Rank Minimization over Finite Fields

Vincent Y. F. Tan

Laura Balzano, Stark C. Draper

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

ISIT 2011

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	Real $\mathbb{R}$, 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	Real $\mathbb{R}$, 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	Real $\mathbb{R}$, Complex $\mathbb{C}$	Finite Field $\mathbb{F}_q$
Techniques	Functional Analysis	Algebraic coding
Decoding	Convex Optimization Nuclear Norm	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	Real $\mathbb{R}$, Complex $\mathbb{C}$	Finite Field $\mathbb{F}_q$
Techniques	Functional Analysis	Algebraic coding
Decoding	Convex Optimization Nuclear Norm	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	Real $\mathbb{R}$, Complex $\mathbb{C}$	Finite Field $\mathbb{F}_q$
Techniques	Functional Analysis	Algebraic coding
Decoding	Convex Optimization Nuclear Norm	Berlekamp-Massey Error Trapping

Can we draw analogies between the two areas of study?

Rank Minimization over finite fields

Minimum rank distance decoding

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

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$
- A rank-metric code $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m\times n}$ endowed with the rank distance
- Receive some received matrix $\mathbf{R} \in \mathbb{F}_q^{m\times n}$

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$
- A rank-metric code $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m\times n}$ endowed with the rank distance
- Receive some received matrix $\mathbf{R} \in \mathbb{F}_q^{m\times n}$
- 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 $\mathbf{C} \in \mathcal{C}$
- A rank-metric code $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m\times n}$ endowed with the rank distance
- Receive some received matrix $\mathbf{R} \in \mathbb{F}_q^{m\times n}$
- 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 – assumed low-rank

Minimum rank distance decoding

- Transmit some codeword $\mathbf{C} \in \mathcal{C}$
- A rank-metric code $\mathcal{C}$ is a non-empty subset of $\mathbb{F}_q^{m \times n}$ endowed with the rank distance
- Receive some received matrix $\mathbf{R} \in \mathbb{F}_q^{m \times n}$
- Decoding problem (under mild conditions) is

$$\hat{\mathbf{C}} = \underset{\mathbf{C} \in \mathcal{C}}{\arg\min} \ \operatorname{rank}(\mathbf{R} - \mathbf{C})$$

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

$$\hat{\mathbf{X}} = \underset{\mathbf{X} \in \text{coset}}{\arg\min} \ \operatorname{rank}(\mathbf{X})$$

When do low-rank errors occur?

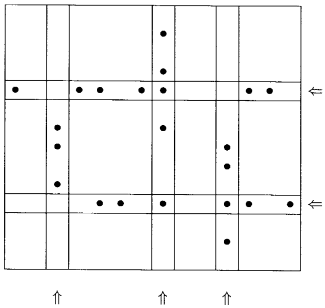
When do low-rank errors occur?

- Probabilistic crisscross error correction

When do low-rank errors occur?

- Probabilistic crisscross error correction

Roth, IT Trans 1997

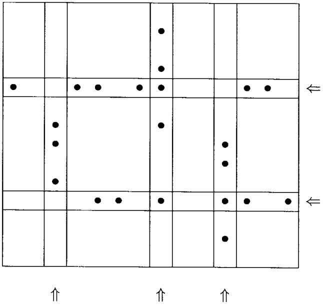


- Data storage applications: Data stored in arrays
- Error patterns confined to two rows and three columns above

When do low-rank errors occur?

- Probabilistic crisscross error correction

Roth, IT Trans 1997



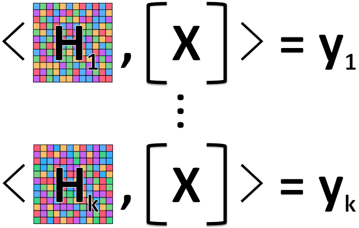
- Data storage applications: Data stored in arrays
- Error patterns confined to two rows and three columns above
- Low-rank errors

Problem Setup: Rank minimization over finite fields

- Square matrix $\mathbf{X} \in \mathbb{F}_q^{n \times n}$ has rank $\leq r \leq n$

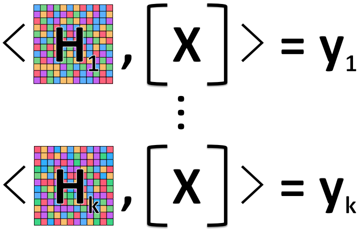
Problem Setup: Rank minimization over finite fields

- Square matrix $\mathbf{X} \in \mathbb{F}_q^{n \times n}$ has rank $\leq r \leq n$



Problem Setup: Rank minimization over finite fields

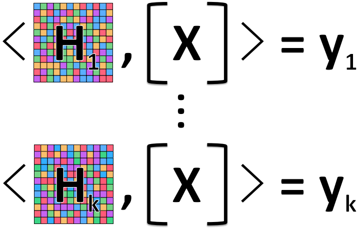
- Square matrix $\mathbf{X} \in \mathbb{F}_q^{n \times n}$ has rank $\leq r \leq n$



- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k \in \mathbb{F}_q^{n \times n}$ are random.

Problem Setup: Rank minimization over finite fields

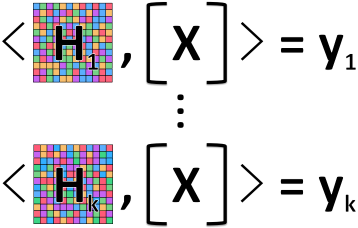
- Square matrix $\mathbf{X} \in \mathbb{F}_q^{n\times n}$ has rank $\leq r \leq n$



- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k \in \mathbb{F}_q^{n\times n}$ are random.
- Arithmetic is performed in the field $\mathbb{F}_q \;\; \Rightarrow \;\; y_a \in \mathbb{F}_q$

Problem Setup: Rank minimization over finite fields

- Square matrix $\mathbf{X} \in \mathbb{F}_q^{n \times n}$ has rank $\leq r \leq n$



- Assume that the sensing matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k \in \mathbb{F}_q^{n \times n}$ are random.
- Arithmetic is performed in the field $\mathbb{F}_q \quad \Rightarrow \quad y_a \in \mathbb{F}_q$
- Given $(y^k, \mathbf{H}^k)$, 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

- k: Num. of linear measurements
- n: Dim. of matrix $\mathbf{X}$
- r: Max. rank of matrix $\mathbf{X}$

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

Critical Quantity

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

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

Critical Quantity

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

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

Critical Quantity

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

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

Critical Quantity

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

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

Critical Quantity

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

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)	$k > (2+\varepsilon)\gamma(1-\gamma/2)n^2$	$\mathbb{P}(\mathcal{E}_n) \to 0$

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

Critical Quantity

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

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

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* $\leq r$
- *Sensing matrices* $\mathbf{H}_a, a = 1, \ldots, k$ *jointly 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 sensing model: Uniform model

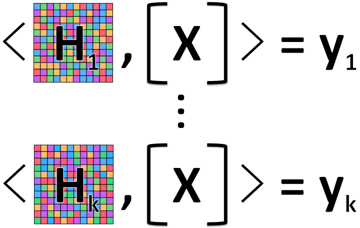
- Now we assume that $\mathbf{X}$ is non-random

A sensing model: Uniform model

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

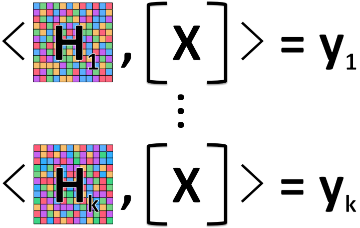
A sensing model: Uniform model

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



A sensing model: Uniform model

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



- 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

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

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

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

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

The min-rank decoder

- We employ the min-rank decoder

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

- 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

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

- 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*
- *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

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

Proof Sketch

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

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

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

Proof Sketch

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

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,

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

Proof Sketch

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

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 $n \times n$ matrices over $\mathbb{F}_q$ of rank $\leq r$ is bounded above by

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

Proof Sketch

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

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,

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

That the number of $n \times n$ matrices over $\mathbb{F}_q$ of rank $\le r$ is bounded above by

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

Remark: Can be extended to noisy case

The Reliability Function

Definition

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

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

The Reliability Function

Definition

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

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

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 rate of a sequence of linear measurement models is defined as

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

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 $1 - k/n^2$.

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

- 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 on $E(R)$ is more tricky. But...

The Reliability Function

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

Proposition (Reliability Function of Min-Rank Decoder)

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 for the min-rank decoder

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 for the min-rank decoder
- de Caen's lower bound: Let $\mathcal{B}_1, \ldots, \mathcal{B}_M$ be events:

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

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 for the min-rank decoder
- de Caen's lower bound: Let $\mathcal{B}_1, \ldots, \mathcal{B}_M$ be events:

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

- Exploit pairwise independence to make statements about error exponents (linear codes achieve capacity in symmetric DMCs)

Sparse sensing model

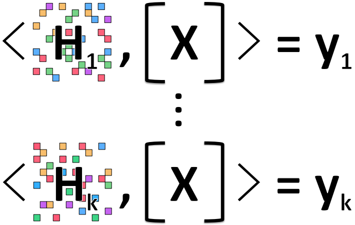
- Assume as usual that $\mathbf{X}$ is non-random

Sparse sensing model

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

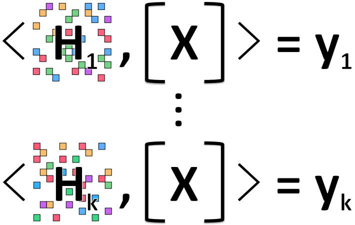
Sparse sensing model

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



Sparse sensing model

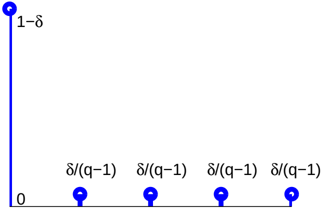
- Assume as usual that $\mathbf{X}$ is non-random
- $\mathrm{rank}(\mathbf{X}) \leq r = \gamma n$
- Sensing matrices are sparse



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

Sparse sensing model

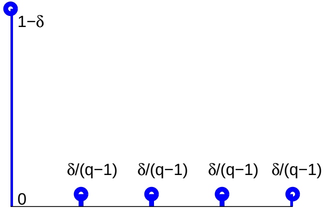
- Each entry of each sensing matrix $\mathbf{H}_a$ is i.i.d. and has a δ-sparse distribution in $\mathbb{F}_q$:



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

Sparse sensing model

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

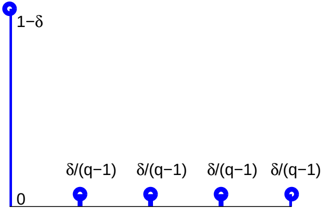


$$\mathbb{P}([\mathbf{H}_a]_{ij} = h) = \begin{cases} 1-\delta & h = 0 \\ \frac{\delta}{q-1} & h \neq 0 \end{cases}$$

- 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 $\mathbb{F}_q$:

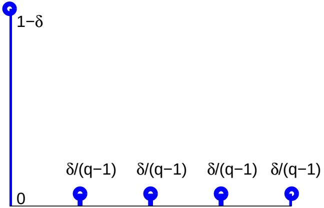


$$\mathbb{P}([\mathbf{H}_a]_{ij} = h) = \begin{cases} 1-\delta & h = 0 \\ \frac{\delta}{q-1} & h \neq 0 \end{cases}$$

- 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

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

  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*
- *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 $\delta = \delta_n$ *satisfies*

$$\delta \in \Omega\left(\frac{\log n}{n}\right) \cap o(1)$$

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 $\delta = \delta_n$ *satisfies*

$$\delta \in \Omega\left(\frac{\log n}{n}\right) \cap o(1)$$

and the number of measurements satisfies

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

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

Proof Strategy

Approximate with uniform model

Proof Strategy

Approximate with uniform model

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

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

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

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

- 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)

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)
- Set $\beta = \Theta(\frac{\delta}{\log n})$ to be the optimal “split”

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
- Analyzed the sparse sensing model

Concluding remarks

- Derived necessary and sufficient conditions for rank minimization over finite fields
- Analyzed the sparse sensing model
- Minimum rank distance properties of rank-metric codes [analogous to Barg and Forney (2002)] were derived

Concluding remarks

- Derived necessary and sufficient conditions for rank minimization over finite fields
- Analyzed the sparse sensing model
- Minimum rank distance properties of rank-metric codes [analogous to Barg and Forney (2002)] were derived
- Drew analogies between number of measurements and minimum distance properties

Concluding remarks

- Derived necessary and sufficient conditions for rank minimization over finite fields
- Analyzed the sparse sensing model
- Minimum rank distance properties of rank-metric codes [analogous to Barg and Forney (2002)] were derived
- Drew analogies between number of measurements and minimum distance properties
- Open Problem: Can the sparsity level of

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

be improved (reduced) further?

Concluding remarks

- Derived necessary and sufficient conditions for rank minimization over finite fields
- Analyzed the sparse sensing model
- Minimum rank distance properties of rank-metric codes [analogous to Barg and Forney (2002)] were derived
- Drew analogies between number of measurements and minimum distance properties
- Open Problem: Can the sparsity level of

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

be improved (reduced) further?

http://arxiv.org/abs/1104.4302