Communication over Finite-Ring Matrix Channels

Chen Feng\textsuperscript{1}  Roberto W. Nóbrega\textsuperscript{2}
Frank R. Kschischang\textsuperscript{1}  Danilo Silva\textsuperscript{2}

\textsuperscript{1}Department of Electrical and Computer Engineering
University of Toronto, Canada

\textsuperscript{2}Department of Electrical Engineering
Federal University of Santa Catarina (UFSC), Brazil

IEEE International Symposium on Information Theory
Istanbul, Turkey, July 12, 2013
Transmitter injects packets (row vectors over $\mathbb{F}_q$)
Intermediate nodes forward random $\mathbb{F}_q$-linear combinations of packets
Errors may also be injected, which randomly mix with the legitimate packets
(Each) receiver gathers as many packets as possible

At any particular receiver:

$$Y = AX + Z$$

where $A$ is a transfer matrix, and $Z$ is some error matrix.
Random-linear network-coding with errors can be formulated as:

\[ Y = AX + Z, \]

where

- all matrices are over \( \mathbb{F}_q \);
- \( X, A, \) and \( Z \) are independent;
- channel law is specified by the distributions of \( A \) and \( Z \).
[SKK10]\textsuperscript{1} considered three variants of \( Y = AX + Z \) over \( \mathbb{F}_q \).

1. \( Y = AX \): \( A \) is invertible, drawn uniformly at random
   exact capacity, code design, encoding-decoding

2. \( Y = X + W \): \( W \) has rank \( t \), drawn uniformly at random
   exact capacity, code design, encoding-decoding

3. \( Y = A(X + W) \): \( A \) invertible, \( W \) rank \( t \), both uniform
   capacity bounds, code design, encoding-decoding

Finite-Ring Matrix Channels

Generalize from finite-field matrix channels to finite-ring matrix channels.

Why?

---

Finite-Ring Matrix Channels

Generalize from finite-field matrix channels to finite-ring matrix channels.

Why?

The motivation comes from physical-layer network coding, in particular, compute-and-forward.²

uncoded modulation:

- $L^2$-QAM $\Rightarrow R = \mathbb{Z}_L[i]$, packet space $= R^m$, where

$$\mathbb{Z}_L[i] \triangleq \{a + bi : a, b \in \mathbb{Z}_L\}.$$ 

nested lattice codes:

- for many practical constructions, we have\(^3\):

$$R = T/\langle \pi^{t_m} \rangle,$$

packet space $= T/\langle \pi^{t_1} \rangle \times \cdots \times T/\langle \pi^{t_m} \rangle$ for some $t_1 \leq \cdots \leq t_m$, where $T$ is a PID.

In all cases, the packet space is $R^\mu$ for some finite chain ring $R$, where

$$R^\mu \triangleq R \times \cdots \times R \times \pi R \times \cdots \times \pi R \times \cdots \times \pi^{s-1} R \times \cdots \times \pi^{s-1} R.$$ 

Finite-Ring Matrix Channels: Packet Space

Example: \( R = \mathbb{Z}_4, \mu = (3, 5), R^\mu = \mathbb{Z}_4^3 \times (2\mathbb{Z}_4)^2 \)

\[
\mathbf{w} = \begin{bmatrix} 1 & 2 & 3 & 0 & 2 \end{bmatrix} \in R^\mu
\]
\[
\mathbf{w} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \end{bmatrix}
\]

So, the packet space \( R^\mu \) can be visualized as

\[
\begin{array}{cccccc}
* & * & * & & & \\
* & * & * & * & & \\
& & & & & \\
\end{array}
\]

In all cases, the packet space is \( R^\mu \) for some finite chain ring \( R \), where

\[
R^\mu \triangleq R \times \cdots \times R \times \pi R \times \cdots \times \pi R \times \cdots \times \pi^{s-1} R \times \cdots \times \pi^{s-1} R. 
\]
The **multiplicative matrix channel (MMC):**

\[ Y = AX \]

where

- \( X, Y \in \mathbb{R}^{n \times \mu} \);
- \( A \) : invertible, uniform;
- \( A \) and \( X \) are independent.
When $R$ reduces to $\mathbb{F}_q$ and $R^{n \times \mu}$ reduces to $\mathbb{F}_q^{n \times m}$:

1. **Exact capacity:** $A$ preserves the row span, so

   \[ C_{\text{MMC}} = \log_q \left( \# \text{ of subspaces of } \mathbb{F}_q^m \right) \]

2. **Capacity-achieving code:** reduced row echelon form (RREF)

3. **Efficient encoding-decoding:**
   - **encoding:**
     \[ X = \begin{bmatrix} I & \text{data} \end{bmatrix} \]
     \[ \begin{array}{c} n \ \ \ m-n \end{array} \]
   - **decoding:** Gaussian elimination (reduction to RREF)
Theorem

The capacity of the MMC, in $q$-ary symbols per channel use, is

$$C_{MMC} = \log_q \left( \text{# of submodules of } R^\mu \right).$$

The # of submodules of $R^\mu$ is $\sum_{\lambda \preceq n, \mu} \binom{\mu}{\lambda}_q$ (see, e.g., [HL00]$^3$), where

$$\binom{\mu}{\lambda}_q = \prod_{i=1}^s q^{(\mu_i - \lambda_i)\lambda_{i-1}} \left[ \frac{\mu_i - \lambda_{i-1}}{\lambda_i - \lambda_{i-1}} \right]_q,$$

and $\binom{m}{k}_q$ is the Gaussian coefficient.

- note: $\lambda \preceq n, \mu$ means $\forall i, \lambda_i \leq n, \mu_i$

---

code design problem $\Rightarrow$ an appropriate generalization of RREF

The presence of zero divisors complicates the matters...

- Over a field, two matrices in echelon form with the same row span will have the same number of nonzero rows—the rank.
- Over a chain ring, this is not the case.

For example, the matrices

$$\begin{bmatrix}
2 & 1 & 1 & 2 \\
0 & 0 & 2 & 2 \\
0 & 0 & 0 & 0
\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}
2 & 1 & 1 & 2 \\
0 & 4 & 0 & 4 \\
0 & 0 & 2 & 2
\end{bmatrix} \quad \text{over} \quad \mathbb{Z}_8$$

have the same row span but not the same number of nonzero rows. So, generalization of RREF seems non-trivial.
code design problem \implies \text{an appropriate generalization of RREF}

There are two matrix canonical forms that generalize RREF:

code design problem ⇒ an appropriate generalization of RREF

There are two matrix canonical forms that generalize RREF:


**Example:** the matrices

\[
\begin{bmatrix}
2 & 1 & 1 & 2 \\
0 & 0 & 2 & 2 \\
0 & 0 & 0 & 0 \\
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
2 & 1 & 1 & 2 \\
0 & 4 & 0 & 4 \\
0 & 0 & 2 & 2 \\
\end{bmatrix}
\]

over \( \mathbb{Z}_8 \)

are Fuller and Howell canonical forms, respectively.
For details, see our paper and/or Kiermaier’s thesis (in German).
MMC: Efficient Encoding-Decoding

First attempt:
- Encoding: transmit a row canonical form (RCF)
- Decoding: reduction to RCF

The decoding complexity is $O(n^2m)$, but the encoding is hard.

Solution:
- Encoding: transmit a principal RCF
- Decoding: reduction to RCF

The encoding complexity is $O(nm)$. Principal RCFs occupy a significant portion of all RCFs.

Hence,
The simple coding scheme asymptotically achieves the capacity.
The additive matrix channel (AMC):

\[ Y = X + W \]

where

- \( X, Y \in \mathbb{R}^{n \times \mu} \);
- \( W \): shape \( \tau \), uniform;
- \( W \) and \( X \) are independent.
The shape is a tuple of non-decreasing integers.

**Example:** \( \mu = (3, 5) \)

\[
R^\mu = R \times \cdots \times R \times \pi R \times \cdots \times \pi R \times \cdots \times \pi^{s-1} R \times \cdots \times \pi^{s-1} R.
\]

The shape of a matrix generalizes the concept of rank.

**Definition** The shape of a matrix \( A \) is defined as the shape of the row span of \( A \), i.e., \( \text{shape}(A) = \text{shape}(\text{row}(A)) \).
Shape of a Matrix

The shape is a tuple of non-decreasing integers.

Example: \( \mu = (3, 5) \)

\[ R^\mu = R \times \cdots \times R \times \pi R \times \cdots \times \pi R \times \cdots \times \pi^{s-1} R \times \cdots \times \pi^{s-1} R. \]

The shape of a module generalizes the concept of dimension.

Theorem

For any finite \( R \)-module \( M \), there is a unique \( \mu \) such that \( M \cong R^\mu \).

We call \( \mu \) the shape of \( M \), and write \( \mu = \text{shape} M \).
The **shape** is a tuple of non-decreasing integers.

**Example:** \( \mu = (3, 5) \)

\[
R^\mu = R \times \cdots \times R \times \pi R \times \cdots \times \pi R \times \cdots \times \pi^{s-1} R \times \cdots \times \pi^{s-1} R.
\]

The **shape of a module** generalizes the concept of **dimension**.

**Theorem**

For any finite \( R \)-module \( M \), there is a unique \( \mu \) such that \( M \cong R^\mu \).

We call \( \mu \) the shape of \( M \), and write \( \mu = \text{shape} \ M \).

The **shape of a matrix** generalizes the concept of **rank**.

**Definition**

The shape of a matrix \( A \) is defined as the shape of the row span of \( A \), i.e., \( \text{shape} \ A = \text{shape} (\text{row} (A)) \).
When $R$ reduces to $\mathbb{F}_q$ and $R^{n \times \mu}$ reduces to $\mathbb{F}_q^{n \times m}$, shape $\tau$ reduces to rank $t$:

1. **Exact capacity:** a discrete symmetric channel

   $$C_{\text{AMC}} = nm - \log_q \left( \# \text{ of matrices of rank } t \text{ in } \mathbb{F}_q^{n \times m} \right)$$

2. **Capacity-approaching code:** $\nu$ is a parameter

   $$X = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{data}$$

3. **Efficient encoding-decoding:**
   - **encoding:** error trapping
   - **decoding:** matrix completion
The AMC is an example of a discrete symmetric channel.

Theorem

The capacity of the AMC, in $q$-ary symbols per channel use, is

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

We need to derive new enumeration results:

- $|R^{n\times\mu}| = q^{n(\mu_1 + \cdots + \mu_s)}$.
- $|\mathcal{T}_\tau(R^{n\times\mu})| = \left[\begin{array}{c} \mu \\
\tau \end{array}\right]_q |R^{n\times\tau}| \prod_{i=0}^{\tau s-1} (1 - q^{i-n})$, where

$$\left[\begin{array}{c} \mu \\
\tau \end{array}\right]_q = \prod_{i=1}^{s} q^{(\mu_i - \tau_i) \tau_i - 1} \left[\begin{array}{c} \mu_i - \tau_i - 1 \\
\tau_i - \tau_i - 1 \end{array}\right]_q.$$
AMC: Capacity-Approaching Code Design

code design problem ⇒ a generalization of error-trapping

Solution: layered error-trapping

Note that every matrix in $\mathbb{R}^{n \times \mu}$ admits a $\pi$-adic decomposition.

Example: $R = \mathbb{Z}_8$, $n = 6$, $\mu = (4, 6, 8)$, $X = X_0 + 2X_1 + 4X_2$

$X_0 = \begin{array}{c|c}
\begin{array}{c}
* * * * \\
* * * * \\
* * * * \\
* * * * \\
\end{array} & 0 \\
\end{array}$

$X_1 = \begin{array}{c|c}
\begin{array}{c}
* * * * * * \\
* * * * * * \\
* * * * * * \\
* * * * * * \\
\end{array} & 0 \\
\end{array}$

$X_2 = \begin{array}{c|c}
\begin{array}{c}
* * * * * * \\
* * * * * * \\
* * * * * * \\
* * * * * * \\
\end{array} & 0 \\
\end{array}$
code design problem ⇒ a generalization of error-trapping

Solution: layered error-trapping

Note that every matrix in $R^{n \times \mu}$ admits a $\pi$-adic decomposition.

**Example:** $R = \mathbb{Z}_8$, $n = 6$, $\mu = (4, 6, 8)$, $X = X_0 + 2X_1 + 4X_2$

\[
\begin{align*}
X_0 &= \begin{array}{c} * * * * * * \\ * * * \\ * * * \\ * * * \\ * * * \\ * * * \\ \end{array} \quad & 0 \\
\mu_1 &\quad \longrightarrow
\end{align*}
\]

\[
\begin{align*}
X_1 &= \begin{array}{c} * * * * * * \\ * * * * * * \\ * * * * * * \\ * * * * * * \\ * * * * * * \\ * * * * * * \\ \end{array} \quad & 0 \\
\mu_2 &\quad \longrightarrow
\end{align*}
\]

\[
\begin{align*}
X_2 &= \begin{array}{c} * * * * * * * * \\ * * * * * * * * \\ * * * * * * * * \\ * * * * * * * * \\ * * * * * * * * \\ * * * * * * * * \\ \end{array} \quad & \begin{array}{c} * * * * * * * * \\ * * * * * * * * \\ * * * * * * * * \\ * * * * * * * * \\ * * * * * * * * \\ * * * * * * * * \\ \end{array} \\
\mu_3 &\quad \longrightarrow
\end{align*}
\]

after error-trapping...

\[
\begin{align*}
X_0 &= \begin{array}{c} * * * \\ * * * \\ * * * \\ * * * \\ * * * \\ * * * \\ \end{array} \quad & 0 \\
\mu_1 &\quad \longrightarrow
\end{align*}
\]

\[
\begin{align*}
X_1 &= \begin{array}{c} * * * \\ * * * \\ * * * \\ * * * \\ * * * \\ * * * \\ \end{array} \quad & 0 \\
\mu_2 &\quad \longrightarrow
\end{align*}
\]

\[
\begin{align*}
X_2 &= \begin{array}{c} * * * * * \\ * * * * * \\ * * * * * \\ * * * * * \\ * * * * * \\ * * * * * \\ \end{array} \quad & \begin{array}{c} * * * * * \\ * * * * * \\ * * * * * \\ * * * * * \\ * * * * * \\ * * * * * \\ \end{array} \\
\mu_3 &\quad \longrightarrow
\end{align*}
\]
Encoding: layered error-trapping, $O(nm)$ complexity
Decoding: multistage matrix completion, $O(n^2m)$ complexity

Example: $R = \mathbb{Z}_8$, $X = X_0 + 2X_1 + 4X_2$. Note that

$$Y = X + W = X_0 + 2X_1 + 4X_2 + W.$$  

1. Take mod 2: $[Y]_2 = X_0 + [W]_2$.
2. Decode $X_0$ by completing $[W]_2$.
3. Clear $X_0$ from $Y$: $Y' = Y - X_0 = 2X_1 + 4X_2 + W$.
4. Take mod 4: $[Y']_4 = 2X_1 + [W]_4$.
5. Decode $2X_1$ by completing $[W]_4$.
6. Clear $X_1$ from $Y'$: $Y'' = Y' - 2X_1 = 4X_2 + W$.
7. We have $Y'' = 4X_2 + W$.
8. Decode $4X_2$ by completing $W$. 
The additive-multiplicative matrix channel (AMMC):

\[ Y = A(X + W) \]

where

- \( X, Y \in R^{n \times \mu} \);
- \( A \): invertible, uniform;
- \( W \): shape \( \tau \), uniform;
- \( A, X \) and \( W \) are independent.

**Remark:** This model is statistically identical to \( Y = AX + Z \).
Theorem

The capacity of the AMMC, in $q$-ary symbols per channel use, is upper-bounded by

$$C_{AMMC} \leq \sum_{i=1}^{s} (\mu_i - \xi_i)\xi_i + \sum_{i=1}^{s} (n - \mu_i)\tau_i + 2s \log_q 4 + \log_q \binom{n+s}{s}$$

$$+ \log_q \binom{\tau s + s}{s} - \log_q \prod_{i=0}^{\tau s - 1} (1 - q^{i-n}), \text{ where } \xi_i = \min\{n, \left\lfloor \mu_i/2 \right\rfloor\}.$$  

In particular, when $\mu \geq 2n$, the upper bound reduces to

$$C_{AMMC} \leq \sum_{i=1}^{s} (n - \tau_i)(\mu_i - n) + 2s \log_q 4$$

$$+ \log_q \binom{n+s}{s} + \log_q \binom{\tau s + s}{s} - \log_q \prod_{i=0}^{\tau s - 1} (1 - q^{i-n}).$$
AMMC: Coding Scheme

coding scheme = principal RCFs + layered error-trapping

However, the combination turns out to be non-trivial. Hence, we focus on the special case when $\tau = (t, \ldots, t)$.

- **Encoding:**
  
  \[
  X = \begin{bmatrix}
  0 & \cdots & 0 \\
  \vdots \\
  0 \\
  \end{bmatrix}
  \]

- **Decoding:** upon receiving $Y = A(X + W)$, the decoder simply computes the RCF of $Y$, which exposes $\bar{X}$ with high probability. This simple coding scheme asymptotically achieves the capacity for the special case when $\tau = (t, \ldots, t)$ and $\mu \geq 2n$. 
Conclusion

- studied three variants of finite-ring matrix channels
  - exact capacities and an upper bound
  - capacity-achieving codes
  - efficient encoding-decoding methods
- refined some linear algebra tools over finite chain rings
  - row canonical form with a new proof for uniqueness
  - construction of principal RCFs
  - new enumeration results
- open problems:
  - Can we handle $Y = A(X + W)$ for general shapes?
  - What if $A$ is not invertible?
Back-up Slides
Let $R$ be a finite chain ring, where $\langle \pi \rangle$ is the unique maximal ideal, $q$ is the order of the residue field $R/\langle \pi \rangle$, and $s$ is the number of proper ideals.

Notation: $(q, s)$ chain ring.

$\pi$-adic decomposition

Let $R(\mathbb{R}, \pi)$ be a complete set of residues with respect to $\pi$. Then every element $r \in R$ can be written uniquely as $r = r_0 + r_1 \pi + r_2 \pi^2 + \cdots + r_{s-1} \pi^{s-1}$ where $r_i \in R(\mathbb{R}, \pi)$. 

$R = \langle \pi^0 \rangle$

$\langle \pi \rangle$

$\langle \pi^2 \rangle$

$\langle \pi^{s-1} \rangle$

$\{0\} = \langle \pi^s \rangle$
Let $R$ be a finite chain ring, where

- $\langle \pi \rangle$ is the unique maximal ideal,
- $q$ is the order of the residue field $R/\langle \pi \rangle$,
- $s$ is the number of proper ideals.

Notation: $(q, s)$ chain ring.
Let $R$ be a finite chain ring, where

- $\langle \pi \rangle$ is the unique maximal ideal,
- $q$ is the order of the residue field $R/\langle \pi \rangle$,
- $s$ is the number of proper ideals.

Notation: $(q, s)$ chain ring.

**$\pi$-adic decomposition**

Let $\mathcal{R}(R, \pi)$ be a complete set of residues with respect to $\pi$. Then every element $r \in R$ can be written uniquely as

$$r = r_0 + r_1 \pi + r_2 \pi^2 + \cdots + r_{s-1} \pi^{s-1}$$

where $r_i \in \mathcal{R}(R, \pi)$. 
Definition

The degree, \( \text{deg}(r) \), of a nonzero element \( r \in R^* \), where

\[
r = r_0 + r_1 \pi + \cdots + r_{s-1} \pi^{s-1},
\]

is defined as the least index \( j \) for which \( r_j \neq 0 \).

- by convention, \( \text{deg}(0) = s \)
- units have degree zero
- elements of the same degree are associates
- \( a \) divides \( b \) if and only if \( \text{deg}(a) \leq \text{deg}(b) \)
Row Canonical Form

A matrix $A$ is in **row canonical form** if it satisfies the following conditions.

1. Nonzero rows of $A$ are above any zero rows.
2. The pivot of a row is of the form $\pi^\ell$, and is the leftmost entry of the least degree.
3. For every pivot (say $\pi^\ell$), 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 residues of $\pi^\ell$.
4. 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.

For example, over $\mathbb{Z}_8$,

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

is in row canonical form.
Reduction to Row Canonical Form: Example

Reduction is a variant of **Gaussian elimination**. An example over $\mathbb{Z}_8$:

\[
A = \begin{bmatrix}
4 & 6 & 2 & 1 \\
0 & 0 & 0 & 2 \\
2 & 4 & 6 & 1 \\
2 & 0 & 2 & 1 \\
\end{bmatrix} \quad \rightarrow \quad A_1 = \begin{bmatrix}
0 & 4 & 4 & 0 \\
6 & 6 & 4 & 0 \\
6 & 2 & 0 & 0 \\
\end{bmatrix} \quad \rightarrow \\
A'_1 = \begin{bmatrix}
4 & 6 & 2 & 1 \\
2 & 2 & 4 & 0 \\
0 & 4 & 4 & 0 \\
6 & 2 & 0 & 0 \\
\end{bmatrix} \quad \rightarrow \quad A_2 = \begin{bmatrix}
0 & 2 & 2 & 1 \\
2 & 2 & 4 & 0 \\
0 & 4 & 4 & 0 \\
0 & 4 & 4 & 0 \\
\end{bmatrix} \\
A_3 = \begin{bmatrix}
0 & 2 & 2 & 1 \\
2 & 2 & 4 & 0 \\
0 & 4 & 4 & 0 \\
0 & 0 & 0 & 0 \\
\end{bmatrix} \quad \text{which is in row canonical form.}
\]

Row canonical form is not necessarily an echelon form!
Construction of Principal RCFs

Definition

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

$$d_1, \ldots, d_r = 1, \ldots, 1, \pi, \ldots, \pi, \ldots, \pi^{s-1}, \ldots, \pi^{s-1}, 0, \ldots, 0.$$ 

All principal RCFs in $\mathcal{T}_\kappa(R^{n \times \mu})$ can be constructed via a \(\pi\)-adic decomposition $X = X_0 + \pi X_1 + \cdots + \pi^{s-1} X_{s-1}$.

Example: $s = 3$, $n = 6$, $\mu = (4, 6, 8)$, and $\kappa = (2, 3, 4)$