High-Girth Matrices and Polarization

Emmanuel Abbe Yuval Wigderson

Princeton University

ISIT 2015

Fundamental Measures on Matrices

Let A be an $m \times n$ matrix, with $m \leq n$.

Fundamental Measures on Matrices

Let A be an $m \times n$ matrix, with $m \leq n$.

- The rank of A is the maximal number of lin. indep. columns.

Fundamental Measures on Matrices

Let A be an $m \times n$ matrix, with $m \leq n$.

- The rank of A is the maximal number of lin. indep. columns.
- The girth of A is the minimal number of lin. dep. columns.

Fundamental Measures on Matrices

Let A be an $m \times n$ matrix, with $m \leq n$.

- The rank of A is the maximal number of lin. indep. columns.
- The girth of A is the minimal number of lin. dep. columns.
- The probabilistic girth of A is the maximal fraction of columns that are lin. indep. with high probability:

  $$\text{girth}_*(A) = \sup\{p \in [0,1] : A[p] \text{ has lin. indep. cols. w.h.p.}\}$$

  where $A[p] = A$ with each column picked independently with probability p.

Fundamental Measures on Matrices

Let A be an $m \times n$ matrix, with $m \leq n$.

- The rank of A is the maximal number of lin. indep. columns.
- The girth of A is the minimal number of lin. dep. columns.
- The probabilistic girth of A is the maximal fraction of columns that are lin. indep. with high probability:

  $$\text{girth}_*(A) = \sup\{p \in [0,1] : A[p] \text{ has lin. indep. cols. w.h.p.}\}$$

  where $A[p] = A$ with each column picked independently with probability p.

- A is high-girth if
  $\text{girth}_*(A) \sim \text{rank}(A)/n$

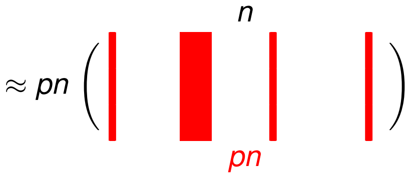
Fundamental Measures on Matrices

Let A be an $m \times n$ matrix, with $m \leq n$.

- The rank of A is the maximal number of lin. indep. columns.
- The girth of A is the minimal number of lin. dep. columns.
- The probabilistic girth of A is the maximal fraction of columns that are lin. indep. with high probability:

  $$\text{girth}_*(A) = \sup\{p \in [0,1] : A[p] \text{ has lin. indep. cols. w.h.p.}\}$$

  where $A[p] = A$ with each column picked independently with probability p.
- A is high-girth if $\text{girth}_*(A) \sim \text{rank}(A)/n$



High-Girth Matrices

High-Girth Matrices

Why we care

How to construct

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\iff$ A is high-girth.

How to construct

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\Longleftrightarrow$ A is high-girth.

0 1 0 0 1 0 1 0 0 0 1 1 0 1 1 0

How to construct

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\Longleftrightarrow$ A is high-girth.

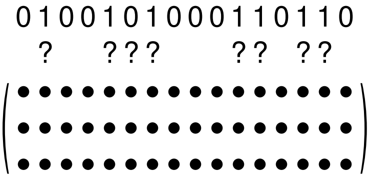
0 1 0 0 1 0 1 0 0 0 1 1 0 1 1 0
? ? ? ? ? ? ? ?

How to construct

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\iff$ A is high-girth.

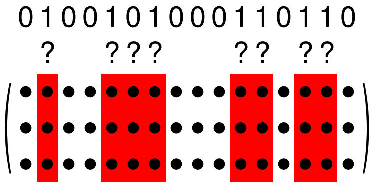


How to construct

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\iff$ A is high-girth.

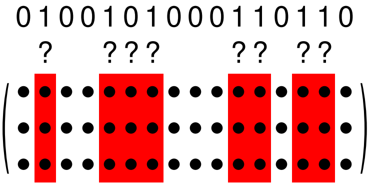


How to construct

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\iff$ A is high-girth.



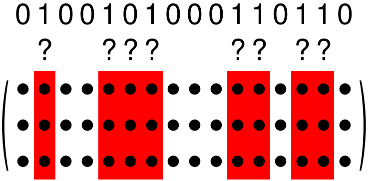
Over $\mathbb{R}$, girth is also called spark, and is related to sparse recovery [Donoho-Elad].

How to construct

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\iff$ A is high-girth.



Over $\mathbb{R}$, girth is also called spark, and is related to sparse recovery [Donoho-Elad].

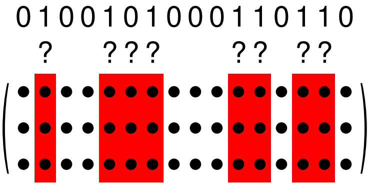
How to construct

Over $\mathbb{F}_2$, a random construction works (take A to have iid Ber($\frac{1}{2}$) entries).

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\iff$ A is high-girth.



Over $\mathbb{R}$, girth is also called spark, and is related to sparse recovery [Donoho-Elad].

How to construct

Over $\mathbb{F}_2$, a random construction works (take A to have iid Ber($\frac{1}{2}$) entries).

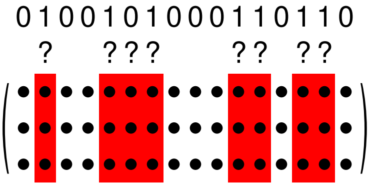
Let $\mathbb{F}$ be a field and

$$G_n = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\otimes \log_2 n}$$

High-Girth Matrices

Why we care

Fact: A linear code C with Parity-Check Matrix A achieves capacity on the BEC(p) $\iff$ A is high-girth.



Over $\mathbb{R}$, girth is also called spark, and is related to sparse recovery [Donoho-Elad].

How to construct

Over $\mathbb{F}_2$, a random construction works (take A to have iid Ber($\frac{1}{2}$) entries).

Let $\mathbb{F}$ be a field and

$$G_n = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{\otimes \log_2 n}$$

Polar codes and Reed-Muller codes are high-girth submatrices of G_n. How about other submatrices? [Arıkan, Kumar-Pfister, Kudekar et al.]

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

$$G_n^{(i)} = \text{first } i \text{ rows of } G_n$$

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

$$G_n^{(i)} = \text{first } i \text{ rows of } G_n$$

$$G_n^{(i)}[p]$$

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

$$G_n^{(i)} = \text{first } i \text{ rows of } G_n$$

$$\text{rank}_{\mathbb{F}}(G_n^{(i)}[p])$$

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

$$G_n^{(i)} = \text{first } i \text{ rows of } G_n$$

Basic quantity:

$$\mathbb{E}(\mathrm{rank}_{\mathbb{F}}(G_n^{(i)}[p]))$$

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

$$G_n^{(i)} = \text{first } i \text{ rows of } G_n$$

Basic quantity:

$$\mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i)}[p]))$$

We define the conditional rank values as

$$\rho_{n,p}(i) = \mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i)}[p])) - \mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i-1)}[p]))$$

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

$$G_n^{(i)} = \text{first } i \text{ rows of } G_n$$

Basic quantity:

$$\mathbb{E}(\mathrm{rank}_{\mathbb{F}}(G_n^{(i)}[p]))$$

We define the conditional rank values as

$$\begin{aligned}\rho_{n,p}(i) &= \mathbb{E}(\mathrm{rank}_{\mathbb{F}}(G_n^{(i)}[p])) - \mathbb{E}(\mathrm{rank}_{\mathbb{F}}(G_n^{(i-1)}[p])) \\ &= \mathbb{P}(i\text{th row of } G_n[p] \text{ indep. of } G_n^{(i-1)}[p])\end{aligned}$$

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

$$G_n^{(i)} = \text{first } i \text{ rows of } G_n$$

Basic quantity:

$$\mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i)}[p]))$$

We define the conditional rank values as

$$\begin{aligned}\rho_{n,p}(i) &= \mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i)}[p])) - \mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i-1)}[p])) \\ &= \mathbb{P}(i\text{th row of } G_n[p] \text{ indep. of } G_n^{(i-1)}[p])\end{aligned}$$

Properties of $\rho_{n,p}(1), \ldots, \rho_{n,p}(n)$?

The conditional rank (COR) process

Our construction is inspired by polar codes, but replaces an information-theoretic measure with a linear-algebraic measure.

$$G_n^{(i)} = \text{first } i \text{ rows of } G_n$$

Basic quantity:

$$\mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i)}[p]))$$

We define the conditional rank values as

$$\begin{aligned}\rho_{n,p}(i) &= \mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i)}[p])) - \mathbb{E}(\text{rank}_{\mathbb{F}}(G_n^{(i-1)}[p])) \\ &= \mathbb{P}(i\text{th row of } G_n[p] \text{ indep. of } G_n^{(i-1)}[p])\end{aligned}$$

Properties of $\rho_{n,p}(1), \ldots, \rho_{n,p}(n)$? As we will see, leaves of a branching process.

The COR process

$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$

The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\rho_{2,p}(1) = 2p - p^2$$

The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\rho_{2,p}(1) = 2p - p^2$$

$$\rho_{2,p}(2) = p^2$$

The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\rho_{2,p}(1) = 2p - p^2$$
$$\rho_{2,p}(2) = p^2$$

Theorem

The COR values are the leaves of the branching process

$$x \mapsto (2x - x^2, x^2)$$

initialized at $x = p$.

The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

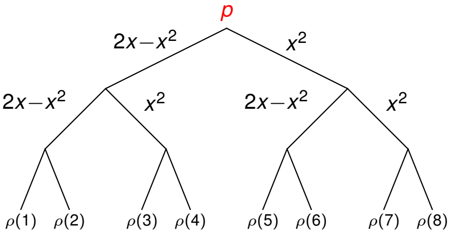
$$\rho_{2,p}(1) = 2p - p^2$$
$$\rho_{2,p}(2) = p^2$$

Theorem

The COR values are the leaves of the branching process

$$x \mapsto (2x - x^2, x^2)$$

initialized at $x = p$.



The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

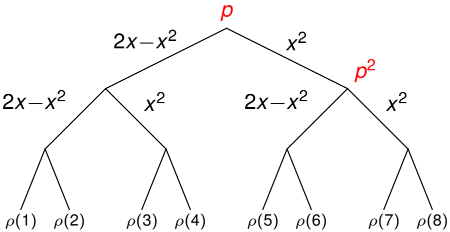
$$\rho_{2,p}(1) = 2p - p^2$$
$$\rho_{2,p}(2) = p^2$$

Theorem

The COR values are the leaves of the branching process

$$x \mapsto (2x - x^2, x^2)$$

initialized at $x = p$.



The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\rho_{2,p}(1) = 2p - p^2$$

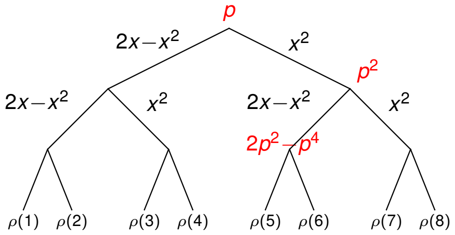
$$\rho_{2,p}(2) = p^2$$

Theorem

The COR values are the leaves of the branching process

$$x \mapsto (2x - x^2, x^2)$$

initialized at $x = p$.



The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\rho_{2,p}(1) = 2p - p^2$$

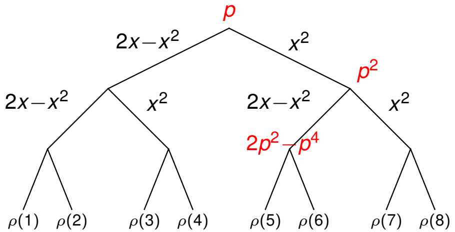
$$\rho_{2,p}(2) = p^2$$

Theorem

The COR values are the leaves of the branching process

$$x \mapsto (2x - x^2, x^2)$$

initialized at $x = p$.



Note 1: This is true over any field!

The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

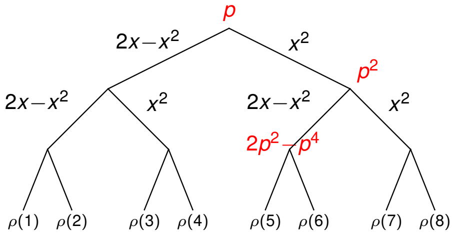
$$\rho_{2,p}(1) = 2p - p^2$$
$$\rho_{2,p}(2) = p^2$$

Theorem

The COR values are the leaves of the branching process

$$x \mapsto (2x - x^2, x^2)$$

initialized at $x = p$.



Note 1: This is true over any field!

Note 2: This is the Bhattacharyya (Z) process for the BEC.

The COR process

$$\rho_{n,p}(i) = \mathbb{P}(\text{rank}_{\mathbb{F}} \text{ increase at row } i)$$

$$G_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\rho_{2,p}(1) = 2p - p^2$$

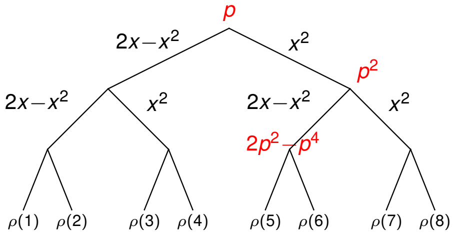
$$\rho_{2,p}(2) = p^2$$

Theorem

The COR values are the leaves of the branching process

$$x \mapsto (2x - x^2, x^2)$$

initialized at $x = p$.



Note 1: This is true over any field!

Note 2: This is the Bhattacharyya (Z) process for the BEC.

Note 3: An upper bound for the Z process on any channel.

Consequences using Polarization Results

Consequences using Polarization Results

1. The COR process polarizes:
   - pn (high) values tend to 1
   - $(1 - p)n$ (low) values tend to 0

Consequences using Polarization Results

1. The COR process polarizes:
   - pn (high) values tend to 1
   - $(1 - p)n$ (low) values tend to 0
2. Take the high rows as a matrix A_n. Then A_n is high-girth over any field.

Consequences using Polarization Results

1. The COR process polarizes:
   - pn (high) values tend to 1
   - $(1 - p)n$ (low) values tend to 0
2. Take the high rows as a matrix A_n. Then A_n is high-girth over any field.
   - ✓ Idea of Proof: $A_n[p]$ has only rows that are likely to be linearly independent of previous rows.

Consequences using Polarization Results

1. The COR process polarizes:
   - pn (high) values tend to 1
   - $(1-p)n$ (low) values tend to 0
2. Take the high rows as a matrix A_n. Then A_n is high-girth over any field.
   - ✓ Idea of Proof: $A_n[p]$ has only rows that are likely to be linearly independent of previous rows.
3. The code with PCM A_n achieves capacity on the BEC(p). (In fact, on the symmetric erasure channel over any field.) This gives a new proof that polar codes achieve capacity on the BEC.

Consequences using Polarization Results

1. The COR process polarizes:
   - pn (high) values tend to 1
   - $(1 - p)n$ (low) values tend to 0
2. Take the high rows as a matrix A_n. Then A_n is high-girth over any field.
   - ✓ Idea of Proof: $A_n[p]$ has only rows that are likely to be linearly independent of previous rows.
3. The code with PCM A_n achieves capacity on the BEC(p). (In fact, on the symmetric erasure channel over any field.) This gives a new proof that polar codes achieve capacity on the BEC.
4. Working over $\mathbb{R}$, these COR matrices are binary matrices with good Sparse Recovery properties (can distinguish most pairs of sparse patterns).

Open Problems

Open Problems

1. We can prove that COR codes achieve a rate

$$1 - 2\sqrt{p(1-p)}$$

   on the BSC(p).

Open Problems

1. We can prove that COR codes achieve a rate

$$1 - 2\sqrt{p(1-p)} \leq 1 - H(p)$$

on the BSC(p).

Open Problems

1. We can prove that COR codes achieve a rate

$$1 - 2\sqrt{p(1-p)} \leq 1 - H(p)$$

on the BSC(p). Can we get a better rate?

Open Problems

1. We can prove that COR codes achieve a rate

$$1 - 2\sqrt{p(1-p)} \leq 1 - H(p)$$

   on the BSC(p). Can we get a better rate?
2. This proof exploits relationship between COR codes and polar codes. Can we prove it from the high-girth property?

Open Problems

1. We can prove that COR codes achieve a rate

$$1 - 2\sqrt{p(1-p)} \leq 1 - H(p)$$

   on the BSC(p). Can we get a better rate?
2. This proof exploits relationship between COR codes and polar codes. Can we prove it from the high-girth property? $\Longleftrightarrow$ Some code achieves a rate on the BEC; does this imply achieving another rate on the BSC?

Open Problems

1. We can prove that COR codes achieve a rate

   $$1 - 2\sqrt{p(1-p)} \leq 1 - H(p)$$

   on the BSC(p). Can we get a better rate?
2. This proof exploits relationship between COR codes and polar codes. Can we prove it from the high-girth property?
   $\Longleftrightarrow$ Some code achieves a rate on the BEC; does this imply achieving another rate on the BSC?
3. What about error channels over larger fields?
   AWGN or other continuous channels?
   Polarization over other combinatorial objects?

Thank you!