The rank of adjacency matrices of sparse Erdős-Rényi graphs over finite fields

Noela Müller

Eindhoven University of Technology

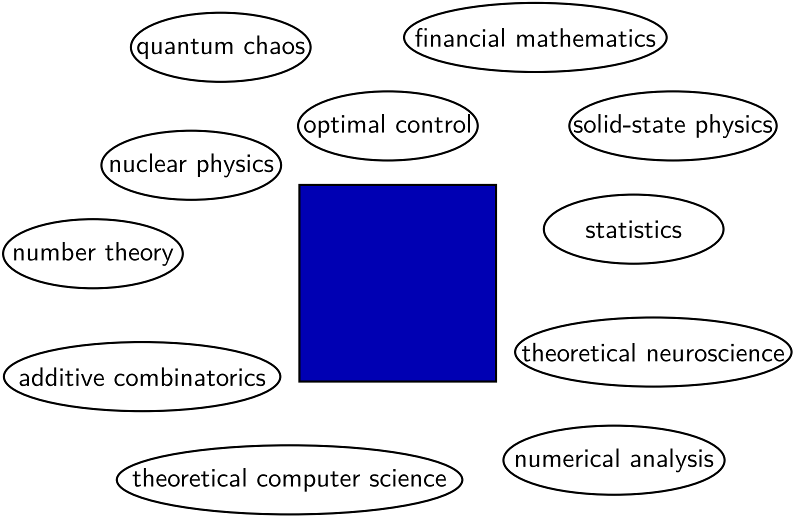
ADYN summer school, Dortmund, July 1 2022

joint work with R. v. d. Hofstad and H. Zhu

Overview

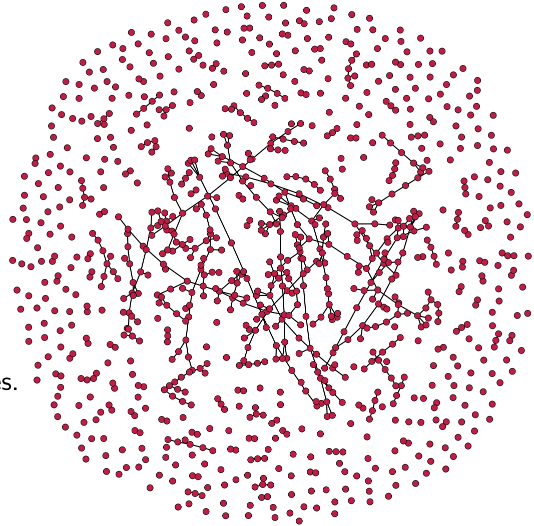
1. Introduction
2. Selected results
3. Rank formula
4. Proof outline

Random matrices



Adjacency matrices

This talk:
Adjacency matrix of
Erdős-Rényi random graph.



Erdős-Rényi random graph $\boldsymbol{G}_{n,p}$:

- $V(\boldsymbol{G}_{n,p}) = \{1, \ldots, n\}$
- $E(\boldsymbol{G}_{n,p})$: Each possible edge is included with probability p, independently of all other edges.

Adjacency Matrix $\boldsymbol{A}_{n,p}$:
$\boldsymbol{A}_{n,p}(i,j) = \mathbb{1}\{\{i,j\} \in E(\boldsymbol{G}_{n,p})\}$

RANK OF $\boldsymbol{A}_{n,p}$

Question: What can be said about the rank of the matrix $\boldsymbol{A}_{n,p}$?

Rank of $\boldsymbol{A}_{n,p}$

Question: What can be said about the rank of the matrix $\boldsymbol{A}_{n,p}$?

The rank of $\boldsymbol{A}_{n,p}$ can be studied over various fields.

To emphasise this, we write $\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_{n,p})$, where $\mathbb{F}$ is an arbitrary field.

Rank of $\boldsymbol{A}_{n,p}$

Question: What can be said about the rank of the matrix $\boldsymbol{A}_{n,p}$?

The rank of $\boldsymbol{A}_{n,p}$ can be studied over various fields.

To emphasise this, we write $\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_{n,p})$, where $\mathbb{F}$ is an arbitrary field.

This talk: Rank of $\boldsymbol{A}_{n,p}$ over finite field $\mathbb{F}$ in the regime $p = d/n$, $d > 0$.

An easy upper bound

Since every isolated vertex contributes a zero-row to the adjacency matrix, we have

$$\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_{n,p}) \leq n - \# \text{ isolated vertices in } \boldsymbol{G}_{n,p}.$$

Moderately sparse setting

Theorem (Costello, Tao & Vu '04, Costello & Vu '06)

Fix $d > 1/2$. Then for any $p \in \left(\frac{d\ln(n)}{n}, \frac{1}{2}\right]$, w.h.p.,

$$\mathrm{rk}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = n - \#\ \textit{isolated vertices of}\ \boldsymbol{G}_{n,p}.$$

Moderately sparse setting

Theorem (Costello, Tao & Vu '04, Costello & Vu '06)

Fix $d > 1/2$. Then for any $p \in \left(\frac{d\ln(n)}{n}, \frac{1}{2}\right]$, w.h.p.,

$$\mathrm{rk}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = n - \#\ \textit{isolated vertices of}\ \boldsymbol{G}_{n,p}.$$

Equivalently: $\mathrm{nul}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = \#$ isolated vertices of $\boldsymbol{G}_{n,p}$

Moderately sparse setting

Theorem (Costello, Tao & Vu '04, Costello & Vu '06)

Fix $d > 1/2$. Then for any $p \in \left(\frac{d\ln(n)}{n}, \frac{1}{2}\right]$, w.h.p.,

$$\mathrm{rk}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = n - \#\ \textit{isolated vertices of}\ \boldsymbol{G}_{n,p}.$$

Equivalently: $\mathrm{nul}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = \#$ isolated vertices of $\boldsymbol{G}_{n,p}$

This implies that

- $\mathrm{rk}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) < n$ w.h.p. for $p \in [0, d\ln(n)/n)$, where $d < 1$, and
- $\mathrm{rk}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = n$ w.h.p. for $p \in (d\ln(n)/n, 1/2]$, where $d > 1$.

Reducing p

Assume that $p \leq d \ln(n)/n$ for $d < 1/2$ (and $p = \omega(n^{-3/2})$ is not too small).

Reducing p

Assume that $p \leq d \ln(n)/n$ for $d < 1/2$ (and $p = \omega(n^{-3/2})$ is not too small).

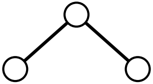
Is it reasonable to believe that
$\mathrm{nul}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = \#$ isolated vertices of $\boldsymbol{G}_{n,p}$?

Reducing p

Assume that $p \leq d \ln(n)/n$ for $d < 1/2$ (and $p = \omega(n^{-3/2})$ is not too small).

Is it reasonable to believe that
$\mathrm{nul}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = \#$ isolated vertices of $\boldsymbol{G}_{n,p}$?

Here, w.h.p., we have more sophisticated types of local dependencies.
Most obvious obstruction:

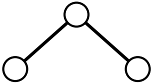


Reducing p

Assume that $p \leq d \ln(n)/n$ for $d < 1/2$ (and $p = \omega(n^{-3/2})$ is not too small).

Is it reasonable to believe that
$\mathrm{nul}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) = \#$ isolated vertices of $\boldsymbol{G}_{n,p}$?

Here, w.h.p., we have more sophisticated types of local dependencies.
Most obvious obstruction:



The rows corresponding to the leaves are not zero, but equal, and this further reduces the rank. So w.h.p.,

$$\mathrm{rk}_{\mathbb{R}}(\boldsymbol{A}_{n,p}) < n - \# \text{ isolated vertices of } \boldsymbol{G}_{n,p}.$$

A modified graph

Each cherry increases the nullity by one, and thus the same as an isolated vertex.

A modified graph

Each cherry increases the nullity by one, and thus the same as an isolated vertex.

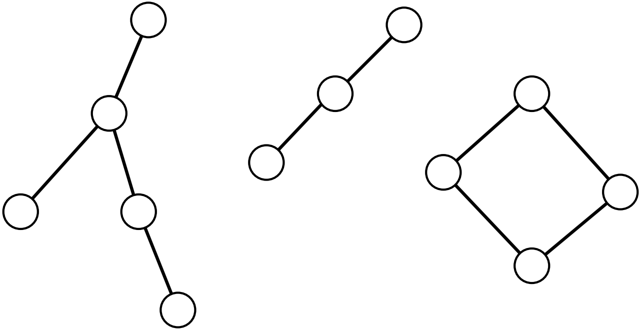
Idea:

Consider an iterative procedure that deletes vertices of degree one and their neighbours
(thus turning cherries into isolated vertices).

Leaf removal process

Procedure:

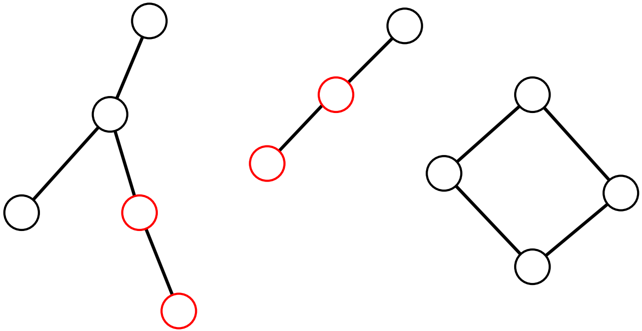
Iteratively remove vertices of degree one along with their unique neighbour.



Leaf removal process

Procedure:

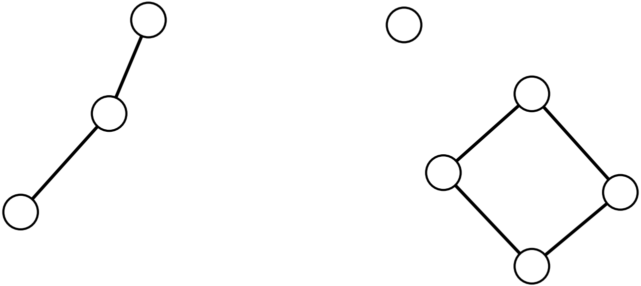
Iteratively remove vertices of degree one along with their unique neighbour.



Leaf removal process

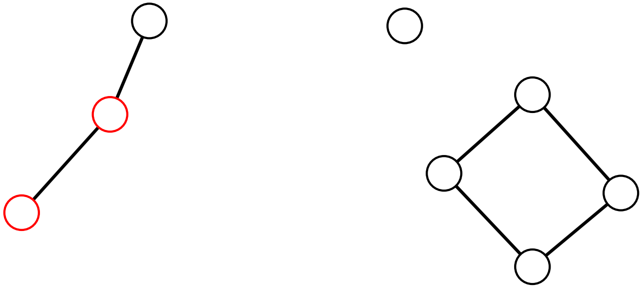
Procedure:

Iterativeły remove vertices of degree one along with their unique neighbour.



Leaf removal process

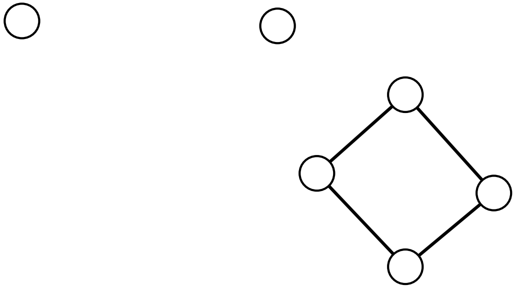
Procedure:
Iteratively remove vertices of degree one along with their unique neighbour.



Leaf removal process

Procedure:

Iteratively remove vertices of degree one along with their unique neighbour.



Leaf removal and nullity

Observation:

Each step of the process does not change the nullity.

Leaf removal and nullity

Observation:
Each step of the process does not change the nullity.

After relabelling the vertices, the adjacency matrix has the following form.

$$A = \begin{pmatrix} 0 & 1 & \mathbb{0}^T \\ 1 & 0 & b^T \\ \mathbb{0} & b & A' \end{pmatrix}$$

Leaf removal and nullity

Observation:

Each step of the process does not change the nullity.

After relabelling the vertices, the adjacency matrix has the following form.

$$A = \begin{pmatrix} 0 & 1 & \mathbb{0}^T \\ 1 & 0 & b^T \\ \mathbb{0} & b & A' \end{pmatrix}$$

Then

$$x = (x_1, x_2, x')^T \in \ker(A) \qquad \Longleftrightarrow \qquad \begin{aligned} x_2 &= 0 \\ x_1 &= -b^T x' \\ A'x' &= (-x_2)b \end{aligned}$$

Leaf removal and nullity

$$x = (x_1, x_2, x')^T \in \ker(A) \qquad \iff \qquad \begin{aligned} x_2 &= 0 \\ x_1 &= -b^T x' \\ x' &\in \ker(A') \end{aligned}$$

Leaf removal and nullity

$$x = (x_1, x_2, x')^T \in \ker(A) \quad \Longleftrightarrow \quad \begin{aligned} x_2 &= 0 \\ x_1 &= -b^T x' \\ x' &\in \ker(A') \end{aligned}$$

Conclusion: $\mathrm{nul}(A) = \mathrm{nul}(A')$.

Leaf removal and nullity

Call the graph obtained from $\boldsymbol{G}_{n,p}$ through the leaf removal process $\boldsymbol{G}_{n,p}^{\mathsf{KS}}$ and its adjacency matrix $\boldsymbol{A}_{n,p}^{\mathsf{KS}}$.

Then $\boldsymbol{G}_{n,p}^{\mathsf{KS}}$ has at least as many isolated vertices as the original graph, and no cherries.

Leaf removal and nullity

Call the graph obtained from $\boldsymbol{G}_{n,p}$ through the leaf removal process $\boldsymbol{G}_{n,p}^{\mathrm{KS}}$ and its adjacency matrix $\boldsymbol{A}_{n,p}^{\mathrm{KS}}$.

Then $\boldsymbol{G}_{n,p}^{\mathrm{KS}}$ has at least as many isolated vertices as the original graph, and no cherries.

Is it true that

$$\mathrm{nul}_{\mathbb{F}}(\boldsymbol{A}_{n,p}) = \mathrm{nul}_{\mathbb{F}}(\boldsymbol{A}_{n,p}^{\mathrm{KS}}) = \#\text{ of isolated vertices in } \boldsymbol{G}_{n,p}^{\mathrm{KS}}?$$

Leaf removal and nullity

Call the graph obtained from $\boldsymbol{G}_{n,p}$ through the leaf removal process $\boldsymbol{G}_{n,p}^{\mathsf{KS}}$ and its adjacency matrix $\boldsymbol{A}_{n,p}^{\mathsf{KS}}$.

Then $\boldsymbol{G}_{n,p}^{\mathsf{KS}}$ has at least as many isolated vertices as the original graph, and no cherries.

Is it true that

$$\mathrm{nul}_{\mathbb{F}}(\boldsymbol{A}_{n,p}) = \mathrm{nul}_{\mathbb{F}}(\boldsymbol{A}_{n,p}^{\mathsf{KS}}) = \#\text{ of isolated vertices in } \boldsymbol{G}_{n,p}^{\mathsf{KS}}?$$

What is the number of isolated vertices in $\boldsymbol{G}_{n,p}^{\mathsf{KS}}$?

Sparse setting

From now on: $p = d/n$ for $d > 0$.

Sparse setting

From now on: $p = d/n$ for $d > 0$.

Let

$$\phi(\alpha) = \exp(d(\alpha - 1))$$

be the probability generating function of a Poisson random variable, and $M_d : [0,1] \to \mathbb{R}$ by

$$M_d(\alpha) = \phi(1 - \phi(\alpha)) + (1 + d(1 - \alpha))\phi(\alpha).$$

Leaf removal and nullity

Theorem (Karp & Sipser (1981), Aronson, Frieze & Pittel (1998))

Fix $d > 0$. Then for $p = d/n$,

$$\frac{\#\ \textit{of isolated vertices in}\ \boldsymbol{G}_{n,p}^{KS}}{n} \xrightarrow{\mathbb{P}} \max_{\alpha \in [0,1]} M_d(\alpha) - 1.$$

Leaf removal and nullity

Theorem (Karp & Sipser (1981), Aronson, Frieze & Pittel (1998))

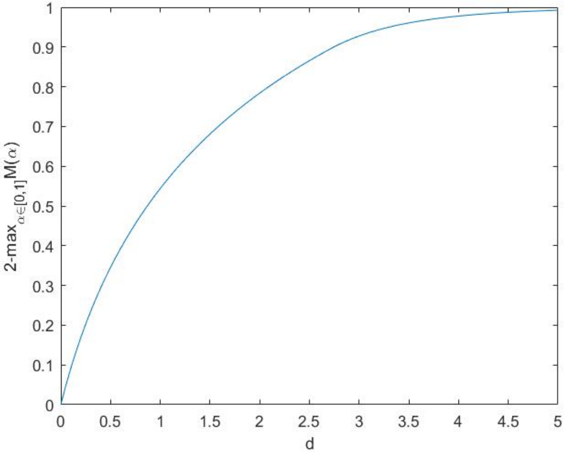
Fix $d > 0$. Then for $p = d/n$,

$$\frac{\#\text{ of isolated vertices in } \boldsymbol{G}_{n,p}^{KS}}{n} \xrightarrow{\mathbb{P}} \max_{\alpha \in [0,1]} M_d(\alpha) - 1.$$

Moreover, for $d \leq e$, the number of non-isolated vertices is $o(n)$ w.h.p. Thus, by the rank-nullity theorem, we obtain the asymptotic rank formula

$$\frac{\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_{n,p})}{n} = 1 - \frac{\mathrm{nul}_{\mathbb{F}}(\boldsymbol{A}_{n,p}^{\mathrm{KS}})}{n} \xrightarrow{\mathbb{P}} 2 - \max_{\alpha \in [0,1]} M_d(\alpha).$$

Rank function



Asymptotic rank formula

For $d > e$, the "2-core"-argument only gives an upper bound on the asymptotic rank.

Asymptotic rank formula

For $d > e$, the "2-core"-argument only gives an upper bound on the asymptotic rank.

Theorem (Bordenave, Lelarge & Salez (2011))

Fix $d > 0$. Then for $p = d/n$,

$$\frac{\mathrm{rk}_{\mathbb{R}}(\boldsymbol{A}_{n,p})}{n} \to 2 - \max_{\alpha \in [0,1]} M_d(\alpha) \qquad a.s.$$

(Bordenave, Lelarge and Salez actually provide an asymptotic rank formula for all random graph sequences $(\boldsymbol{G}'_n)_{n \geq 1}$ converging locally to a rooted Galton-Watson tree, provided the latter satisfies a certain degree condition.)

A matter of field

Does the result of Bordenave, Lelarge and Salez transfer to an asymptotic rank formula over other fields than $\mathbb{R}$?

A matter of field

Does the result of Bordenave, Lelarge and Salez transfer to an asymptotic rank formula over other fields than $\mathbb{R}$?

Their proof crucially relies on the fact that because of the spectral theorem for symmetric matrices over $\mathbb{R}$, $\boldsymbol{A}_{n,p}$ is diagonalisable by an orthogonal matrix.

(They then analyse the associated spectral measure and provide an explicit formula for the asymptotic multiplicity of the eigenvalue 0 in terms of the degree generating function.)

A matter of field

Does the result of Bordenave, Lelarge and Salez transfer to an asymptotic rank formula over other fields than $\mathbb{R}$?

Their proof crucially relies on the fact that because of the spectral theorem for symmetric matrices over $\mathbb{R}$, $\boldsymbol{A}_{n,p}$ is diagonalisable by an orthogonal matrix.

(They then analyse the associated spectral measure and provide an explicit formula for the asymptotic multiplicity of the eigenvalue 0 in terms of the degree generating function.)

However, symmetric matrices over finite fields are not necessarily diagonalisable, and the approach through an empirical spectral measure becomes unfeasible.

An upper bound

Since for any field $\mathbb{F}$ and $A \in \{0,1\}^{n\times n}$,

$$\mathrm{rk}_{\mathbb{F}}(A) \le \mathrm{rk}_{\mathbb{Q}}(A) = \mathrm{rk}_{\mathbb{R}}(A),$$

the result of Bordenave, Lelarge and Salez immediately implies the upper bound

$$\limsup_{n\to\infty} \frac{\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_n)}{n} \le 2 - \max_{\alpha\in[0,1]} M_d(\alpha) \qquad \text{a.s.}$$

Alternatively, the same upper bound follows from the leaf removal argument and the result of Karp & Sipser.

An upper bound

Since for any field $\mathbb{F}$ and $A \in \{0,1\}^{n \times n}$,

$$\mathrm{rk}_{\mathbb{F}}(A) \leq \mathrm{rk}_{\mathbb{Q}}(A) = \mathrm{rk}_{\mathbb{R}}(A),$$

the result of Bordenave, Lelarge and Salez immediately implies the upper bound

$$\limsup_{n\to\infty} \frac{\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_n)}{n} \leq 2 - \max_{\alpha\in[0,1]} M_d(\alpha) \qquad \text{a.s.}$$

Alternatively, the same upper bound follows from the leaf removal argument and the result of Karp & Sipser.

Question: Is this bound tight for finite fields and $d > e$?

Towards a more combinatorial approach

A non-symmetric random matrix model

- $(\boldsymbol{k}_i)_{i\geq 1}$ be i.i.d. copies of a random variable $\boldsymbol{k}$ (`ROW`),
- $(\boldsymbol{d}_i)_{i\geq 1}$ be i.i.d. copies of a random variable $\boldsymbol{d}$ (`COLUMN`).

$$
\mathbb{A} = \left(\begin{array}{ccccccccc}
 & & * & & & * & & & \\
* & & & * & & & & * & \\
 & & & & & & & & * \\
 & & & & & & & & \\
 & * & & & & * & & & \\
* & & & & & * & & & * \\
 & & & & & & & & \\
* & & & & * & & & & \\
 & & & * & & & & & \\
 & & & & & & * & * &
\end{array}\right)
\begin{array}{l}
\boldsymbol{k}_1 \\ \boldsymbol{k}_2 \\ \\ \\ \\ \\ \vdots \\ \\ \\ \boldsymbol{k}_m
\end{array}
$$

$$
\begin{array}{ccccccccc}
\boldsymbol{d}_1 & \boldsymbol{d}_2 & & & \cdots & & & & \boldsymbol{d}_n
\end{array}
$$

Asymptotic rank formula for non-symmetric model

Theorem (Coja-Oghlan, Ergür, Gao, Hetterich, Rolvien '22+)

For any field $\mathbb{F}$,

$$\frac{\mathrm{rk}(\mathbb{A})}{n} \xrightarrow{\mathbb{P}} 1 - \max_{z \in [0,1]} \Phi(z) \qquad \text{as } n \to \infty.$$

Rank result for finite fields

Theorem (v. d. Hofstad, M., Zhu '22+)

For any $d > 0$, any coupling $(\boldsymbol{A}_{n,p})_{n\geq 1}$ and any field $\mathbb{F}$,

$$\frac{\mathrm{rk}_{\mathbb{F}}\left(\boldsymbol{A}_{n,p}\right)}{n} \xrightarrow{a.s.} 2 - \max_{\alpha\in[0,1]} M_d(\alpha), \qquad n \to \infty.$$

This recovers the result of Bordenave, Lelarge and Salez for the case $\mathbb{F} = \mathbb{R}$.

Proof Overview

Proof structure

In light of the previous discussion, there are three steps:

1. Upper bound on the expected rank: Via rank result over $\mathbb{R}$ or leaf removal argument.
2. Reduce almost sure lower bound to lower bound on expectations: Via martingale argument.
3. Lower bound on expected rank: Via Aizenman-Sims-Starr scheme.

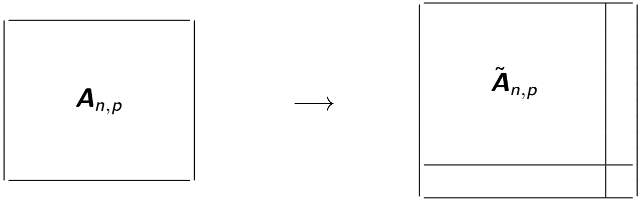
Today: Sketch of third part.

Aizenman-Sims-Starr scheme

Basic idea following Coja-Oghlan, Ergür, Gao, Hetterich, Rolvien:
Grow the matrices one-by-one and observe the change in rank.

By the Stolz-Cesàro Theorem,

$$\liminf_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_{n,p})\right] \geq \liminf_{n\to\infty} \left(\mathbb{E}\left[\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_{n+1,p})\right] - \mathbb{E}\left[\mathrm{rk}_{\mathbb{F}}(\boldsymbol{A}_{n,p})\right]\right).$$



Intermediate step

We break this into the two steps

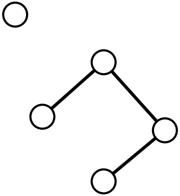
$$\left|\begin{array}{c} \\ \boldsymbol{A}_{n,p} \\ \\ \end{array}\right| \longrightarrow \left|\begin{array}{c} \\ \tilde{\boldsymbol{A}}_{n,p} \\ \\ \end{array}\right|$$

and

$$\left|\begin{array}{c} \\ \tilde{\boldsymbol{A}}_{n,p} \\ \\ \end{array}\right| \longrightarrow \left|\begin{array}{c|c} \tilde{\boldsymbol{A}}_{n,p} & \\ \hline & \\ \end{array}\right|$$

Rank of $\boldsymbol{A}_{n,p}$

The rank may behave in a complex way under the addition of edges:

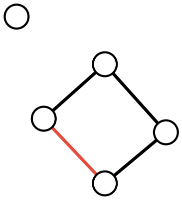


$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathrm{rk}_{\mathbb{F}}(A) = 4$$

Rank of $\boldsymbol{A}_{n,p}$

The rank may behave in a complex way under the addition of edges:

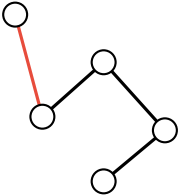


$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathrm{rk}_{\mathbb{F}}(A) = 2$$

Rank of $\boldsymbol{A}_{n,p}$

The rank may behave in a complex way under the addition of edges:

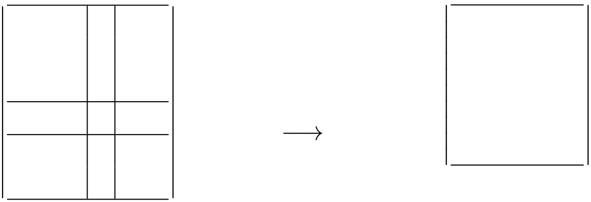


$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathrm{rk}_{\mathbb{F}}(A) = 5$$

Symmetric row and column removal

Observation: The core of both operations is symmetric row and column removal.



$\mathsf{rk}_{\mathbb{F}}(A)$ $\longrightarrow$ $\mathsf{rk}_{\mathbb{F}}(A)\langle i; i\rangle$

Frozen variables

The rank increase can be related to very specific properties of coordinate i under row removal

Frozen variables

The rank increase can be related to very specific properties of coordinate i under row removal

Preparation: Some linear algebra

Coja-Oghlan, Ergür, Gao, Hetterich & Rolvien considered the following concept (with minor modification):

Definition (Linear e-relations)

Let $A \in \mathbb{F}^{m \times n}$.
A set $I \subseteq [n]$ is an **e-relation** of A if there exists a row vector $y \in \mathbb{F}^{1 \times m}$ such that $\mathrm{supp}(yA) = I$. In this case, we call y a **representation** of I in A.

Atomic Linear Relations

Definition (Frozen variables)

Let $A \in \mathbb{F}^{m \times n}$.
If $I = \{i\}$ is an e-relation of A, then we call i **frozen** in A. We denote the set of all frozen variables by $\mathcal{F}(A)$.

Atomic Linear Relations

Definition (Frozen variables)

Let $A \in \mathbb{F}^{m \times n}$.
If $I = \{i\}$ is an e-relation of A, then we call i **frozen** in A. We denote the set of all frozen variables by $\mathcal{F}(A)$.

Observation 1: $i \in \mathcal{F}(A) \iff y_i = 0$ for all $y \in \ker_{\mathbb{F}}(A)$.

Atomic Linear Relations

Definition (Frozen variables)

Let $A \in \mathbb{F}^{m \times n}$.
If $I = \{i\}$ is an e-relation of A, then we call i **frozen** in A. We denote the set of all frozen variables by $\mathcal{F}(A)$.

Observation 1: $i \in \mathcal{F}(A) \iff y_i = 0$ for all $y \in \ker_{\mathbb{F}}(A)$.

"$\Longrightarrow$" If $i \in \mathcal{F}(A)$, the equation $x_i = 0$ can be obtained through elementary row operations.

Atomic Linear Relations

Definition (Frozen variables)

Let $A \in \mathbb{F}^{m \times n}$.
If $I = \{i\}$ is an e-relation of A, then we call i **frozen** in A. We denote the set of all frozen variables by $\mathcal{F}(A)$.

Observation 1: $i \in \mathcal{F}(A) \iff y_i = 0$ for all $y \in \ker_{\mathbb{F}}(A)$.

"$\Longrightarrow$" If $i \in \mathcal{F}(A)$, the equation $x_i = 0$ can be obtained through elementary row operations.

"$\Longleftarrow$" If $y_i = 0$ for all $y \in \ker_{\mathbb{F}}(A)$, then adding the equation $x_i = 0$ does not change the set of solutions to the homogeneous system, so it can be obtained through elementary row operations.

More kinds of frozen variables

Definition

For any matrix $A \in \mathbb{F}^{m \times n}$ and $1 \leq i \leq m \wedge n$, we say that

1. i is **firmly frozen** in A, if i is frozen in the matrix that arises from A through deletion of row i.

More kinds of frozen variables

Definition

For any matrix $A \in \mathbb{F}^{m \times n}$ and $1 \leq i \leq m \wedge n$, we say that

1. i is **firmly frozen** in A, if i is frozen in the matrix that arises from A through deletion of row i.

More kinds of frozen variables

Definition

For any matrix $A \in \mathbb{F}^{m \times n}$ and $1 \leq i \leq m \wedge n$, we say that

1. i is **firmly frozen** in A, if i is frozen in the matrix that arises from A through deletion of row i.
2. i is **completely frozen** in A, if i is firmly frozen both in A and A^T.

More kinds of frozen variables

Definition

For any matrix $A \in \mathbb{F}^{m \times n}$ and $1 \leq i \leq m \wedge n$, we say that

1. i is **firmly frozen** in A, if i is frozen in the matrix that arises from A through deletion of row i.
2. i is **completely frozen** in A, if i is firmly frozen both in A and A^T.

Observation: Variables that are firmly frozen in A are frozen in A.

Rank and symmetric row and column removal

With the previous definitions, we have

$$\mathrm{rk}(A) - \mathrm{rk}(A\langle i; i\rangle) =$$

Rank and symmetric row and column removal

With the previous definitions, we have

$$\mathrm{rk}(A) - \mathrm{rk}(A\langle i; i\rangle) = 1$$

Rank and symmetric row and column removal

With the previous definitions, we have

$$\mathrm{rk}(A) - \mathrm{rk}(A\langle i; i\rangle) = 1 + \mathbf{1}\{i \text{ is completely frozen in } A\}$$

Rank and symmetric row and column removal

With the previous definitions, we have

$$\mathrm{rk}(A) - \mathrm{rk}(A\langle i; i\rangle) = 1 + \mathbf{1}\left\{i \text{ is completely frozen in } A\right\} - \mathbf{1}\left\{i \text{ is neither frozen in } A \text{ nor in } A^t\right\}.$$

Rank and symmetric row and column removal

With the previous definitions, we have

$$\mathrm{rk}(A) - \mathrm{rk}(A\langle i; i\rangle) = \quad + \mathbf{1}\{i \text{ is completely frozen in } A\}$$
$$- \mathbf{1}\left\{i \text{ is neither frozen in } A \text{ nor in } A^t\right\}.$$

$\Longrightarrow$ Next step: Get a handle on completely frozen and "nowhere frozen" variables.

Even more kinds of frozen variables

The notions of completely frozen and “nowhere frozen” variables are not exhaustive.
We next introduce a **partition of the coordinates** into $p+3$ different types of variables.

Even more kinds of frozen variables

The notions of completely frozen and "nowhere frozen" variables are not exhaustive.
We next introduce a **partition of the coordinates** into $p+3$ different types of variables.

Definition

For any matrix $A \in \mathbb{F}^{m \times n}$ and $1 \leq i \leq m \wedge n$, we say that i is **frailly frozen** in A if i is frozen in A, but not in the matrix that arises from A through deletion of row i.

Even more kinds of frozen variables

The notions of completely frozen and “nowhere frozen” variables are not exhaustive.
We next introduce a **partition of the coordinates** into $p+3$ different types of variables.

Definition

For any matrix $A \in \mathbb{F}^{m\times n}$ and $1 \le i \le m \wedge n$, we say that i is **frailly frozen** in A if i is frozen in A, but not in the matrix that arises from A through deletion of row i.

Observation:

Proposition

Let $A \in \mathbb{F}^{m\times n}$ and $i \in [m \wedge n]$. Then

$$i \text{ is frailly frozen in } A \iff i \text{ is frailly frozen in } A^T.$$

Signature

Observation: If $\{i\}$ is either frailly frozen or completely frozen in A, then for any two representations $y = (y_1, \ldots, y_m)$, $y' = (y'_1, \ldots, y'_m)$ of $\{i\}$ in A,

$$y_i = y'_i.$$

Signature

Observation: If $\{i\}$ is either fraily frozen or completely frozen in A, then for any two representations $y = (y_1, \ldots, y_m)$, $y' = (y'_1, \ldots, y'_m)$ of $\{i\}$ in A,

$$y_i = y'_i.$$

Definition (Signature)

Assume that $i \in [m \wedge n]$ is frailly frozen or completely frozen in $A \in \mathbb{F}^{m \times n}$. We call the **unique** value y_i in any representation y of i in A the signature of i in A and denote it by $\mathrm{sig}(A, i)$.

Signature

Observation: If $\{i\}$ is either frailly frozen or completely frozen in A, then for any two representations $y = (y_1, \ldots, y_m)$, $y' = (y'_1, \ldots, y'_m)$ of $\{i\}$ in A,

$$y_i = y'_i.$$

Definition (Signature)

Assume that $i \in [m \wedge n]$ is frailly frozen or completely frozen in $A \in \mathbb{F}^{m \times n}$. We call the **unique** value y_i in any representation y of i in A the signature of i in A and denote it by $\mathrm{sig}(A, i)$.

Remark: If $\{i\}$ is completely frozen in A, then

$$\mathrm{sig}(A, i) = 0.$$

Partition of the set of variables

For a symmetric matrix $A \in \mathbb{F}^{n \times n}$, we thus have the following partition of $[n]$:

Let $i \in [n]$. Then i belongs to exactly one of the following sets:

- The set of variables that are **frailly frozen** in A **with signature** $j \in [p-1]$.
- The set of variables that are **completely frozen** in A.
- The set of variables that are **neither frozen** in A nor A^t.
- The set of variables that are **not frozen in** A**,** but **firmly frozen in** A^t.
- The set of variables that are **firmly frozen in** A, but **not frozen in** A^t.

Fixed point equations for the proportions

Now as $n \to \infty$, what are the various proportions of variables in $\boldsymbol{A}_{n,p}$?

Fixed point equations for the proportions

Now as $n \to \infty$, what are the various proportions of variables in $\boldsymbol{A}_{n,p}$?

To characterise them, we derive equations on the change of the various proportions under the operation of symmetrically attaching a row and a column to a general matrix A:

$$A \qquad \longrightarrow \qquad A^{\boldsymbol{b}} = \begin{pmatrix} A & \boldsymbol{b}^t \\ \boldsymbol{b} & 0 \end{pmatrix}.$$

Fixed point equations for the proportions

Now as $n \to \infty$, what are the various proportions of variables in $\boldsymbol{A}_{n,p}$?

To characterise them, we derive equations on the change of the various proportions under the operation of symmetrically attaching a row and a column to a general matrix A:

$$A \qquad \longrightarrow \qquad A^{\boldsymbol{b}} = \begin{pmatrix} A & \boldsymbol{b}^t \\ \boldsymbol{b} & 0 \end{pmatrix}.$$

More precisely, we derive the corresponding equations for a perturbed version of a general matrix A that eliminates most short linear relations.

Symmetric perturbation

Idea:

Apply a random perturbation to A that does not change the rank by a significant amount, but eliminates most of the short linear relations.

Symmetric perturbation

Idea:

Apply a random perturbation to A that does not change the rank by a significant amount, but eliminates most of the short linear relations.

$$A[3,4] = \left(\begin{array}{cccccccccc|cccc}
 & & & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 0 & 1 & 0 & 0 \\
 & & & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 0 & 0 & 0 & 1 \\
 & & & & A & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 1 & 0 & 0 & 0 \\
 & & & & & & & & & & 0 & 0 & 1 & 0 \\
 & & & & & & & & & & 0 & 0 & 0 & 0 \\
 & & & & & & & & & & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right)$$

Extended perturbation

Actually, we derive equations on the change of the various proportions under the operation of symmetrically attaching a row and a column to $A[\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]$:

$$A[\boldsymbol{\theta}_1, \boldsymbol{\theta}_2] \qquad \longrightarrow \qquad A[\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]^{\boldsymbol{b}} = \begin{pmatrix} A[\boldsymbol{\theta}_1, \boldsymbol{\theta}_2] & \boldsymbol{b}^t \\ \boldsymbol{b} & 0 \end{pmatrix}.$$

Narrowing down the possible scenarios

For large n, one of the following two scenarios will be true in the perturbed matrix:

Narrowing down the possible scenarios

For large n, one of the following two scenarios will be true in the perturbed matrix:

Case 1:

- The proportions of frailly frozen as well as "one-sided" firmly frozen variables are negligible.
- The proportion of frozen variables approximately satisfies the equation $\alpha = 1 - \phi(1 - \phi(\alpha))$.

Narrowing down the possible scenarios

For large n, one of the following two scenarios will be true in the perturbed matrix:

Case 1:

- The proportions of frailly frozen as well as "one-sided" firmly frozen variables are negligible.
- The proportion of frozen variables approximately satisfies the equation $\alpha = 1 - \phi(1 - \phi(\alpha))$.

Case 2:

- The proportion of frailly frozen variables is not negligible, but a specific function of the proportion of frozen variables.
- The proportion of completely frozen variables is not negligible, but a specific function of the proportion of frozen variables.
- The proportion of frozen variables approximately satisfies the equation

$$\alpha = 1 - \frac{1}{p}\phi(\alpha) - \frac{p-1}{p}\phi\left(\frac{p}{p-1}\left(1 - \phi\left(\alpha\right)\right) - \frac{1}{p-1}\alpha\right).$$

Lower bound on the rank

With this a priori estimate, it is possible to show that in any case, the desired lower bound holds:

Theorem

For any $d > 0$ and any prime p,

$$\liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\,\mathrm{rk}_{\mathbb{F}_p}\left(\boldsymbol{A}_{n,p}\right)\right] \geq 2 - \max_{\alpha\in[0,1]} M_d(\alpha).$$

Future directions

- Extension to general entries from $\mathbb{F}_p$
- Extension to a broader class of symmetric matrices
- Rank of the adjacency matrix of the $k-$core of $\boldsymbol{G}_{n,p}$ over finite fields ($k \geq 3$)