Rank Minimization over Finite Fields: Fundamental Limits and Coding-Theoretic Interpretations

Vincent Tan

Department of Electrical and Computer Engineering,
University of Wisconsin-Madison

UIUC Comm Seminar (Apr 25, 2011)

Joint work with

Laura Balzano

Stark C. Draper

Outline

1. Motivation

Outline

1. Motivation
2. Problem Setup and Summary of Main Results

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse
4. Achievability for Uniform Model

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse
4. Achievability for Uniform Model
5. Achievability for Sparse Model

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse
4. Achievability for Uniform Model
5. Achievability for Sparse Model
6. Coding-Theoretic Interpretations

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse
4. Achievability for Uniform Model
5. Achievability for Sparse Model
6. Coding-Theoretic Interpretations

Connection between Two Areas of Study

Area of Study	Matrix Completion Rank Minimization	Rank-Metric Codes

Connection between Two Areas of Study

Area of Study	**Matrix Completion Rank Minimization**	**Rank-Metric Codes**
Applications	Collaborative Filtering Minimal Realization	Crisscross Error Correction Network Coding

Connection between Two Areas of Study

Area of Study	Matrix Completion Rank Minimization	Rank-Metric Codes
Applications	Collaborative Filtering Minimal Realization	Crisscross Error Correction Network Coding
Field	Reals $\mathbb{R}$, Complex $\mathbb{C}$	Finite Field $\mathbb{F}_q$

Connection between Two Areas of Study

Area of Study	Matrix Completion Rank Minimization	Rank-Metric Codes
Applications	Collaborative Filtering Minimal Realization	Crisscross Error Correction Network Coding
Field	Reals $\mathbb{R}$, Complex $\mathbb{C}$	Finite Field $\mathbb{F}_q$
Techniques	Functional Analysis	Algebraic coding

Connection between Two Areas of Study

Area of Study	**Matrix Completion Rank Minimization**	**Rank-Metric Codes**
Applications	Collaborative Filtering Minimal Realization	Crisscross Error Correction Network Coding
Field	Reals $\mathbb{R}$, Complex $\mathbb{C}$	Finite Field $\mathbb{F}_q$
Techniques	Functional Analysis	Algebraic coding
Decoding	Nuclear Norm Convex Optimization	Berlekamp-Massey Error Trapping

Connection between Two Areas of Study

Area of Study	Matrix Completion Rank Minimization	Rank-Metric Codes
Applications	Collaborative Filtering Minimal Realization	Crisscross Error Correction Network Coding
Field	Reals $\mathbb{R}$, Complex $\mathbb{C}$	Finite Field $\mathbb{F}_q$
Techniques	Functional Analysis	Algebraic coding
Decoding	Nuclear Norm Convex Optimization	Berlekamp-Massey Error Trapping

Can we draw analogies between the two areas of study?

Connection between Two Areas of Study

Area of Study	Matrix Completion Rank Minimization	Rank-Metric Codes
Applications	Collaborative Filtering Minimal Realization	Crisscross Error Correction Network Coding
Field	Reals $\mathbb{R}$, Complex $\mathbb{C}$	Finite Field $\mathbb{F}_q$
Techniques	Functional Analysis	Algebraic coding
Decoding	Nuclear Norm Convex Optimization	Berlekamp-Massey Error Trapping

Can we draw analogies between the two areas of study?

Rank Minimization over finite fields

Rank-metric codes

Definition

The *rank distance* *between matrices* $\mathbf{A}, \mathbf{B} \in \mathbb{F}_q^{m \times n}$ *is defined as*

$$\Delta(\mathbf{A}, \mathbf{B}) := \mathrm{rank}(\mathbf{A} - \mathbf{B}).$$

Rank-metric codes

Definition

The rank distance between matrices $\mathbf{A}, \mathbf{B} \in \mathbb{F}_q^{m \times n}$ *is defined as*

$$\Delta(\mathbf{A}, \mathbf{B}) := \mathrm{rank}(\mathbf{A} - \mathbf{B}).$$

Fact: $\Delta(\mathbf{A}, \mathbf{B})$ is a metric

Rank-metric codes

Definition

The rank distance between matrices $\mathbf{A}, \mathbf{B} \in \mathbb{F}_q^{m \times n}$ is defined as

$$\Delta(\mathbf{A}, \mathbf{B}) := \mathrm{rank}(\mathbf{A} - \mathbf{B}).$$

Fact: $\Delta(\mathbf{A}, \mathbf{B})$ is a metric

Definition

A rank-metric code is $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m \times n}$ endowed with the rank distance.

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$
- Receive some received word $\mathbf{R} \in \mathbb{F}_q^{m \times n}$

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$
- Receive some received word $\mathbf{R} \in \mathbb{F}_q^{m \times n}$
- Decoding problem (under mild conditions) is

$$\hat{\mathbf{C}} = \arg\min_{\mathbf{C} \in \mathcal{C}} \ \mathrm{rank}(\mathbf{R} - \mathbf{C})$$

- Minimum distance decoding since rank induces a metric

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$
- Receive some received word $\mathbf{R} \in \mathbb{F}_q^{m \times n}$
- Decoding problem (under mild conditions) is

$$\hat{\mathbf{C}} = \arg\min_{\mathbf{C} \in \mathcal{C}} \ \mathrm{rank}(\mathbf{R} - \mathbf{C})$$

- Minimum distance decoding since rank induces a metric
- $\mathbf{X} \equiv \mathbf{R} - \mathbf{C}$ is known as the error matrix – low-rank

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$
- Receive some received word $\mathbf{R} \in \mathbb{F}_q^{m \times n}$
- Decoding problem (under mild conditions) is

$$\hat{\mathbf{C}} = \arg\min_{\mathbf{C} \in \mathcal{C}} \ \mathrm{rank}(\mathbf{R} - \mathbf{C})$$

- Minimum distance decoding since rank induces a metric
- $\mathbf{X} \equiv \mathbf{R} - \mathbf{C}$ is known as the error matrix – low-rank
- Assume linear

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$
- Receive some received word $\mathbf{R} \in \mathbb{F}_q^{m \times n}$
- Decoding problem (under mild conditions) is

$$\hat{\mathbf{C}} = \arg\min_{\mathbf{C} \in \mathcal{C}} \ \mathrm{rank}(\mathbf{R} - \mathbf{C})$$

- Minimum distance decoding since rank induces a metric
- $\mathbf{X} \equiv \mathbf{R} - \mathbf{C}$ is known as the error matrix – low-rank
- Assume linear
- Rank minimization over finite field problem

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X} \in \text{coset}} \ \mathrm{rank}(\mathbf{X})$$

When do low-rank errors occur?

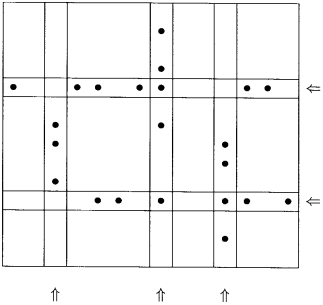
When do low-rank errors occur?

- Probabilistic crisscross error correction

When do low-rank errors occur?

- Probabilistic crisscross error correction

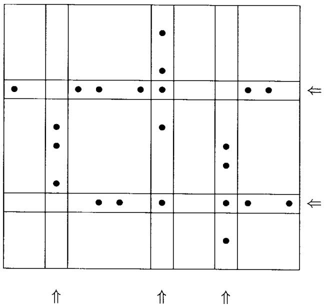
Roth, IT Trans 1997



When do low-rank errors occur?

- Probabilistic crisscross error correction

Roth, IT Trans 1997

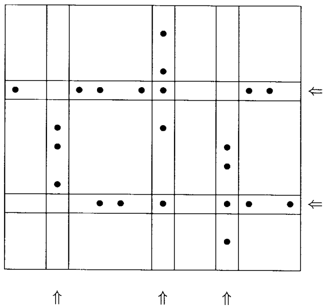


- Data storage applications: Data stored in arrays

When do low-rank errors occur?

- Probabilistic crisscross error correction

Roth, IT Trans 1997

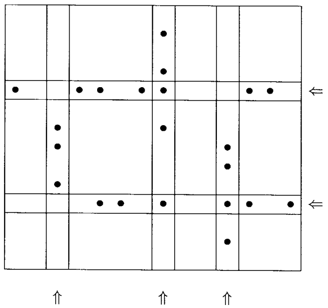


- Data storage applications: Data stored in arrays
- Error patterns confined to two rows and three columns above

When do low-rank errors occur?

- Probabilistic crisscross error correction

Roth, IT Trans 1997



- Data storage applications: Data stored in arrays
- Error patterns confined to two rows and three columns above
- Low-rank errors

Index coding with side information

Index coding with side information

- k-receiver deterministic broadcast channel
- Birk and Kol (IT Trans 06) and Bar-Yossef et al. (FOCS 06)

Index coding with side information

- k-receiver deterministic broadcast channel
- Birk and Kol (IT Trans 06) and Bar-Yossef et al. (FOCS 06)
- Transmitter has k bits of information $x^k = \{0, 1\}^k$

Index coding with side information

- k-receiver deterministic broadcast channel
- Birk and Kol (IT Trans 06) and Bar-Yossef et al. (FOCS 06)
- Transmitter has k bits of information $x^k = \{0,1\}^k$
- Wishes to transmit x_i to receiver i

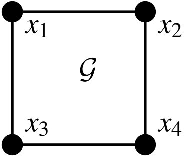
Index coding with side information

- k-receiver deterministic broadcast channel
- Birk and Kol (IT Trans 06) and Bar-Yossef et al. (FOCS 06)
- Transmitter has k bits of information $x^k = \{0,1\}^k$
- Wishes to transmit x_i to receiver i
- Receiver i also has subset of other coordinates

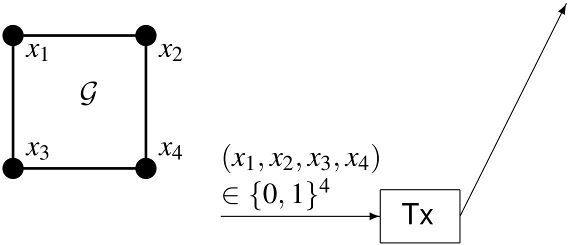
Index coding with side information

- k-receiver deterministic broadcast channel
- Birk and Kol (IT Trans 06) and Bar-Yossef et al. (FOCS 06)
- Transmitter has k bits of information $x^k = \{0,1\}^k$
- Wishes to transmit x_i to receiver i
- Receiver i also has subset of other coordinates
- Side information $\{x_j : (i,j) \in \mathcal{G}\}$

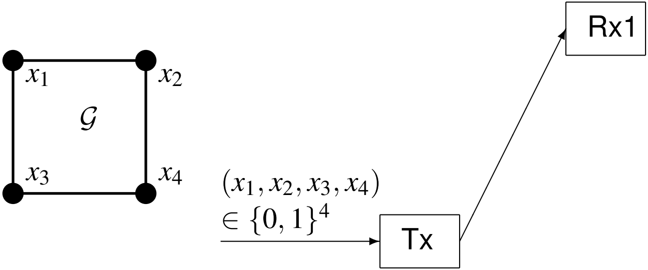
Index coding with side information



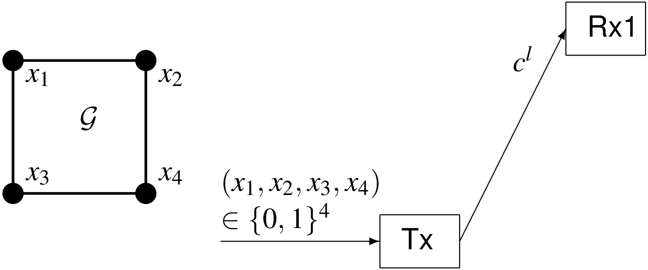
Index coding with side information



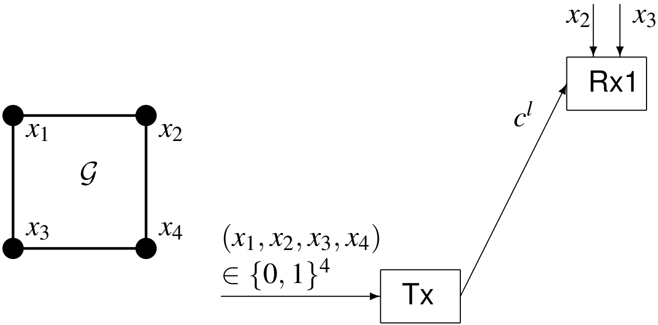
Index coding with side information



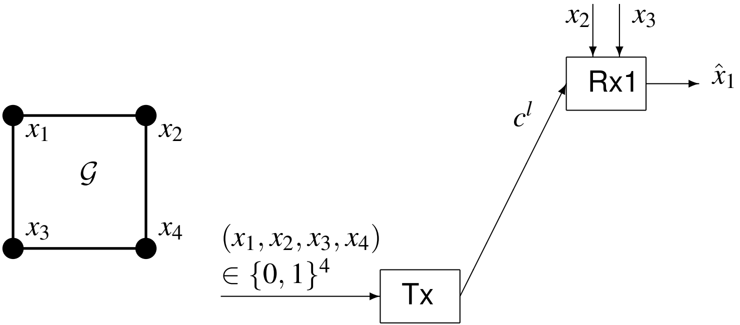
Index coding with side information



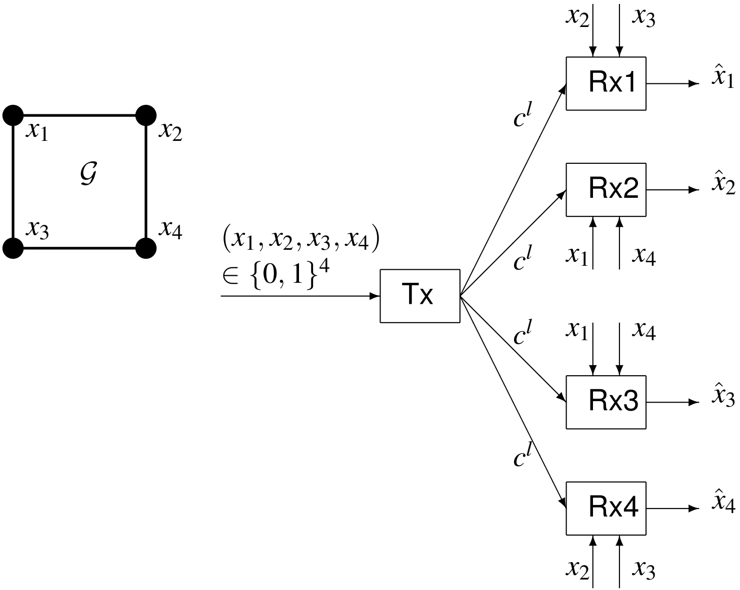
Index coding with side information



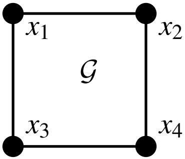
Index coding with side information

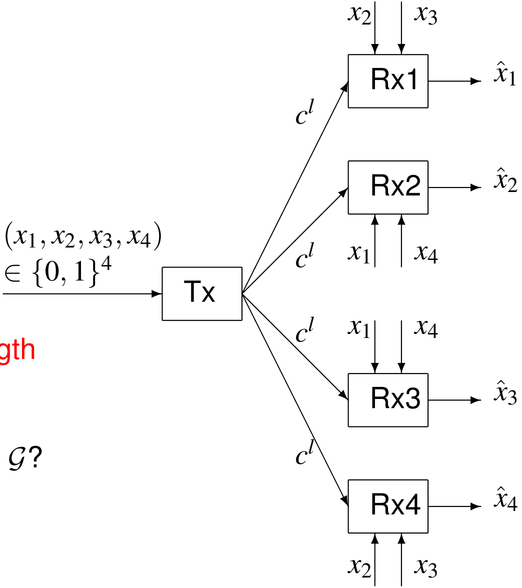


Index coding with side information



Index coding with side information





What's the minimum length l required for reliable reconstruction of (x_1, x_2, x_3, x_4) given graph $\mathcal{G}$?

Index coding with side information

We say that a matrix $\mathbf{A} \in \mathbb{F}_2^{k \times k}$ fits a graph $\mathcal{G}$ if

$$a_{ii} = 1, \qquad a_{ij} = 0 \quad \text{if} \quad (i,j) \notin \mathcal{G}$$

Index coding with side information

We say that a matrix $\mathbf{A} \in \mathbb{F}_2^{k\times k}$ fits a graph $\mathcal{G}$ if

$$a_{ii} = 1, \qquad a_{ij} = 0 \quad \text{if} \quad (i,j) \notin \mathcal{G}$$

Definition

The minimum rank of a graph $\mathcal{G}$ is defined as

$$\mathrm{minrk}_2(\mathcal{G}) := \min\{\mathrm{rank}(\mathbf{A}) : \mathbf{A} \text{ fits } \mathcal{G}\}$$

This is an example of a rank minimization problem over $\mathbb{F}_2$

Index coding with side information

We say that a matrix $\mathbf{A} \in \mathbb{F}_2^{k\times k}$ fits a graph $\mathcal{G}$ if

$$a_{ii} = 1, \qquad a_{ij} = 0 \quad \text{if} \quad (i,j) \notin \mathcal{G}$$

Definition

The minimum rank of a graph $\mathcal{G}$ is defined as

$$\mathrm{minrk}_2(\mathcal{G}) := \min\{\mathrm{rank}(\mathbf{A}) : \mathbf{A} \text{ fits } \mathcal{G}\}$$

This is an example of a rank minimization problem over $\mathbb{F}_2$

Theorem (Bar-Yossef et al. 06)

The optimal length of a linear index code for graph $\mathcal{G}$ equals $\mathrm{minrk}_2(\mathcal{G})$.

Matrix completion in the reals

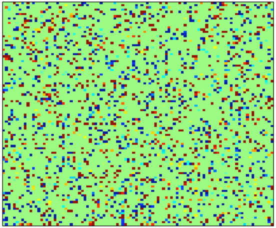
- Very popular and well-studied problem

Matrix completion in the reals

- Very popular and well-studied problem
- Given a subset of entries from a low-rank matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$

Matrix completion in the reals

- Very popular and well-studied problem
- Given a subset of entries from a low-rank matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$

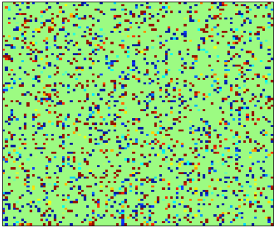


Green = Missing entry

- Goal: Recover the matrix $\mathbf{X}$

Matrix completion in the reals

- Very popular and well-studied problem
- Given a subset of entries from a low-rank matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$



Green = Missing entry

- Goal: Recover the matrix $\mathbf{X}$
- The nuclear norm heuristic has enjoyed tremendous successes

Rank minimization in the reals

- $\mathbf{X} \in \mathbb{R}^{m \times n}$ a low-rank matrix

Rank minimization in the reals

- $\mathbf{X} \in \mathbb{R}^{m \times n}$ a low-rank matrix
- Given k linear measurements:

$$\begin{aligned} y_1 &= \langle \mathbf{H}_1, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_1 \mathbf{X}^T) \\ &\vdots \\ y_k &= \langle \mathbf{H}_k, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_k \mathbf{X}^T) \end{aligned}$$

Rank minimization in the reals

- $\mathbf{X} \in \mathbb{R}^{m \times n}$ a low-rank matrix
- Given k linear measurements:

$$
\begin{aligned}
y_1 &= \langle \mathbf{H}_1, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_1 \mathbf{X}^T) \\
&\vdots \\
y_k &= \langle \mathbf{H}_k, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_k \mathbf{X}^T)
\end{aligned}
$$

- A generalization of matrix completion

Rank minimization in the reals

- $\mathbf{X} \in \mathbb{R}^{m \times n}$ a low-rank matrix
- Given k linear measurements:

$$\begin{aligned} y_1 &= \langle \mathbf{H}_1, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_1 \mathbf{X}^T) \\ &\vdots \\ y_k &= \langle \mathbf{H}_k, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_k \mathbf{X}^T) \end{aligned}$$

- A generalization of matrix completion
- Goal: Recover the matrix $\mathbf{X}$

Rank minimization in the reals

- $\mathbf{X} \in \mathbb{R}^{m \times n}$ a low-rank matrix
- Given k linear measurements:

$$\begin{aligned} y_1 &= \langle \mathbf{H}_1, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_1 \mathbf{X}^T) \\ &\vdots \\ y_k &= \langle \mathbf{H}_k, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_k \mathbf{X}^T) \end{aligned}$$

- A generalization of matrix completion
- Goal: Recover the matrix $\mathbf{X}$
- The nuclear norm heuristic has also enjoyed tremendous successes

Rank minimization in the reals

- $\mathbf{X} \in \mathbb{R}^{m \times n}$ a low-rank matrix
- Given k linear measurements:

$$\begin{aligned} y_1 &= \langle \mathbf{H}_1, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_1 \mathbf{X}^T) \\ &\vdots \\ y_k &= \langle \mathbf{H}_k, \mathbf{X} \rangle = \mathrm{Trace}(\mathbf{H}_k \mathbf{X}^T) \end{aligned}$$

- A generalization of matrix completion
- Goal: Recover the matrix $\mathbf{X}$
- The nuclear norm heuristic has also enjoyed tremendous successes
- Wide applicability in for e.g., collaborative filtering, minimal realization of LTI system etc.

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse
4. Achievability for Uniform Model
5. Achievability for Sparse Model
6. Coding-Theoretic Interpretations

Problem Setup: Rank minimization over finite fields

- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k$ are random.

Problem Setup: Rank minimization over finite fields

- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k$ are random.
- Assume for simplicity that $\mathbf{X}$ is square ($n \times n$)

Problem Setup: Rank minimization over finite fields

- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k$ are random.
- Assume for simplicity that $\mathbf{X}$ is square ($n \times n$)
- $\mathbf{X}$ has rank $\leq r \leq n$

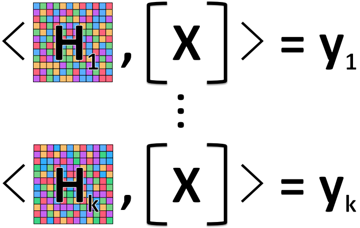
Problem Setup: Rank minimization over finite fields

- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k$ are random.
- Assume for simplicity that $\mathbf{X}$ is square ($n \times n$)
- $\mathbf{X}$ has rank $\leq r \leq n$

$$\langle \mathbf{H}_1, \mathbf{X} \rangle = y_1$$
$$\vdots$$
$$\langle \mathbf{H}_k, \mathbf{X} \rangle = y_k$$

Problem Setup: Rank minimization over finite fields

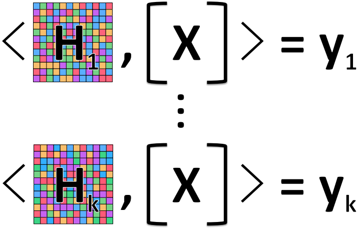
- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k$ are random.
- Assume for simplicity that $\mathbf{X}$ is square ($n \times n$)
- $\mathbf{X}$ has rank $\leq r \leq n$



- Arithmetic is performed in the field $\mathbb{F}_q$

Problem Setup: Rank minimization over finite fields

- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k$ are random.
- Assume for simplicity that $\mathbf{X}$ is square ($n \times n$)
- $\mathbf{X}$ has rank $\leq r \leq n$



- Arithmetic is performed in the field $\mathbb{F}_q$
- Each measurement $y_a, a = 1, \ldots, k$ belongs to $\mathbb{F}_q$

Problem Setup: Rank minimization over finite fields

- Problem Statement: Given $(y^k, \mathbf{H}^k)$ and the measurement model, estimate $\mathbf{X}$.

Problem Setup: Rank minimization over finite fields

- Problem Statement: Given $(y^k, \mathbf{H}^k)$ and the measurement model, estimate $\mathbf{X}$.
- Find necessary and sufficient conditions on k and sensing model such that recovery is reliable, i.e.,

$$\mathbb{P}(\mathcal{E}_n) \to 0$$

Main Results

- Converse without assuming linear model
  - Lower bound on number of measurements k

Main Results

- Converse without assuming linear model
  - Lower bound on number of measurements k
- Achievability under uniform sampling model
  - Sufficient condition that matches lower bound
  - Reliability function for the particular decoder

Main Results

- Converse without assuming linear model
  - Lower bound on number of measurements k
- Achievability under uniform sampling model
  - Sufficient condition that matches lower bound
  - Reliability function for the particular decoder
- Achievability under sparse sampling model
  - Sufficient condition that matches lower bound

Main Results

- Converse without assuming linear model
  - Lower bound on number of measurements k
- Achievability under uniform sampling model
  - Sufficient condition that matches lower bound
  - Reliability function for the particular decoder
- Achievability under sparse sampling model
  - Sufficient condition that matches lower bound
- Coding-theoretic interpretations: Geometric insights

Main Results

- k: Num. of linear measurements
- n: Dim. of matrix $\mathbf{X}$
- r: Max. rank of matrix $\mathbf{X}$
- $\gamma = \frac{r}{n}$: Rank-dimension ratio

Main Results

- k: Num. of linear measurements
- n: Dim. of matrix $\mathbf{X}$
- r: Max. rank of matrix $\mathbf{X}$
- $\gamma = \frac{r}{n}$: Rank-dimension ratio

$$2\gamma\left(1 - \gamma/2\right)n^2 = 2rn - r^2$$

Main Results

- k: Num. of linear measurements
- n: Dim. of matrix $\mathbf{X}$
- r: Max. rank of matrix $\mathbf{X}$
- $\gamma = \frac{r}{n}$: Rank-dimension ratio

$$2\gamma\left(1-\gamma/2\right)n^2 = 2rn - r^2$$

Result	Statement	Consequence

Main Results

- k: Num. of linear measurements
- n: Dim. of matrix $\mathbf{X}$
- r: Max. rank of matrix $\mathbf{X}$
- $\gamma = \frac{r}{n}$: Rank-dimension ratio

$$2\gamma\left(1-\gamma/2\right)n^2 = 2rn - r^2$$

Result	Statement	Consequence
Converse	$k < (2-\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \nrightarrow 0$

Main Results

- k: Num. of linear measurements
- n: Dim. of matrix $\mathbf{X}$
- r: Max. rank of matrix $\mathbf{X}$
- $\gamma = \frac{r}{n}$: Rank-dimension ratio

$$2\gamma\left(1-\gamma/2\right)n^2 = 2rn - r^2$$

Result	Statement	Consequence
Converse	$k < (2-\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \nrightarrow 0$
Achievability (Uniform)	$k > (2+\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \to 0$ $\mathbb{P}(\mathcal{E}_n) \approx q^{-n^2 E(R)}$

Main Results

- k: Num. of linear measurements
- n: Dim. of matrix $\mathbf{X}$
- r: Max. rank of matrix $\mathbf{X}$
- $\gamma = \frac{r}{n}$: Rank-dimension ratio

$$2\gamma\left(1 - \gamma/2\right) n^2$$
$$= 2rn - r^2$$

Result	Statement	Consequence
Converse	$k < (2-\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \not\to 0$
Achievability (Uniform)	$k > (2+\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \to 0$ $\mathbb{P}(\mathcal{E}_n) \approx q^{-n^2 E(R)}$
Achievability (Sparse)	$k > (2+\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \to 0$

Main Results

- k: Num. of linear measurements
- n: Dim. of matrix $\mathbf{X}$
- r: Max. rank of matrix $\mathbf{X}$
- $\gamma = \frac{r}{n}$: Rank-dimension ratio

$$2\gamma\left(1-\gamma/2\right)n^2 = 2rn - r^2$$

Result	Statement	Consequence
Converse	$k < (2-\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \nrightarrow 0$
Achievability (Uniform)	$k > (2+\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \to 0$ $\mathbb{P}(\mathcal{E}_n) \approx q^{-n^2 E(R)}$
Achievability (Sparse)	$k > (2+\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \to 0$
Achievability (Noisy)	$k \gtrapprox (3+\varepsilon)(\gamma+\sigma)n^2$ (q assumed large)	$\mathbb{P}(\mathcal{E}_n) \to 0$

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse
4. Achievability for Uniform Model
5. Achievability for Sparse Model
6. Coding-Theoretic Interpretations

A necessary condition on number of measurements

Given k measurements $y_a \in \mathbb{F}_q$ and sensing matrices $\mathbf{H}_a \in \mathbb{F}_q^{n \times n}$, we want a necessary condition for reliable recovery of $\mathbf{X}$.

A necessary condition on number of measurements

Given k measurements $y_a \in \mathbb{F}_q$ and sensing matrices $\mathbf{H}_a \in \mathbb{F}_q^{n\times n}$, we want a necessary condition for reliable recovery of $\mathbf{X}$.

Proposition (Converse)

Assume

- $\mathbf{X}$ *drawn uniformly at random from all matrices in* $\mathbb{F}_q^{n\times n}$ *of rank* $\leq r$

A necessary condition on number of measurements

Given k measurements $y_a \in \mathbb{F}_q$ and sensing matrices $\mathbf{H}_a \in \mathbb{F}_q^{n\times n}$, we want a necessary condition for reliable recovery of $\mathbf{X}$.

Proposition (Converse)

Assume

- $\mathbf{X}$ *drawn uniformly at random from all matrices in* $\mathbb{F}_q^{n\times n}$ *of rank* $\leq r$
- *Sensing matrices* $\mathbf{H}_a$ *independent of* $\mathbf{X}$

A necessary condition on number of measurements

Given k measurements $y_a \in \mathbb{F}_q$ and sensing matrices $\mathbf{H}_a \in \mathbb{F}_q^{n\times n}$, we want a necessary condition for reliable recovery of $\mathbf{X}$.

Proposition (Converse)

Assume

- $\mathbf{X}$ *drawn uniformly at random from all matrices in* $\mathbb{F}_q^{n\times n}$ *of rank* $\leq r$
- *Sensing matrices* $\mathbf{H}_a$ *independent of* $\mathbf{X}$
- $r/n \to \gamma$ *(constant)*

A necessary condition on number of measurements

Given k measurements $y_a \in \mathbb{F}_q$ and sensing matrices $\mathbf{H}_a \in \mathbb{F}_q^{n\times n}$, we want a necessary condition for reliable recovery of $\mathbf{X}$.

Proposition (Converse)

Assume

- $\mathbf{X}$ *drawn uniformly at random from all matrices in* $\mathbb{F}_q^{n\times n}$ *of rank* $\leq r$
- *Sensing matrices* $\mathbf{H}_a$ *independent of* $\mathbf{X}$
- $r/n \to \gamma$ *(constant)*

If the number of measurements satisfies

$$k < (2-\varepsilon)\gamma\left(1-\frac{\gamma}{2}\right)n^2$$

then $\mathbb{P}(\hat{\mathbf{X}} \neq \mathbf{X}) \geq \varepsilon/2$ *for all* n *sufficiently large.*

A necessary condition on number of measurements

Ramifications:

- We need at least

$$k \geq 2\gamma \left(1 - \frac{\gamma}{2}\right) n^2$$

  measurements for successful recovery

A necessary condition on number of measurements

Ramifications:

- We need at least
$$k \geq 2\gamma \left(1 - \frac{\gamma}{2}\right) n^2$$
measurements for successful recovery
- Need as many measurements as there are degrees of freedom

A necessary condition on number of measurements

Ramifications:

- We need at least

$$k \geq 2\gamma \left(1 - \frac{\gamma}{2}\right) n^2$$

  measurements for successful recovery

- Need as many measurements as there are degrees of freedom

- Fano's inequality and counting

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse
4. Achievability for Uniform Model
5. Achievability for Sparse Model
6. Coding-Theoretic Interpretations

A sensing model: Uniform model

- Now we assume that $\mathbf{X}$ is non-random

A sensing model: Uniform model

- Now we assume that $\mathbf{X}$ is non-random
- $\mathrm{rank}(\mathbf{X}) \leq r = \gamma n$

A sensing model: Uniform model

- Now we assume that $\mathbf{X}$ is non-random
- $\mathrm{rank}(\mathbf{X}) \leq r = \gamma n$

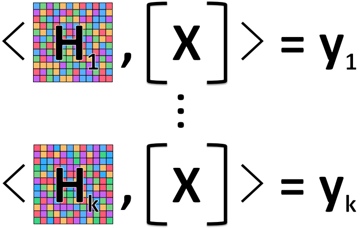
$$\left\langle \mathbf{H}_1 , \left[\mathbf{X} \right] \right\rangle = \mathbf{y}_1$$

$$\vdots$$

$$\left\langle \mathbf{H}_k , \left[\mathbf{X} \right] \right\rangle = \mathbf{y}_k$$

A sensing model: Uniform model

- Now we assume that $\mathbf{X}$ is non-random
- $\text{rank}(\mathbf{X}) \leq r = \gamma n$



- Each entry of each sensing matrix $\mathbf{H}_a$ is i.i.d. and has a uniform distribution in $\mathbb{F}_q$:

$$\mathbb{P}([\mathbf{H}_a]_{i,j} = h) = \frac{1}{q}, \qquad \forall\, h \in \mathbb{F}_q$$

The min-rank decoder

- We employ the min-rank decoder

$$\begin{aligned} \text{minimize} \quad & \operatorname{rank}(\tilde{\mathbf{X}}) \\ \text{subject to} \quad & \langle \mathbf{H}_a, \tilde{\mathbf{X}} \rangle = y_a, \quad a = 1, \ldots, k \end{aligned}$$

The min-rank decoder

- We employ the min-rank decoder

$$\begin{aligned} \text{minimize} \quad & \mathrm{rank}(\tilde{\mathbf{X}}) \\ \text{subject to} \quad & \langle \mathbf{H}_a, \tilde{\mathbf{X}} \rangle = y_a, \quad a = 1, \ldots, k \end{aligned}$$

- NP-hard, combinatorial

The min-rank decoder

- We employ the min-rank decoder

$$\begin{aligned} \text{minimize} \quad & \operatorname{rank}(\tilde{\mathbf{X}}) \\ \text{subject to} \quad & \langle \mathbf{H}_a, \tilde{\mathbf{X}} \rangle = y_a, \quad a = 1, \ldots, k \end{aligned}$$

- NP-hard, combinatorial
- Denote the set of optimizers as $\mathcal{S}$

The min-rank decoder

- We employ the min-rank decoder

$$\begin{aligned} \text{minimize} \quad & \operatorname{rank}(\tilde{\mathbf{X}}) \\ \text{subject to} \quad & \langle \mathbf{H}_a, \tilde{\mathbf{X}} \rangle = y_a, \quad a = 1, \ldots, k \end{aligned}$$

- NP-hard, combinatorial
- Denote the set of optimizers as $\mathcal{S}$
- Define the error event:

$$\mathcal{E}_n := \{|\mathcal{S}| > 1\} \cup (\{|\mathcal{S}| = 1\} \cap \{\mathbf{X}^* \neq \mathbf{X}\})$$

The min-rank decoder

- We employ the min-rank decoder

$$\begin{array}{ll} \text{minimize} & \operatorname{rank}(\tilde{\mathbf{X}}) \\ \text{subject to} & \langle \mathbf{H}_a, \tilde{\mathbf{X}} \rangle = y_a, \quad a = 1, \ldots, k \end{array}$$

- NP-hard, combinatorial
- Denote the set of optimizers as $\mathcal{S}$
- Define the error event:

$$\mathcal{E}_n := \{|\mathcal{S}| > 1\} \cup (\{|\mathcal{S}| = 1\} \cap \{\mathbf{X}^* \neq \mathbf{X}\})$$

- We want the solution to be unique and correct

Achievability under uniform model

Proposition (Achievability under uniform model)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn uniformly*

Achievability under uniform model

Proposition (Achievability under uniform model)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn uniformly*
- *Min-rank decoder is used*

Achievability under uniform model

Proposition (Achievability under uniform model)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn uniformly*
- *Min-rank decoder is used*
- $r/n \to \gamma$ *(constant)*

Achievability under uniform model

Proposition (Achievability under uniform model)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn* *uniformly*
- *Min-rank decoder is used*
- $r/n \to \gamma$ *(constant)*

If the number of measurements satisfies

$$k > (2+\varepsilon)\gamma\left(1-\frac{\gamma}{2}\right)n^2$$

then $\mathbb{P}(\mathcal{E}_n) \to 0$.

Proof Sketch

By the union bound, the error probability can be bounded as

$$\mathbb{P}(\mathcal{E}_n) \le \sum_{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X})} \mathbb{P}(\langle \mathbf{Z}, \mathbf{H}_a \rangle = \langle \mathbf{X}, \mathbf{H}_a \rangle, \forall a = 1, \ldots, k)$$

Proof Sketch

By the union bound, the error probability can be bounded as

$$\mathbb{P}(\mathcal{E}_n) \le \sum_{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X})} \mathbb{P}(\langle\mathbf{Z},\mathbf{H}_a\rangle = \langle\mathbf{X},\mathbf{H}_a\rangle, \forall a = 1,\ldots,k)$$

At the same time, by uniformity and independence,

$$\mathbb{P}(\langle\mathbf{Z},\mathbf{H}_a\rangle = \langle\mathbf{X},\mathbf{H}_a\rangle, \forall a = 1,\ldots,k) = q^{-k}$$

Proof Sketch

By the union bound, the error probability can be bounded as

$$\mathbb{P}(\mathcal{E}_n) \leq \sum_{\mathbf{Z} \neq \mathbf{X} : \mathrm{rank}(\mathbf{Z}) \leq \mathrm{rank}(\mathbf{X})} \mathbb{P}(\langle \mathbf{Z}, \mathbf{H}_a \rangle = \langle \mathbf{X}, \mathbf{H}_a \rangle, \forall a = 1, \ldots, k)$$

At the same time, by uniformity and independence,

$$\mathbb{P}(\langle \mathbf{Z}, \mathbf{H}_a \rangle = \langle \mathbf{X}, \mathbf{H}_a \rangle, \forall a = 1, \ldots, k) = q^{-k}$$

That the number of matrices of rank $\leq r$ is bounded above as

$$4q^{2\gamma(1-\gamma/2)n^2}$$

Remarks for the sufficient condition

- Recall the lower bound:

$$k \geq 2\gamma \left(1 - \frac{\gamma}{2}\right) n^2$$

Remarks for the sufficient condition

- Recall the lower bound:

$$k \geq 2\gamma \left(1 - \frac{\gamma}{2}\right) n^2$$

- Thus, min-rank decoder matches the lower bound

Remarks for the sufficient condition

- Recall the lower bound:

$$k \geq 2\gamma \left(1 - \frac{\gamma}{2}\right) n^2$$

- Thus, min-rank decoder matches the lower bound
- Optimal measurement complexity

Remarks for the sufficient condition

- Recall the lower bound:

$$k \geq 2\gamma \left(1 - \frac{\gamma}{2}\right) n^2$$

- Thus, min-rank decoder matches the lower bound
- Optimal measurement complexity
- Can we provide a precise rate of convergence of $\mathbb{P}(\mathcal{E}_n)$?

Remarks for the sufficient condition

- Recall the lower bound:

$$k \geq 2\gamma\left(1 - \frac{\gamma}{2}\right) n^2$$

- Thus, min-rank decoder matches the lower bound
- Optimal measurement complexity
- Can we provide a precise rate of convergence of $\mathbb{P}(\mathcal{E}_n)$?
- Reliability Function?

The Reliability Function

Definition

The rate of a sequence of linear measurement models is defined as

$$R := \lim_{n\to\infty} \frac{n^2 - k}{n^2} = \lim_{n\to\infty} 1 - \frac{k}{n^2}, \qquad k = \#\ \textit{measurements}$$

The Reliability Function

Definition

The rate of a sequence of linear measurement models is defined as

$$R := \lim_{n\to\infty} \frac{n^2 - k}{n^2} = \lim_{n\to\infty} 1 - \frac{k}{n^2}, \qquad k = \#\ \textit{measurements}$$

Analogy to coding: The rate of the code

$$\mathcal{C} = \{\mathbf{C} : \langle \mathbf{C}, \mathbf{H}_a \rangle = 0, a = 1, \ldots, k\}$$

is lower bounded by $1 - k/n^2$.

The Reliability Function

Definition

The reliability function of the min-rank decoder is defined as

$$E(R) := \lim_{n\to\infty} -\frac{1}{n^2}\log_q \mathbb{P}(\mathcal{E}_n)$$

The Reliability Function

Definition

The reliability function of the min-rank decoder is defined as

$$E(R) := \lim_{n\to\infty} -\frac{1}{n^2} \log_q \mathbb{P}(\mathcal{E}_n)$$

Roughly speaking,

$$\mathbb{P}(\mathcal{E}_n) \approx q^{-n^2 E(R)}$$

The Reliability Function

Definition

The reliability function of the min-rank decoder is defined as

$$E(R) := \lim_{n\to\infty} -\frac{1}{n^2}\log_q \mathbb{P}(\mathcal{E}_n)$$

Roughly speaking,

$$\mathbb{P}(\mathcal{E}_n) \approx q^{-n^2 E(R)}$$

Reliability function also known as error exponent

The Reliability Function

- Our achievability proof gives a lower bound on $E(R)$

The Reliability Function

- Our achievability proof gives a lower bound on $E(R)$
- Upper bound is more tricky. But...

The Reliability Function

- Our achievability proof gives a lower bound on $E(R)$
- Upper bound is more tricky. But...

Proposition (Reliability Function)

Assume

- *Sensing matrices $\mathbf{H}_a$ drawn uniformly*

The Reliability Function

- Our achievability proof gives a lower bound on $E(R)$
- Upper bound is more tricky. But...

Proposition (Reliability Function)

Assume

- *Sensing matrices $\mathbf{H}_a$ drawn uniformly*
- *Min-rank decoder is used*

The Reliability Function

- Our achievability proof gives a lower bound on $E(R)$
- Upper bound is more tricky. But...

Proposition (Reliability Function)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn uniformly*
- *Min-rank decoder is used*
- $r/n \to \gamma$ *(constant)*

The Reliability Function

- Our achievability proof gives a lower bound on $E(R)$
- Upper bound is more tricky. But...

Proposition (Reliability Function)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn uniformly*
- *Min-rank decoder is used*
- $r/n \to \gamma$ *(constant)*

Then,

$$E(R) = \left|(1-R) - 2\gamma\left(1 - \frac{\gamma}{2}\right)\right|^{+}$$

Note $|x|^{+} := \max\{x, 0\}$.

Interpretation

$$E(R) \approx \left| \frac{k}{n^2} - 2\gamma \left(1 - \frac{\gamma}{2}\right) \right|^+$$

Interpretation

$$E(R) \approx \left| \frac{k}{n^2} - 2\gamma \left(1 - \frac{\gamma}{2}\right) \right|^+$$

- The more the ratio $\frac{k}{n^2}$ exceeds $2\gamma\left(1 - \frac{\gamma}{2}\right)$

Interpretation

$$E(R) \approx \left| \frac{k}{n^2} - 2\gamma \left(1 - \frac{\gamma}{2}\right) \right|^{+}$$

- The more the ratio $\frac{k}{n^2}$ exceeds $2\gamma\left(1 - \frac{\gamma}{2}\right)$
- The larger $E(R)$

Interpretation

$$E(R) \approx \left| \frac{k}{n^2} - 2\gamma\left(1 - \frac{\gamma}{2}\right) \right|^+$$

- The more the ratio $\frac{k}{n^2}$ exceeds $2\gamma\left(1 - \frac{\gamma}{2}\right)$
- The larger $E(R)$
- The faster $\mathbb{P}(\mathcal{E}_n)$ decays

Interpretation

$$E(R) \approx \left| \frac{k}{n^2} - 2\gamma \left(1 - \frac{\gamma}{2}\right) \right|^+$$

- The more the ratio $\frac{k}{n^2}$ exceeds $2\gamma\left(1 - \frac{\gamma}{2}\right)$
- The larger $E(R)$
- The faster $\mathbb{P}(\mathcal{E}_n)$ decays
- Linear relationship

Remarks on the reliability function

- $E(R) \approx \left| \frac{k}{n^2} - 2\gamma \left(1 - \frac{\gamma}{2}\right) \right|^+$
- Proof of upper bound on $E(R)$ is interesting

Remarks on the reliability function

- $E(R) \approx \left| \frac{k}{n^2} - 2\gamma \left(1 - \frac{\gamma}{2}\right) \right|^+$
- Proof of upper bound on $E(R)$ is interesting
- It utilizes de Caen's lower bound: Let $\mathcal{B}_1, \ldots, \mathcal{B}_M$ be events:

$$\mathbb{P}\left(\bigcup_{m=1}^{M} \mathcal{B}_m\right) \geq \sum_{m=1}^{M} \frac{\mathbb{P}(\mathcal{B}_m)^2}{\sum_{m'=1}^{M} \mathbb{P}(\mathcal{B}_m \cap \mathcal{B}_{m'})}.$$

Remarks on the reliability function

- $E(R) \approx \left| \frac{k}{n^2} - 2\gamma \left(1 - \frac{\gamma}{2}\right) \right|^+$
- Proof of upper bound on $E(R)$ is interesting
- It utilizes de Caen's lower bound: Let $\mathcal{B}_1, \ldots, \mathcal{B}_M$ be events:

$$\mathbb{P}\left(\bigcup_{m=1}^{M} \mathcal{B}_m \right) \ge \sum_{m=1}^{M} \frac{\mathbb{P}(\mathcal{B}_m)^2}{\sum_{m'=1}^{M} \mathbb{P}(\mathcal{B}_m \cap \mathcal{B}_{m'})}.$$

- Pairwise independence of error events

Remarks on the reliability function

- $E(R) \approx \left| \frac{k}{n^2} - 2\gamma\left(1 - \frac{\gamma}{2}\right)\right|^+$
- Proof of upper bound on $E(R)$ is interesting
- It utilizes de Caen's lower bound: Let $\mathcal{B}_1, \ldots, \mathcal{B}_M$ be events:

$$\mathbb{P}\left(\bigcup_{m=1}^{M} \mathcal{B}_m\right) \geq \sum_{m=1}^{M} \frac{\mathbb{P}(\mathcal{B}_m)^2}{\sum_{m'=1}^{M} \mathbb{P}(\mathcal{B}_m \cap \mathcal{B}_{m'})}.$$

- Pairwise independence of error events
- Analogy: Linear codes achieve capacity for symmetric DMCs

Remarks on the reliability function

- $E(R) \approx \left| \frac{k}{n^2} - 2\gamma\left(1 - \frac{\gamma}{2}\right)\right|^+$
- Proof of upper bound on $E(R)$ is interesting
- It utilizes de Caen's lower bound: Let $\mathcal{B}_1, \ldots, \mathcal{B}_M$ be events:

$$\mathbb{P}\left(\bigcup_{m=1}^{M} \mathcal{B}_m\right) \ge \sum_{m=1}^{M} \frac{\mathbb{P}(\mathcal{B}_m)^2}{\sum_{m'=1}^{M} \mathbb{P}(\mathcal{B}_m \cap \mathcal{B}_{m'})}.$$

- Pairwise independence of error events
- Analogy: Linear codes achieve capacity for symmetric DMCs
- de Caen's inequality allows us to exploit pairwise independence to make statements about error exponents

Remarks on the reliability function

- $E(R) \approx \left| \frac{k}{n^2} - 2\gamma\left(1 - \frac{\gamma}{2}\right)\right|^{+}$
- Proof of upper bound on $E(R)$ is interesting
- It utilizes de Caen's lower bound: Let $\mathcal{B}_1, \ldots, \mathcal{B}_M$ be events:

$$\mathbb{P}\left(\bigcup_{m=1}^{M} \mathcal{B}_m\right) \geq \sum_{m=1}^{M} \frac{\mathbb{P}(\mathcal{B}_m)^2}{\sum_{m'=1}^{M} \mathbb{P}(\mathcal{B}_m \cap \mathcal{B}_{m'})}.$$

- Pairwise independence of error events
- Analogy: Linear codes achieve capacity for symmetric DMCs
- de Caen's inequality allows us to exploit pairwise independence to make statements about error exponents
- Union bound in achievability is tight!

Extensions

- All the preceding analyses can be extended to the noisy case:

$$y_a = \langle \mathbf{H}_a, \mathbf{X} \rangle + w_a$$

Extensions

- All the preceding analyses can be extended to the noisy case:

$$y_a = \langle \mathbf{H}_a, \mathbf{X} \rangle + w_a$$

- Noise w_a can be random or deterministic but assume $\|\mathbf{w}\|_0 = \sigma n^2$

Extensions

- All the preceding analyses can be extended to the noisy case:

$$y_a = \langle \mathbf{H}_a, \mathbf{X} \rangle + w_a$$

- Noise w_a can be random or deterministic but assume $\|\mathbf{w}\|_0 = \sigma n^2$

$$\begin{aligned} \text{minimize} \quad & \operatorname{rank}(\tilde{\mathbf{X}}) + \lambda \|\tilde{\mathbf{w}}\|_0 \\ \text{subject to} \quad & \langle \mathbf{H}_a, \tilde{\mathbf{X}} \rangle + \tilde{w}_a = y_a, \quad a = 1, \ldots, k \end{aligned}$$

Extensions

- All the preceding analyses can be extended to the noisy case:

$$y_a = \langle \mathbf{H}_a, \mathbf{X} \rangle + w_a$$

- Noise w_a can be random or deterministic but assume $\|\mathbf{w}\|_0 = \sigma n^2$

$$\begin{aligned} \text{minimize} \quad & \mathrm{rank}(\tilde{\mathbf{X}}) + \lambda \|\tilde{\mathbf{w}}\|_0 \\ \text{subject to} \quad & \langle \mathbf{H}_a, \tilde{\mathbf{X}} \rangle + \tilde{w}_a = y_a, \quad a = 1, \ldots, k \end{aligned}$$

- Choose $\lambda = 1/n$

Extensions

- All the preceding analyses can be extended to the noisy case:

$$y_a = \langle \mathbf{H}_a, \mathbf{X} \rangle + w_a$$

- Noise w_a can be random or deterministic but assume $\|\mathbf{w}\|_0 = \sigma n^2$

$$\begin{aligned} \text{minimize} \quad & \text{rank}(\tilde{\mathbf{X}}) + \lambda \|\tilde{\mathbf{w}}\|_0 \\ \text{subject to} \quad & \langle \mathbf{H}_a, \tilde{\mathbf{X}} \rangle + \tilde{w}_a = y_a, \quad a = 1, \ldots, k \end{aligned}$$

- Choose $\lambda = 1/n$
- If

$$k \gtrsim (3 + \varepsilon)(\gamma + \sigma) n^2,$$

$\mathbb{P}(\mathcal{E}_n) \to 0$. See preprint

Outline

1. Motivation
2. Problem Setup and Summary of Main Results
3. Converse
4. Achievability for Uniform Model
5. Achievability for Sparse Model
6. Coding-Theoretic Interpretations

Sparse sensing model

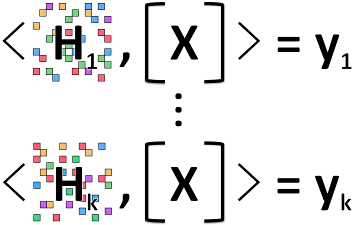
- Assume as usual that $\mathbf{X}$ is non-random

Sparse sensing model

- Assume as usual that $\mathbf{X}$ is non-random
- $\mathrm{rank}(\mathbf{X}) \leq r = \gamma n$

Sparse sensing model

- Assume as usual that $\mathbf{X}$ is non-random
- $\mathrm{rank}(\mathbf{X}) \leq r = \gamma n$
- Sensing matrices are sparse



Sparse sensing model

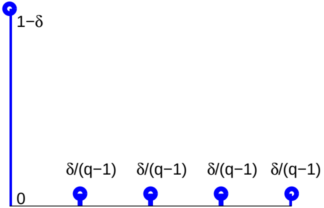
- Assume as usual that $\mathbf{X}$ is non-random
- $\mathrm{rank}(\mathbf{X}) \le r = \gamma n$
- Sensing matrices are sparse

$$\langle \mathbf{H}_1, [\,\mathbf{X}\,] \rangle = \mathbf{y}_1$$
$$\vdots$$
$$\langle \mathbf{H}_k, [\,\mathbf{X}\,] \rangle = \mathbf{y}_k$$

- Arithmetic still performed in $\mathbb{F}_q$

Sparse sensing model

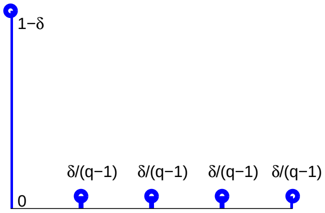
- Each entry of each sensing matrix $\mathbf{H}_a$ is i.i.d. and has a δ-sparse distribution in $\mathbb{F}_q$:



$$\mathbb{P}([\mathbf{H}_a]_{ij} = h) = \begin{cases} 1-\delta & h = 0 \\ \frac{\delta}{q-1} & h \neq 0 \end{cases}$$

Sparse sensing model

- Each entry of each sensing matrix $\mathbf{H}_a$ is i.i.d. and has a δ-sparse distribution in $\mathbb{F}_q$:

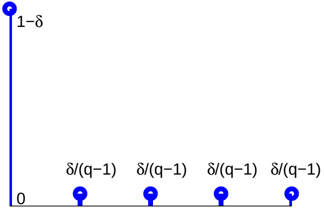


$$\mathbb{P}([\mathbf{H}_a]_{ij} = h) = \begin{cases} 1-\delta & h = 0 \\ \frac{\delta}{q-1} & h \neq 0 \end{cases}$$

- Fewer adds and multiplies since $\mathbf{H}_a$ sparse $\Rightarrow$ Encoding cheaper

Sparse sensing model

- Each entry of each sensing matrix $\mathbf{H}_a$ is i.i.d. and has a δ-sparse distribution in $\mathbb{F}_q$:

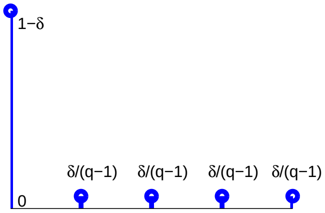


$$\mathbb{P}([\mathbf{H}_a]_{ij} = h) = \begin{cases} 1-\delta & h = 0 \\ \frac{\delta}{q-1} & h \neq 0 \end{cases}$$

- Fewer adds and multiplies since $\mathbf{H}_a$ sparse $\Rightarrow$ Encoding cheaper
- May help in decoding via message-passing algorithms

Sparse sensing model

- Each entry of each sensing matrix $\mathbf{H}_a$ is i.i.d. and has a δ-sparse distribution in $\mathbb{F}_q$:



$$\mathbb{P}([\mathbf{H}_a]_{ij} = h) = \begin{cases} 1-\delta & h = 0 \\ \frac{\delta}{q-1} & h \neq 0 \end{cases}$$

- Fewer adds and multiplies since $\mathbf{H}_a$ sparse $\Rightarrow$ Encoding cheaper
- May help in decoding via message-passing algorithms
- How fast can δ, the sparsity factor, decay with n for reliable recovery?

Sparse sensing model

- Problem becomes more challenging

Sparse sensing model

- Problem becomes more challenging
- $\mathbf{X}$ is not sensed "as much"

Sparse sensing model

- Problem becomes more challenging
- $\mathbf{X}$ is not sensed "as much"
- Measurements y^k do not contain as much information about $\mathbf{X}$

Sparse sensing model

- Problem becomes more challenging
- $\mathbf{X}$ is not sensed "as much"
- Measurements y^k do not contain as much information about $\mathbf{X}$
- The equality

$$\mathbb{P}(\langle \mathbf{Z}, \mathbf{H}_a \rangle = \langle \mathbf{X}, \mathbf{H}_a \rangle, \forall a = 1, \ldots, k) = q^{-k}$$

  no longer holds

Sparse sensing model

- Problem becomes more challenging
- $\mathbf{X}$ is not sensed "as much"
- Measurements y^k do not contain as much information about $\mathbf{X}$
- The equality

$$\mathbb{P}(\langle \mathbf{Z}, \mathbf{H}_a \rangle = \langle \mathbf{X}, \mathbf{H}_a \rangle, \forall a = 1, \ldots, k) = q^{-k}$$

  no longer holds
- Nevertheless....

Achievability under sparse model

Theorem (Achievability under sparse model)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn according to* δ*-sparse distribution*

Achievability under sparse model

Theorem (Achievability under sparse model)

Assume

- *Sensing matrices $\mathbf{H}_a$ drawn according to δ-sparse distribution*
- *Min-rank decoder is used*

Achievability under sparse model

Theorem (Achievability under sparse model)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn according to* δ*-sparse distribution*
- *Min-rank decoder is used*
- $r/n \to \gamma$ *(constant)*

Achievability under sparse model

Theorem (Achievability under sparse model)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn according to* δ*-sparse distribution*
- *Min-rank decoder is used*
- $r/n \to \gamma$ *(constant)*

If the sparsity factor satisfies

$$\delta \in \Omega\left(\frac{\log n}{n}\right)$$

Achievability under sparse model

Theorem (Achievability under sparse model)

Assume

- *Sensing matrices* $\mathbf{H}_a$ *drawn according to* δ*-sparse distribution*
- *Min-rank decoder is used*
- $r/n \to \gamma$ *(constant)*

If the sparsity factor satisfies

$$\delta \in \Omega\left(\frac{\log n}{n}\right)$$

and the number of measurements satisfies

$$k > (2+\varepsilon)\gamma\left(1-\frac{\gamma}{2}\right)n^2$$

then $\mathbb{P}(\mathcal{E}_n) \to 0$.

Remarks on achievability under sparse model

- No loss in measurement complexity

Remarks on achievability under sparse model

- No loss in measurement complexity
- Average number of entries in $\mathbf{H}_a$ is $\Omega(n \log n)$

Remarks on achievability under sparse model

- No loss in measurement complexity
- Average number of entries in $\mathbf{H}_a$ is $\Omega(n \log n)$
- Number of measurements matches uniform model and information-theoretic (Fano's) lower bound

Remarks on achievability under sparse model

- No loss in measurement complexity
- Average number of entries in $\mathbf{H}_a$ is $\Omega(n \log n)$
- Number of measurements matches uniform model and information-theoretic (Fano's) lower bound
- Roughly speaking, if

$$\delta \in \Omega\left(\frac{\log n}{n}\right)$$

probability of error events still "look like" the uniform model

Proof Strategy

Approximate with uniform model

Proof Strategy

Approximate with uniform model

$$\mathbb{P}(\mathcal{E}_n) \le \sum_{\substack{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X})\\ \|\mathbf{Z}-\mathbf{X}\|_0\le\beta n^2}} \mathbb{P}(\langle\mathbf{Z},\mathbf{H}_a\rangle = \langle\mathbf{X},\mathbf{H}_a\rangle, \forall a = 1,\ldots,k)$$

Proof Strategy

Approximate with uniform model

$$\mathbb{P}(\mathcal{E}_n) \le \sum_{\substack{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X}) \\ \|\mathbf{Z}-\mathbf{X}\|_0\le\beta n^2}} \mathbb{P}(\langle \mathbf{Z}, \mathbf{H}_a\rangle = \langle \mathbf{X}, \mathbf{H}_a\rangle, \forall a = 1,\ldots,k)$$

$$+ \sum_{\substack{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X}) \\ \|\mathbf{Z}-\mathbf{X}\|_0>\beta n^2}} \mathbb{P}(\langle \mathbf{Z}, \mathbf{H}_a\rangle = \langle \mathbf{X}, \mathbf{H}_a\rangle, \forall a = 1,\ldots,k)$$

Proof Strategy

Approximate with uniform model

$$\mathbb{P}(\mathcal{E}_n) \le \sum_{\substack{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X})\\ \|\mathbf{Z}-\mathbf{X}\|_0\le\beta n^2}} \mathbb{P}(\langle\mathbf{Z},\mathbf{H}_a\rangle = \langle\mathbf{X},\mathbf{H}_a\rangle, \forall a = 1,\ldots,k)$$

$$+ \sum_{\substack{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X})\\ \|\mathbf{Z}-\mathbf{X}\|_0>\beta n^2}} \mathbb{P}(\langle\mathbf{Z},\mathbf{H}_a\rangle = \langle\mathbf{X},\mathbf{H}_a\rangle, \forall a = 1,\ldots,k)$$

- First term
  - Few terms
  - Low Hamming weight – can afford loose bound on probability

Proof Strategy

Approximate with uniform model

$$\mathbb{P}(\mathcal{E}_n) \le \sum_{\substack{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X})\\ \|\mathbf{Z}-\mathbf{X}\|_0\le\beta n^2}} \mathbb{P}(\langle\mathbf{Z},\mathbf{H}_a\rangle = \langle\mathbf{X},\mathbf{H}_a\rangle, \forall a = 1,\ldots,k)$$

$$+ \sum_{\substack{\mathbf{Z}\neq\mathbf{X}:\mathrm{rank}(\mathbf{Z})\le\mathrm{rank}(\mathbf{X})\\ \|\mathbf{Z}-\mathbf{X}\|_0>\beta n^2}} \mathbb{P}(\langle\mathbf{Z},\mathbf{H}_a\rangle = \langle\mathbf{X},\mathbf{H}_a\rangle, \forall a = 1,\ldots,k)$$

- First term
  - Few terms
  - Low Hamming weight – can afford loose bound on probability
- Second term
  - Many terms
  - Tight bound on probability (circular convolution of δ-sparse pmf)

Outline

1 Motivation

2 Problem Setup and Summary of Main Results

3 Converse

4 Achievability for Uniform Model

5 Achievability for Sparse Model

6 Coding-Theoretic Interpretations

Interpreting our results from the rank-metric lens

- Recall that a rank-metric code $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m \times n}$ endowed with the rank distance

Interpreting our results from the rank-metric lens

- Recall that a rank-metric code $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m \times n}$ endowed with the rank distance
- The minimum rank distance decoding problem

$$\min_{\mathbf{C} \in \mathcal{C}} \ \mathrm{rank}(\mathbf{R} - \mathbf{C})$$

  is in one-to-one correspondence to the rank minimization problem

Interpreting our results from the rank-metric lens

- Recall that a rank-metric code $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m\times n}$ endowed with the rank distance
- The minimum rank distance decoding problem

$$\min_{\mathbf{C}\in\mathcal{C}} \ \mathrm{rank}(\mathbf{R}-\mathbf{C})$$

is in one-to-one correspondence to the rank minimization problem

$$\begin{aligned} &\text{minimize} \quad && \mathrm{rank}(\mathbf{X}) \\ &\text{subject to} \quad && \langle \mathbf{H}_a, \mathbf{X}\rangle = y_a, \quad a = 1,\ldots,k \end{aligned}$$

with the identification $\mathbf{X} \equiv \mathbf{R}-\mathbf{C}$ and

Interpreting our results from the rank-metric lens

- Recall that a rank-metric code $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m\times n}$ endowed with the rank distance
- The minimum rank distance decoding problem

$$\min_{\mathbf{C}\in\mathcal{C}} \ \mathrm{rank}(\mathbf{R}-\mathbf{C})$$

is in one-to-one correspondence to the rank minimization problem

$$\begin{array}{ll} \text{minimize} & \mathrm{rank}(\mathbf{X}) \\ \text{subject to} & \langle \mathbf{H}_a, \mathbf{X}\rangle = y_a, \quad a = 1,\ldots,k \end{array}$$

with the identification $\mathbf{X} \equiv \mathbf{R}-\mathbf{C}$ and

$$\mathcal{C} := \{\mathbf{C} : \langle \mathbf{C}, \mathbf{H}_a\rangle = 0, a = 1,\ldots,k\}$$

Minimum rank distance

Definition

The *minimum rank distance* *of a rank-metric code* $\mathcal{C}$ *is*

$$d(\mathcal{C}) := \min_{\mathbf{C}_1, \mathbf{C}_2 \in \mathcal{C} : \mathbf{C}_1 \neq \mathbf{C}_2} \operatorname{rank}(\mathbf{C}_1 - \mathbf{C}_2)$$

Minimum rank distance

Definition

The *minimum rank distance* *of a rank-metric code* $\mathcal{C}$ *is*

$$d(\mathcal{C}) := \min_{\mathbf{C}_1, \mathbf{C}_2 \in \mathcal{C} : \mathbf{C}_1 \neq \mathbf{C}_2} \operatorname{rank}(\mathbf{C}_1 - \mathbf{C}_2)$$

If the code is linear (as it is)

$$d(\mathcal{C}) := \min_{\mathbf{C} \in \mathcal{C} : \mathbf{C} \neq \mathbf{0}} \operatorname{rank}(\mathbf{C})$$

How does $d(\mathcal{C})$ behave for the random linear rank-metric code under the uniform and sparse models?

Minimum rank distance

The rate of the code is defined as

$$R = 1 - \frac{k}{n^2}$$

Minimum rank distance

The rate of the code is defined as

$$R = 1 - \frac{k}{n^2}$$

Proposition (Concentration of minimum rank distance)

For the uniform model, with probability approaching 1 as $n \to \infty$, the relative minimum distance satisfies

$$\frac{d(\mathcal{C})}{n} \in (1 - \sqrt{R} - \varepsilon, 1 - \sqrt{R} + \varepsilon)$$

Minimum rank distance: Interpretation

- The rate-dependent function

$$\gamma_{\mathsf{GV}}(R) := 1 - \sqrt{R}$$

can be regarded as the "Gilbert-Varshamov distance" of the random rank-metric code

Minimum rank distance: Interpretation

- The rate-dependent function
$$\gamma_{\mathsf{GV}}(R) := 1 - \sqrt{R}$$
can be regarded as the "Gilbert-Varshamov distance" of the random rank-metric code

- Barg and Forney (2002) derived similar properties for the binary Hamming case

Minimum rank distance: Interpretation

- The rate-dependent function

$$\gamma_{\mathsf{GV}}(R) := 1 - \sqrt{R}$$

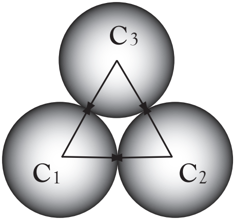
can be regarded as the "Gilbert-Varshamov distance" of the random rank-metric code

- Barg and Forney (2002) derived similar properties for the binary Hamming case

Uniform rank-metric code has minimum distance

$$n\gamma_{\mathsf{GV}}(R) = n(1 - \sqrt{R})$$

with high probability for n large



Ramification in terms of measurement complexity for uniform model

For successful recovery of $\mathbf{X}$ on average,

$$\text{relative minimum distance} \;\geq\; \text{rank-dimension ratio}$$

Ramification in terms of measurement complexity for uniform model

For successful recovery of $\mathbf{X}$ on average,

relative minimum distance $\geq$ rank-dimension ratio

Achievability result may be derived by setting $R = 1 - \frac{k}{n^2}$ and considering

$$\gamma_{\mathsf{GV}}(R) = 1 - \sqrt{R} - \varepsilon \geq \gamma = \frac{r}{n}$$

Ramification in terms of measurement complexity for uniform model

For successful recovery of $\mathbf{X}$ on average,

relative minimum distance $\geq$ rank-dimension ratio

Achievability result may be derived by setting $R = 1 - \frac{k}{n^2}$ and considering

$$\gamma_{\mathsf{GV}}(R) = 1 - \sqrt{R} - \varepsilon \geq \gamma = \frac{r}{n}$$

which is equivalent to

$$k \geq (2 + \varepsilon')\gamma \left(1 - \frac{\gamma}{2}\right) n^2$$

Minimum distances for the sparse model

Proposition (Concentration of minimum rank distance)

For the

$$\delta = \Omega\left(\frac{\log n}{n}\right)$$

sparse model, with probability approaching 1 as $n \to \infty$, the relative minimum distance satisfies

$$\frac{d(\mathcal{C})}{n} \geq 1 - \sqrt{R} - \varepsilon$$

Minimum distances for the sparse model

Proposition (Concentration of minimum rank distance)

For the

$$\delta = \Omega\left(\frac{\log n}{n}\right)$$

sparse model, with probability approaching 1 as $n \to \infty$, the relative minimum distance satisfies

$$\frac{d(\mathcal{C})}{n} \geq 1 - \sqrt{R} - \varepsilon$$

- The minimum distance properties of the uniform and sparse models are identical

Minimum distances for the sparse model

Proposition (Concentration of minimum rank distance)

For the

$$\delta = \Omega\left(\frac{\log n}{n}\right)$$

sparse model, with probability approaching 1 as $n \to \infty$, the relative minimum distance satisfies

$$\frac{d(\mathcal{C})}{n} \geq 1 - \sqrt{R} - \varepsilon$$

- The minimum distance properties of the uniform and sparse models are identical
- So are their measurement complexities!

$$k > (2+\varepsilon)\gamma\left(1 - \frac{\gamma}{2}\right)n^2$$

Concluding remarks

- Derived necessary and sufficient conditions for rank minimization over finite fields

Concluding remarks

- Derived necessary and sufficient conditions for rank minimization over finite fields
- Derived minimum distance properties of rank-metric codes

Concluding remarks

- Derived necessary and sufficient conditions for rank minimization over finite fields
- Derived minimum distance properties of rank-metric codes
- Drawn analogies between number of measurements and minimum distance properties

Concluding remarks

- Derived necessary and sufficient conditions for rank minimization over finite fields
- Derived minimum distance properties of rank-metric codes
- Drawn analogies between number of measurements and minimum distance properties
- Reduced complexity of the min-rank decoder vis-à-vis exhaustive search

Open Problems

- Can the sparsity level of
$$\Omega\left(\frac{\log n}{n}\right)$$
be improved (reduced) further?

Open Problems

- Can the sparsity level of

$$\Omega\left(\frac{\log n}{n}\right)$$

  be improved (reduced) further?

- Tradeoff between sparsity and number of measurements?

Open Problems

- Can the sparsity level of

$$\Omega\left(\frac{\log n}{n}\right)$$

  be improved (reduced) further?

- Tradeoff between sparsity and number of measurements?

Check out the preprint: `http://arxiv.org/abs/1104.4302`