Algebraic Codes for Random Linear Network Coding

Maximilien Gadouleau

Crestic
Université de Reims Champagne-Ardenne

August 21, 2009
Outline

Introduction
- Random Linear Network Coding
- Error control codes for RLNC
- Overview of Results

Rank Metric Codes
- Rank Metric Codes

CDCs and Subspace Codes
- DEP of CDCs
- Packing Properties
- Covering Properties

Conclusion and Future Work
Outline

Introduction

Random Linear Network Coding
Error control codes for RLNC
Overview of Results

Rank Metric Codes

Rank Metric Codes

CDCs and Subspace Codes

DEP of CDCs
Packing Properties
Covering Properties

Conclusion and Future Work
Network coding

- Traditional routing: store and forward. Does not achieve the highest throughput for multicast
- Network coding allows intermediate nodes to combine packets
- Only perform linear combinations of packets. $\Rightarrow$ Linear network coding
- Results in higher throughput for multicast, robustness, and adaptability
Network coding example
Random linear network coding (RLNC)

- Fixing the linear combinations requires knowledge of the network and is too rigid
- Solution: choosing linear combinations at random
- In example, 1 chance out of 3 to be successful. However, probability of success tends to 1 with field size
- Header to keep track of linear combinations.
  \[ (\mathbf{I}_r | \mathbf{M}) \text{, where } \mathbf{M} = \begin{pmatrix} a \\ b \end{pmatrix} \]
- Easy to decode: receive \((\mathbf{L} | \mathbf{LM})\), compute \(\mathbf{L}^{-1}(\mathbf{LM})\)
Error control for RLNC

- RLNC highly sensitive to errors
- Many types of errors
  - Link errors
  - Packet losses
  - Malfunctioning nodes
  - Adversary on the network
- Errors are highly detrimental: one packet error can corrupt all packets after linear combination
Error in network coding example

Error in network coding example

\[ b' = a + b - a' \]
Operator channel

- RLNC preserves the row space of the input matrix.  
  ⇒ Modeled as transmission of a linear subspace
- Two ways to modify the subspace
  - Erasure: deletion of a dimension
  - Error: addition of a dimension
- Operator channel: send $U$, receive

$$V = H(U) \oplus E.$$  

$H$ performs erasures, $E$ are errors
- Error control in RLNC is a coding theory problem, where codewords are subspaces
Projective space and subspace codes

- **Subspace code**: set of subspaces of $\text{GF}(q)^n$.
  $\Rightarrow$ Subset of the projective space
- **Metrics**:
  - subspace distance $d_S$ for network errors
  - injection distance $d_I$ for adversarial channels
- $V$ obtained from $U$ after $\epsilon$ erasures and $\rho$ errors.
  
  \[
  d_S(U, V) = \epsilon + \rho = \dim(U + V) - \dim(U \cap V) \\
  d_I(U, V) = \max\{\epsilon, \rho\} = \max\{\dim(U), \dim(V)\} - \dim(U \cap V)
  \]
- For any subspace code, $d_I(C) \leq d_S(C) \leq 2d_I(C)$
Grassmannian and CDCs

- Constant-dimension code (CDC): set of subspaces of $\text{GF}(q)^n$ with dimension $r$.
  ⇒ Subset of the Grassmannian
- $d_S(U^\perp, V^\perp) = d_S(U, V)$ ⇒ assume $r \leq \frac{n}{2}$
- Advantages:
  - $d_S(C) = 2d_1(C)$ for a CDC.
    ⇒ Versatility, and maximum $d_S$
  - Simplified decoding procedure
- Relation to rank metric codes through the lifting operation
Rank metric codes

- Rank metric code: set of matrices in $\text{GF}(q)^{m \times n}$
- Applications to data storage, cryptography, space-time coding
- Rank distance: $d_R(M, N) = \text{rk}(M - N)$.
  $d_R(M^T, N^T) = d_R(M, N) \Rightarrow$ assume $m \geq n$
- Optimal codes: MRD codes. Gabidulin codes: RS-like MRD codes. Decoding algorithms for Gabidulin codes
Liftings of rank metric codes

- Lifting: $M \in \text{GF}(q)^{m \times n}$, $I(M)$ is the row space of $(I_m | M)$
- $d_I(I(M), I(N)) = d_R(M, N)$
- $C$ rank metric code in $\text{GF}(q)^{r \times (n-r)}$.
  $\Rightarrow I(C)$ CDC, and $d_S(I(C)) = 2d_I(I(C)) = 2d_R(C)$
- Error control in RLNC using liftings is a rank metric problem
- KK code: lifting of (transposed) Gabidulin code.
  Nearly optimal CDC, bounded subspace distance decoding algorithm
Overview of results

- **Rank metric codes**
  - Covering properties \((IT\ 08\ +\ to\ appear\ in\ CL\ 09)\)
  - MacWilliams identity \((EURASIP\ 08)\)
  - DEP \((IT\ 08)\)
  - Constant-rank codes \((submitted\ to\ IT)\)

- **CDCs**
  - Construction and covering properties of CDCs \((submitted\ to\ IT)\)
  - DEP of CDCs \((submitted\ to\ IT)\)

- **Subspace codes**: Packing and covering properties \((submitted\ to\ IT)\)
Outline

Introduction
  Random Linear Network Coding
  Error control codes for RLNC
  Overview of Results

Rank Metric Codes
  Rank Metric Codes

CDCs and Subspace Codes
  DEP of CDCs
  Packing Properties
  Covering Properties

Conclusion and Future Work
Rank metric codes

- Sphere covering
  - Geometric properties (volumes of balls and their intersections)
  - Bounds and constructions
  - Covering radius of known codes
- MacWilliams identity
  - Polynomial form previously unknown
  - Binomial and power moments of the rank distribution
- Decoder error probability
  - Exact DEP of codes given their distance distribution
  - Upper bounds on the DEP
  - MRD codes have the highest DEP up to a scalar
CRCs

- Motivation: study CDCs and their relations to rank metric codes
- Constant-rank code (CRC): rank metric code where all codewords have the same rank
- Row space of a CRC is a CDC with related minimum distance
- Optimal CRC in GF($q^{m\times n}$) with rank $r$ and $d_R = d + r$ for $m$ large enough.
  $\Rightarrow$ Row space: optimal CDC of length $n$ and dimension $r$ with $d_1 = d$
- We also study CRCs. Many results for $d_R \leq r$, fewer for $d_R > r$
Outline

Introduction
  Random Linear Network Coding
  Error control codes for RLNC
  Overview of Results

Rank Metric Codes
  Rank Metric Codes

CDCs and Subspace Codes
  DEP of CDCs
  Packing Properties
  Covering Properties

Conclusion and Future Work
DEP of CDCs

- Two decoders: bounded subspace distance, bounded injection distance
- Bounded distance decoder (BDD): finds the unique nearest codeword at distance $\leq \left\lfloor \frac{d-1}{2} \right\rfloor$
- Decoder has three outputs: success, failure, or error. Success iff received subspace at distance $\leq \left\lfloor \frac{d-1}{2} \right\rfloor$. Error or failure otherwise
- Error more detrimental than failure.
  $\Rightarrow$ Focus on decoder error probability (DEP)
DEP of CDCs

- Symmetric operator channel: all subspaces after $\epsilon$ erasures and $\rho$ errors equiprobable.
  $\Rightarrow$ Same dimension and same distance are equiprobable
- Results:
  - Exact DEP for all CDCs based on distance distribution
  - Bounds on the DEP of all liftings
  - Liftings of MRD have maximum DEP up to a scalar
- Bounded subspace distance decoder corrects more errors and has higher DEP than bounded injection decoder
Sphere packing

- How many spheres of same radius can be packed in a space? ⇒ Max cardinality of a code with a given minimum distance $A(d)$
- Crucial for error correcting code design
Construction of CDCs

- **KK code** $\mathcal{E}^0$: $\mathcal{D}^0 \subseteq \text{GF}(q)^{r \times (n-r)}$ is a (transposed) Gabidulin code with $d_R = d$, where $d_S = 2d_I = 2d$.

  $E_{0,j}^0 = R(\mathbf{I}_r|\mathbf{D}_j^0)$

- **Layer** $\mathcal{E}^k$ for $1 \leq k \leq \left\lfloor \frac{r}{d} \right\rfloor$: $\mathcal{C}^k \subseteq \text{GF}(q)^{(r-kd) \times kd}$, $\mathcal{D}^k \subseteq \text{GF}(q)^{r \times (n-r-kd)}$ are (transposed) Gabidulin codes with $d_R = d$ or trivial codes.

  $E_{i,j}^k = R \left( \begin{array}{c|c|c}
  \mathbf{I}_{r-kd} & \mathbf{C}_i^k & 0 \\
  0 & 0 & \mathbf{I}_{kd} \\
  \end{array} \right) \mathbf{D}_j^k$

- **Augmented KK code** $\mathcal{E} = \bigcup_{k=0}^{\left\lfloor \frac{r}{d} \right\rfloor} \mathcal{E}^k$
Augmented KK codes

- $d_I(R(A), R(B)) \geq d_I(R(A_1), R(B_1))$, where $A_1, B_1$ are subsets of columns
  $\Rightarrow d_I = d, d_S = 2d_I = 2d$

- Properly contains a KK code. Greatest cardinality for $d < r$

- Enhanced BDD for augmented KK code:
  - Corrects more than bounded subspace distance decoder
  - Complexity on the same order as BDD of KK code
**EBDD**($k, A$)

Input: $k$ and $A = (A_1|A_2|A_3)$, $A_1 \in \text{GF}(q)^a \times (r-kd)$, $A_2 \in \text{GF}(q)^a \times kd$, $A_3 \in \text{GF}(q)^a \times (n-r)$

Output: $(E_{i,j}^k, d_k, f_k)$

1. If $k = 0$, decoder of $E^0 \Rightarrow E_{0,j}^0$, calculate $d_k = d_S(R(A), E_{0,j}^0)$, and return $(E_{0,j}^0, d_k, 0)$

2. Decoder of $I(C^k)$ on $(A_1|A_2) \Rightarrow C_i^k$

3. Decoder of $I(D^k)$ on $(A_1|A_3) \Rightarrow D_j^k$

4. Calculate $d_k = d_S(R(A), E_{i,j}^k)$,
   $$f_k = 2d - \max\{d_S(R(A_1|A_2), I(C_i^k)), d_S(R(A_1|A_3), I(D_j^k))\},$$
   and return $(E_{i,j}^k, d_k, f_k)$
Decoder for $\mathcal{E}$

Input: $A = (A_0|A_3)$, $A_0 \in \text{GF}(q)^{a \times r}$, $A_3 \in \text{GF}(q)^{a \times (n-r)}$
Output: The unique nearest codeword in $\mathcal{E}$ from $R(A)$ or failure

1. If $\text{rk}(A) < r - d + 1$, return failure

2. Calculate $r - \text{rk}(A_0) = ld + m$, with $0 \leq l \leq \lfloor \frac{r}{d} \rfloor$ and $0 \leq m < d$

3. $\text{EBDD}(l, A) \Rightarrow (E_{i,j}^l, d_l, f_l)$. If $d_l \leq d - 1$, return $E_{i,j}^l$

4. If $m = 0$, return failure.
   Otherwise $\text{EBDD}(l + 1, A) \Rightarrow (E_{s,t}^{l+1}, d_{l+1}, f_{l+1})$.
   If $d_{l+1} \leq d - 1$, return $E_{s,t}^{l+1}$

5. If $d_l < \min\{d + m, f_l, d_{l+1}, f_{l+1}, 2d - m\}$, return $E_{i,j}^l$.
   If $d_{l+1} < \min\{d + m, d_l, f_l, f_{l+1}, 2d - m\}$, return $E_{s,t}^{l+1}$

6. Return failure
Packing properties of subspace codes

- Max cardinality of a subspace code (resp., CDC) with minimum injection distance $d$: $A_I(d)$ (resp., $A_C(r, d)$)
- Upper bound:
  \[
  A_I(d) \leq \sum_r A_C(r, d) \sim A_C \left( \left\lfloor \frac{n}{2} \right\rfloor, d \right)
  \]
- Lower bound: $A_I(d) \geq A_C(r, d)$. Nearly tight for $r = \left\lfloor \frac{n}{2} \right\rfloor$
- Same situation for the subspace metric: $A_S(d) \sim A_C \left( \left\lfloor \frac{n}{2} \right\rfloor, \left\lceil \frac{d}{2} \right\rceil \right)$
- CDCs optimal up to a scalar for both metrics.
  \[\Rightarrow\] Nearly optimal codes in the injection metric with $d_S = 2d_I$
Sphere covering

- How many spheres needed to cover a space?
  \[ \Rightarrow \text{Min cardinality of a code with a given covering radius } K(\rho) \]
- Applications: minimum distance decoding, decoding with erasures, data compression
- Sphere covering bound: \( K(\rho) \geq \frac{|E|}{\max V(\rho)} \)
- Greedy upper bound: \( K(\rho) \lesssim \frac{|E|}{\min V(\rho)} \)
Sphere covering example

- Sphere covering bound: \( K(1) \geq \lceil \frac{7}{4} \rceil = 2 \)
- Greedy construction: \( K(1) \leq 4 \)
- However, \( K(1) = 3 \)
Covering properties of CDCs

- We study how CDCs cover the Grassmannian. Min cardinality of a CDC with covering radius $\rho$: $K_C(r, \rho)$
- Volume of a ball equal for all centers. $\Rightarrow$ Sphere covering bound asymptotically tight
- The Grassmannian with the injection metric is an association scheme. Intersection of two spheres with radii $u$ and $s$ and distance between their centers $d$: $J_C(u, s, d)$
- Refinement of the sphere covering bound based on linear programming
Refined sphere covering bound for CDCs

- For $0 \leq \delta \leq \rho$, let $T_\delta = \min \sum_{i=0}^{r} A_i(\delta)$, subject to

  $A_i(\delta) = 0$ for $0 \leq i \leq \delta - 1$,

  $1 \leq A_\delta(\delta) \leq N_C(\delta)$,

  $0 \leq A_i(\delta) \leq N_C(i)$ for $\delta + 1 \leq i \leq r$,

  $\sum_{i=0}^{r} A_i(\delta) \sum_{s=0}^{\rho} J_C(l, s, i) \geq N_C(l)$ for $0 \leq l \leq r$.

Then $K_C(r, \rho) \geq \max_{0 \leq \delta \leq \rho} T_\delta$

- Proof: $U$ subspace at distance $\delta$ from a CDC with covering radius $\rho$, $A_i(\delta)$ distance distribution from $U$
Covering with liftings

- Liftings have covering radius $r$. Proof: consider $R(0_r|M)$.  
  $\Rightarrow$ Liftings are not optimal packing CDCs

- Construction of covering CDCs based on liftings.  
  Min cardinality of a code with rank radius $\rho$: $K_R(\rho)$

- $C$ optimal rank metric covering code.  
  $L(C)$: all possible permutations of liftings of codewords in $C$.  
  $\Rightarrow |L(C)| \leq \binom{n}{r} K_R(\rho)$

- $L(C)$ has covering radius $\rho$. Proof: all subspaces are permutations of liftings

- Asymptotically optimal but not good for finite values
Covering in the injection metric

- Min cardinality of a subspace code with injection covering radius $\rho$: $K_I(\rho)$
- Volume of a ball depends on the dimension of its center
- However, $V_I(r, \rho) \sim f(\rho)$.
  $\Rightarrow$ Sphere covering bound asymptotically tight
- Construction using covering CDCs: $K_I(\rho) \leq \sum_{r=0}^{n} K_C(r, \rho)$. Asymptotically optimal
Covering in the subspace metric

- Min cardinality of a subspace code with subspace covering radius $\rho$: $K_S(\rho)$
- Volume of a ball depends on the dimension of its center and $V_S(r, \rho) \sim g(r, \rho)$
- Sphere covering bound, construction using CDCs: not asymptotically tight
- We derive asymptotically tight lower and upper bounds
Outline

Introduction
  Random Linear Network Coding
  Error control codes for RLNC
  Overview of Results

Rank Metric Codes
  Rank Metric Codes

CDCs and Subspace Codes
  DEP of CDCs
  Packing Properties
  Covering Properties

Conclusion and Future Work
Conclusion and Future Work

- Summary of results
- Future work:
  - RLNC: implementation and related problems
  - Constructions of packing and covering CDCs and subspace codes
  - Connection between subspace codes and rank metric codes
  - Decoding of error control codes for RLNC
Definitions of subspace and injection metrics

\[ d_S(U, V) = \dim(U + V) - \dim(U \cap V) \]
\[ = \dim(U) + \dim(V) - 2 \dim(U \cap V) \]
\[ = 2 \dim(U + V) - \dim(U) - \dim(V) \]

\[ d_I(U, V) = \frac{1}{2} d_S(U, V) + \frac{1}{2} |\dim(U) - \dim(V)| \]
\[ = \max\{\dim(U), \dim(V)\} - \dim(U \cap V) \]
\[ = \dim(U + V) - \min\{\dim(U), \dim(V)\} \]

\[ |\dim(U) - \dim(V)| \leq d_I(U, V) \leq d_S(U, V) \leq \dim(U) + \dim(V) \]
Injection distance of liftings

For $M, N \in \text{GF}(q)^{r \times (n-r)}$,

$$d_I(I(M), I(N)) = \text{rk} \left( \begin{array}{c|c} I_r & M \\ \hline I_r & N \end{array} \right) - r$$

$$= \text{rk} \left( \begin{array}{c|c} I_r & M \\ \hline 0 & M - N \end{array} \right) - r$$

$$= \text{rk}(M - N)$$
CRCs

Theorem

For all $X, Y \in \text{GF}(q)^{m \times n}$,

$$
d_I(R(X), R(Y)) + d_I(C(X), C(Y)) - |\text{rk}(X) - \text{rk}(Y)|
\leq d_R(X, Y)
\leq \min\{d_I(R(X), R(Y)), d_I(C(X), C(Y))\} + \min\{\text{rk}(X), \text{rk}(Y)\}.
$$

Therefore, for a CRC $C$:

$$
d_I(R(C)) + d_I(C(C)) \leq d_R(C) \leq \min\{d_I(R(C)), d_I(C(C))\} + r.
$$
Optimal CRCs

Proposition

For all $q$, $2 \leq d \leq r \leq n \leq m$,

$$\min\{A_C(q, n, r, d), A_C(q, m, r, r)\} \leq A_R(q, m, n, r, d+r) \leq A_C(q, n, r, d).$$

For $2r \leq n$, $A_C(q, n, r, d) \sim q^{(n-r)(r-d+1)}$, $A_C(q, m, r, r) \sim q^{m-r}$.

$\Rightarrow$ For $m \geq (n-r)(r-d+1)+r+1$, $A_C(q, m, r, r) \geq A_C(q, n, r, d)$
Lemma

Suppose the output of $EBDD(k, A)$ is $(E_{i,j}^k, d_k, f_k)$, then $d_S(R(A), E_{u,v}^k) \geq f_k$ for any $E_{u,v}^k \in \mathcal{E}^k$ provided $(u, v) \neq (i, j)$.

Proof.

We have

$$f_k = \min\{2d - d_S(R(A_1|A_2), I(C_i^k)), 2d - d_S(R(A_1|A_3), I(D_j^k))\}.$$  

When $u \neq i$,

$$d_S(R(A), E_{u,v}^k) \geq d_S(R(A_1|A_2), I(C_u^k))$$
$$\geq d_S(I(C_i^k), I(C_u^k)) - d_S(R(A_1|A_2), I(C_i^k))$$
$$\geq 2d - d_S(R(A_1|A_2), I(C_i^k)) \geq f_k.$$

Similarly, when $v \neq j$, we obtain
$$d_S(R(A), E_{u,v}^k) \geq 2d - d_S(R(A_1|A_3), I(D_j^k)) \geq f_k.$$  