f = \Pr[\text{shape}(W_{11}) \neq t]$, which is upper-bounded by $P_f < \frac{2t}{q^{1+v-t}}$. Thus, if we set v such that $v - t \to \infty$ and $\frac{v-t}{m} \to 0$, as $m \to \infty$, we have $P_f \to 0$, and $\bar{R} = \frac{R}{n|\mu|} \to \frac{\sum_{i=1}^{s}(\bar{n}-\bar{t})(\bar{\mu}_i - \bar{n})}{\bar{n}|\bar{\mu}|}$ for $\mu \succeq 2n$.

Theorem 10: When $\tau = t$ and $\mu \succeq 2n$, the coding scheme described above can achieve the upper-bound (11).

Although the special case $\mu \succeq 2n$ is of practical importance (as suggested by several existing constructions for physical-layer network coding [3]), the general case remains open.

VIII. Conclusion

In this work, we have studied the matrix channel $Y = AX + BZ$ over a finite chain ring. Under the assumption that A is uniform over all invertible matrices and BZ is uniform over all matrices of shape t, we have derived tight capacity results and provided polynomial-complexity capacity-achieving coding schemes, which naturally extend the work of [4] from finite fields to finite chain rings. Our extension requires the use of row canonical forms, as well as several new enumeration results and construction methods, for matrices over finite chain rings, which may be of independent interest.

We believe that there is still much work to be done in this area. One direction would be to relax the assumption on A and BZ. An initial step towards this direction was taken in [12].

Acknowledgment

The authors would like to thank Michael Kiermaier for enlightening discussions on row canonical forms for matrices over finite chain rings, and for useful suggestions that greatly improved this work.

References

[1] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[2] M. P. Wilson, K. Narayanan, H. D. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5641–5654, Nov. 2010.

[3] C. Feng, D. Silva, and F. R. Kschischang, "An algebraic approach to physical-layer network coding," *Computing Research Repository (CoRR)*, Aug. 2011, to appear in the IEEE Trans. Inf. Theory. [Online]. Available: http://arxiv.org/abs/1108.1695

[4] D. Silva, F. R. Kschischang, and R. Kötter, "Communication over finite-field matrix channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1296–1305, Mar. 2010.

[5] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. New York: Springer-Verlag, 1999.

[6] C. Feng, R. W. Nóbrega, D. Silva, and F. R. Kschischang, "Communication over finite-chain-ring matrix channels," *Computing Research Repository (CoRR)*, Apr. 2013, submitted to the IEEE Trans. Inf. Theory. [Online]. Available: http://arxiv.org/abs/1304.2523

[7] B. R. McDonald, *Finite Rings with Identity*. Marcel Dekker, Inc., 1974.

[8] T. Honold and I. Landjev, "Linear codes over finite chain rings," *The Electronic Journal of Combinatorics*, vol. 7, 2000.

[9] M. Kiermaier, "Geometric constructions of linear codes over Galois rings of characteristic 4 of high homogeneous minimum distance," Ph.D. dissertation, Universität Bayreuth, 2012.

[10] L. E. Fuller, "A canonical set for matrices over a principal ideal ring modulo m," *Canad. J. Math.*, pp. 54–59, 1955.

[11] G. Birkhoff, "Subgroups of abelian groups," *Proc. London Math. Soc.*, pp. 385–401, 1934.

[12] R. W. Nóbrega, C. Feng, D. Silva, and B. F. Uchôa-Filho, "On multiplicative matrix channels over finite chain rings," to appear in Proc. of IEEE Int. Symp. on Netw. Coding (NetCod'13).