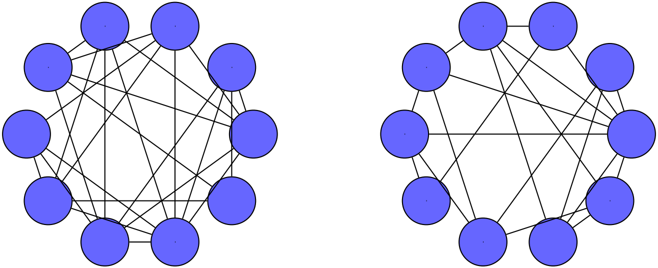
S&DS 684 Lecture 12: Random Graph Matching: Information-theoretic Limits

Yihong Wu

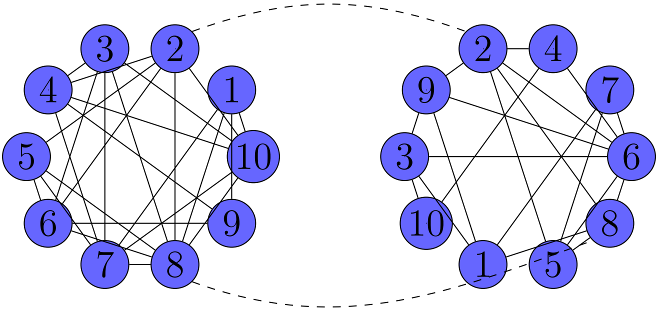
Department of Statistics and Data Science
Yale University

Apr 18, 2023

Graph matching (network alignment)



Graph matching (network alignment)



Goal: find a mapping between two node sets that maximally aligns the edges (i.e. minimizes # of adjacency disagreements)

QAP (1)

Given symmetric $n \times n$ matrices A, B, solve

$$\text{Quadratic Assignment Problem (QAP)}: \quad \max_{\pi \in S_n} \sum_{i<j} A_{\pi(i)\pi(j)} B_{ij}$$

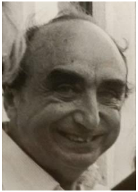
- Introduced by Koopmans-Beckmann '57 (Yale Econon)

COWLES FOUNDATION DISCUSSION PAPER, NO. 4*

<u>Assignment Problems and the Location of Economic Activities**</u>

by

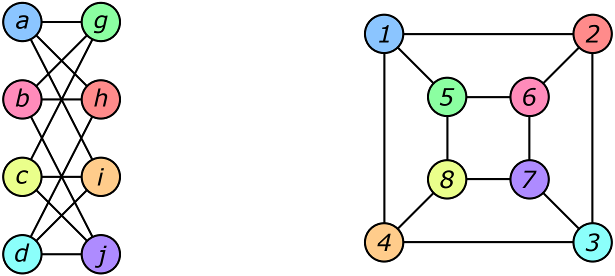
Tjalling C. Koopmans and Martin Beckmann

QAP (2)

Noiseless case: QAP $\iff$ **Graph isomorphism**
Given two graphs A and B, decide whether $A \cong B$, i.e., there exists a bijection $\pi : V(A) \to V(B)$ such that

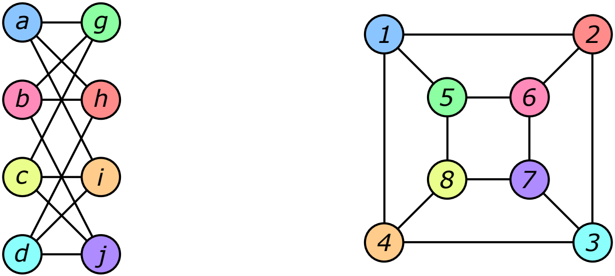
$$(u, v) \in E(A) \Leftrightarrow (\pi(u), \pi(v)) \in E(B)$$



QAP (2)

Noiseless case: QAP $\iff$ **Graph isomorphism**
Given two graphs A and B, decide whether $A \cong B$, i.e., there exists a bijection $\pi : V(A) \to V(B)$ such that

$$(u, v) \in E(A) \Leftrightarrow (\pi(u), \pi(v)) \in E(B)$$



- Not known to be solvable in polynomial time in the worst case
- In practice, two graphs are often not exactly isomorphic, but still want to tell whether their topologies are similar

QAP (3)

QAP includes many problems as special cases: $A =$ adj matrix of observed graph

- Planted clique (Part I):

$$B = \text{ adj matrix of a fixed } k\text{-clique}$$

- Minumum bisection (Part II):

$$B = \xi\xi^\top, \qquad \xi = (1, \dots, 1, -1, \dots, -1)^\top.$$

- TSP (Lec 12):

$$B = \text{ adj matrix of a fixed Hamiltonian cycle}$$

QAP (3)

QAP includes many problems as special cases: $A =$ adj matrix of observed graph

- Planted clique (Part I):

$$B = \text{ adj matrix of a fixed } k\text{-clique}$$

- Minumum bisection (Part II):

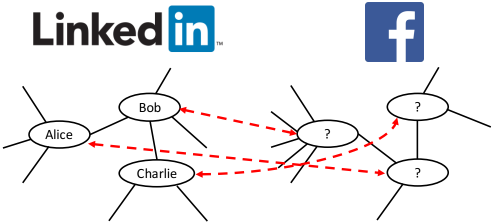
$$B = \xi\xi^\top, \qquad \xi = (1, \ldots, 1, -1, \ldots, -1)^\top.$$

- TSP (Lec 12):

$$B = \text{ adj matrix of a fixed Hamiltonian cycle}$$

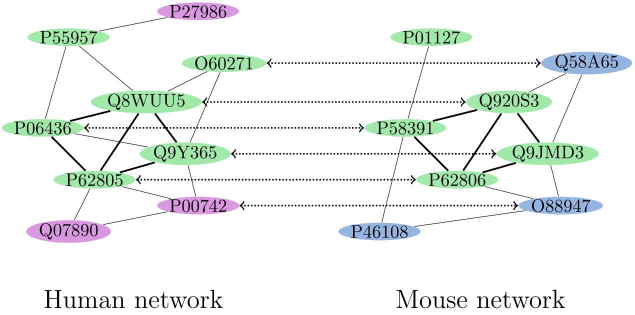
Here we will be dealing with B being Erdős-Rényi as well.

Application 1: Network de-anonymization



- Successfully de-anonymize Netflix dataset by matching it to IMDB [Narayanan-Shmatikov '08]
- Correctly identify 30.8% of shared users between Twitter and Flickr [Narayanan-Shmatikov '09]

Application 2: Protein-Protein Interaction network

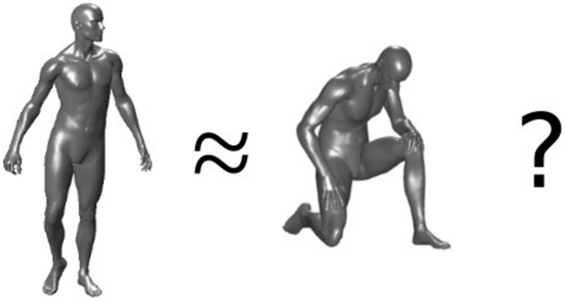


[Kazemi-Hassani-Grossglauser-Modarres '16]

Graph matching for aligning PPI networks between different species, to identify conserved components and genes with common function [Singh-Xu-Berger '08]

Application 3: Computer vision

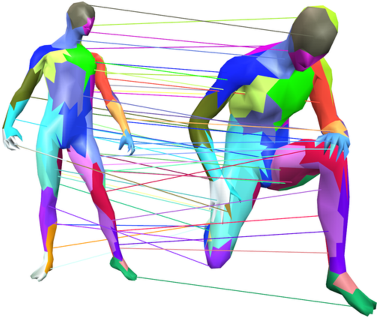
A fundamental problem in computer vision: Detect and match similar objects that undergo different deformations



Shape REtrieval Contest (SHREC) dataset [Lähner et al '16]

Application 3: Computer vision

A fundamental problem in computer vision: Detect and match similar objects that undergo different deformations

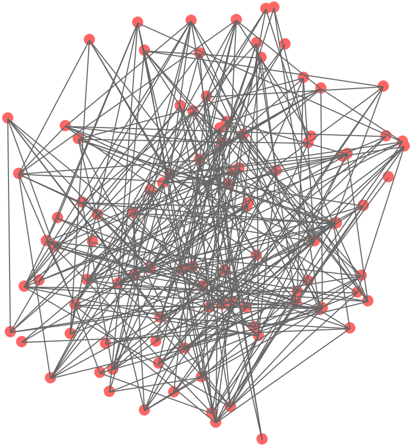


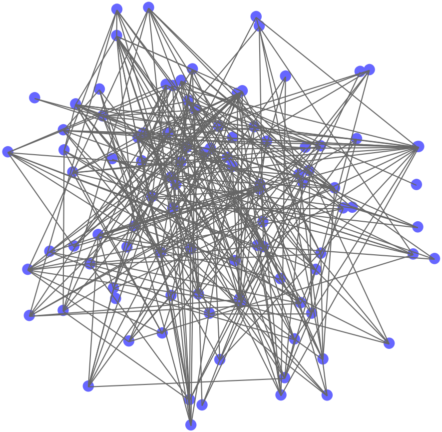
Shape REtrieval Contest (SHREC) dataset [Lähner et al '16]

3-D shapes → geometric graphs (features → nodes, distances → edges)

Two key challenges

- Statistical: two graphs may not be the same
- Computational: # of possible node mappings is $n!$ ($100! \approx 10^{158}$)





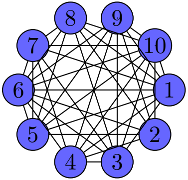
Beyond worst-case intractability

- NP-hard for matching two graphs in worst case
  - QAP is hard to approximate within $\exp(\text{polylog}(n))$ multiplicative factor [Makarychev-Manokaran-Sviridenko '15]
- However, real networks are not arbitrary and have latent structures

Beyond worst-case intractability

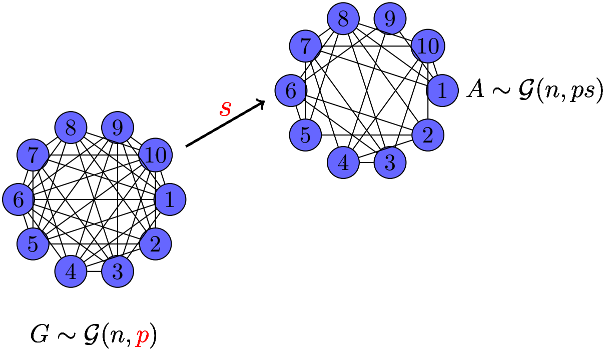
- NP-hard for matching two graphs in worst case
  - QAP is hard to approximate within $\exp(\text{polylog}(n))$ multiplicative factor [Makarychev-Manokaran-Sviridenko '15]
- However, real networks are not arbitrary and have latent structures
- Recent surge of interests on the average-case analysis of matching correlated random graphs [Feizi at el.'16, Lyzinski at el'16, Cullina-Kiyavash'16,17, Ding-Ma-W-Xu'18, Barak-Chou-Lei-Schramm-Sheng'19, Fan-Mao-W-Xu'19a,19b, Ganassali-Massoulié'20, Hall-Massoulié'20, ...]
  - CS-style average-case analysis: under null model, aiming to understand "what's the fraction of bad instances"
  - Stat-style average-case analysis: under planted model (meaningful statistical model).
- Focus on correlated Erdős-Rényi graphs model [Pedarsani-Grossglauser '11]

Correlated Erdős-Rényi graphs model $\mathcal{G}(n,p,s)$

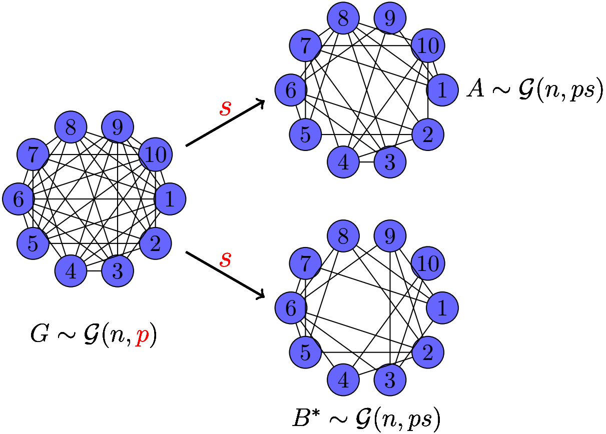


$G \sim \mathcal{G}(n, p)$

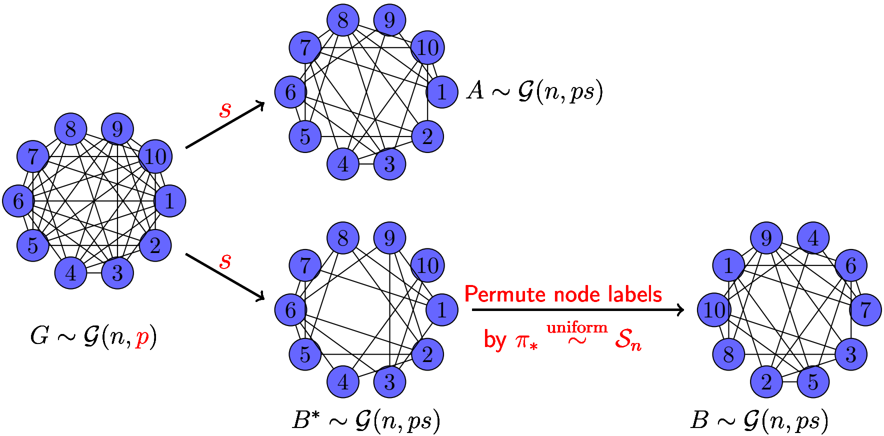
Correlated Erdős-Rényi graphs model $\mathcal{G}(n, p, s)$



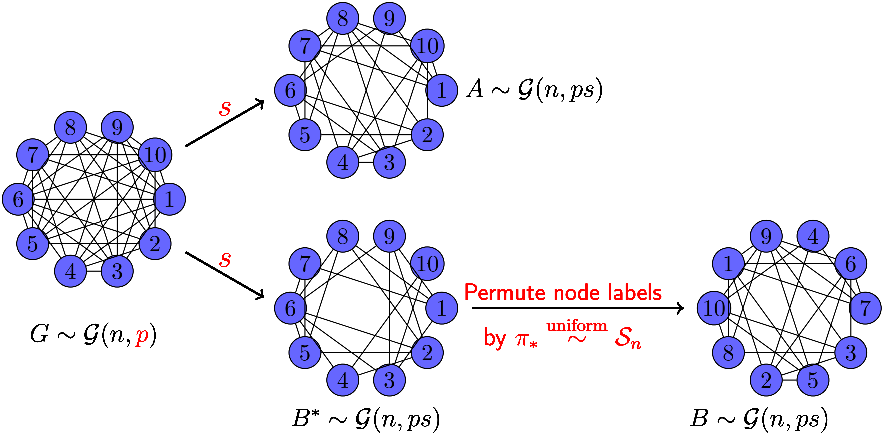
Correlated Erdős-Rényi graphs model $\mathcal{G}(n, p, s)$



Correlated Erdős-Rényi graphs model $\mathcal{G}(n, p, s)$



Correlated Erdős-Rényi graphs model $\mathcal{G}(n,p,s)$



- $(A_{\pi_*(i)\pi_*(j)}, B_{ij})$ are iid pairs of correlated $\mathrm{Bern}(ps)$
- Key parameter nps^2: average degree of intersection graph $A \wedge B^*$;

Correlated Gaussian model

$$B = \rho A^{\pi_*} + \sqrt{1-\rho^2} Z \,,$$

where

- A and Z are independent Gaussian Wigner matrices with iid standard normal entries;
- $A^{\pi_*} = (A_{\pi_*(i)\pi_*(j)})$ denotes the relabeled version of A
- Conditional on π_*, for any $1 \le i < j \le n$,

$$(A_{\pi_*(i)\pi_*(j)}, B_{ij}) \overset{\text{iid}}{\sim} \mathcal{N}\left(\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right), \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right).$$

Two statistical tasks: detection and estimation

- Detection:
  - $\mathcal{H}_0$: A and B are independent Erdős-Rényi graphs $\mathcal{G}(n, ps)$
  - $\mathcal{H}_1$: A and B are correlated Erdős-Rényi graphs $\mathcal{G}(n, p, s)$
  - Test between $\mathcal{H}_0$ and $\mathcal{H}_1$ based on observation of (A, B)

- Estimation:
  - Observe two correlated Erdős-Rényi graphs $A, B \sim \mathcal{G}(n, p, s)$
  - Recover the underlying true vertex correspondence π_*

(Similarly for Gaussian model)

Two statistical tasks: detection and estimation

- Detection:
  - $\mathcal{H}_0$: A and B are independent Erdős-Rényi graphs $\mathcal{G}(n, ps)$
  - $\mathcal{H}_1$: A and B are correlated Erdős-Rényi graphs $\mathcal{G}(n, p, s)$
  - Test between $\mathcal{H}_0$ and $\mathcal{H}_1$ based on observation of (A, B)

- Estimation:
  - Observe two correlated Erdős-Rényi graphs $A, B \sim \mathcal{G}(n, p, s)$
  - Recover the underlying true vertex correspondence π_*

  (Similarly for Gaussian model)

Focus of this lecture

What are the information-theoretic limits of detection and estimation?
(Next Tuesday: Algorithms.)

Two statistical tasks: detection and estimation

- Detection:
  - $\mathcal{H}_0$: A and B are independent Erdős-Rényi graphs $\mathcal{G}(n, ps)$
  - $\mathcal{H}_1$: A and B are correlated Erdős-Rényi graphs $\mathcal{G}(n, p, s)$
  - Test between $\mathcal{H}_0$ and $\mathcal{H}_1$ based on observation of (A, B)
- Estimation:
  - Observe two correlated Erdős-Rényi graphs $A, B \sim \mathcal{G}(n, p, s)$
  - Recover the underlying true vertex correspondence π_*

  (Similarly for Gaussian model)

Focus of this lecture

What are the information-theoretic limits of detection and estimation? (Next Tuesday: Algorithms.)

Progress in the recent decade: [Pedarsani-Grossglauser '11], [Cullina-Kiyavash '16,17], [Hall-Massoulié '20], [Ganassali '20], [W-Xu-Yu '20,21], [Ganassali-Lelarge-Massoulié '21], [Ding-Du '21 22]

Maximum likelihood estimation as quadratic assignment

Maximum likelihood estimation reduces to quadratic assignment (QAP):

$$\widehat{\pi}_{\mathsf{ML}} \in \arg\max_{\pi} \sum_{i<j} A_{\pi(i)\pi(j)} B_{ij} \, .$$

- QAP is NP-hard in worst case
- How much does $\widehat{\pi}_{\mathrm{ML}}$ have in common with π^*?

$$\mathrm{overlap}(\pi_*, \widehat{\pi}) \triangleq \frac{1}{n} \Big| \{ i \in [n] : \widehat{\pi}(i) = \pi_*(i) \} \Big|$$

i.e., fraction of correctly classified nodes

Sharp threshold for detection: Gaussian

Theorem (W-Xu-Yu '20)

$$n\rho^2 \geq (4+\epsilon)\log n \implies \mathrm{TV}(\mathcal{P}, \mathcal{Q}) = 1 - o(1)\,(\textit{test error}{=}o(1))$$

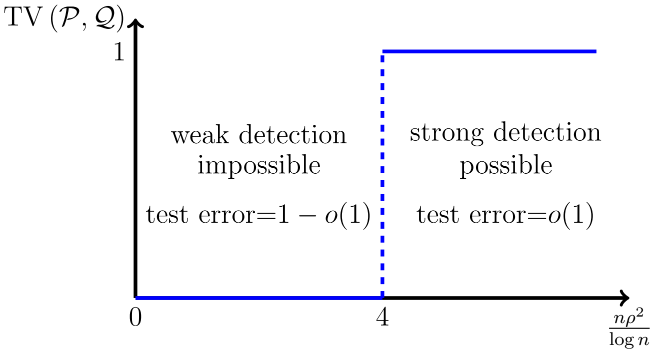
$$n\rho^2 \leq (4-\epsilon)\log n \implies \mathrm{TV}(\mathcal{P}, \mathcal{Q}) = o(1)\,(\textit{test error}{=}1 - o(1))$$

Sharp threshold for detection: Gaussian

Theorem (W-Xu-Yu '20)

$$n\rho^2 \geq (4+\epsilon)\log n \implies \mathrm{TV}(\mathcal{P},\mathcal{Q}) = 1-o(1)\ (\textit{test error}{=}o(1))$$

$$n\rho^2 \leq (4-\epsilon)\log n \implies \mathrm{TV}(\mathcal{P},\mathcal{Q}) = o(1)\ (\textit{test error}{=}1-o(1))$$

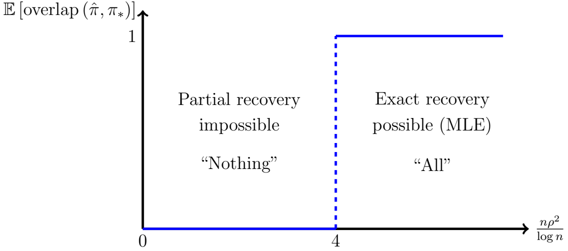


Sharp threshold for recovery: Gaussian model

Theorem (W-Xu-Yu '21)

$$n\rho^2 \geq (4+\epsilon)\log n \implies \widehat{\pi}_{\mathsf{ML}} = \pi_* \ \textit{whp}$$

$$n\rho^2 \leq (4-\epsilon)\log n \implies \text{overlap}\,(\widehat{\pi}, \pi_*) = o(1), \ \textit{whp}, \forall \ \textit{estimator}\ \widehat{\pi}$$

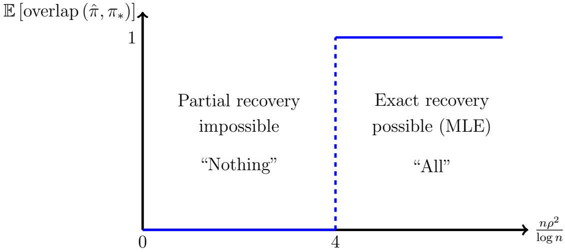


Sharp threshold for recovery: Gaussian model

Theorem (W-Xu-Yu '21)

$$n\rho^2 \geq (4+\epsilon)\log n \implies \widehat{\pi}_{\mathsf{ML}} = \pi_* \ \textit{whp}$$

$$n\rho^2 \leq (4-\epsilon)\log n \implies \mathrm{overlap}(\widehat{\pi}, \pi_*) = o(1), \ \textit{whp}, \forall \textit{ estimator } \widehat{\pi}$$



- Exact recovery threshold is derived in [Ganassali '20]
- Exhibits a stronger form of "all or nothing" phenomenon
- Only a vanishing amount of correlation allows detection and recovery

Sharp detection threshold: dense Erdős-Rényi graphs

Theorem (W-Xu-Yu '20)

Suppose $n^{-o(1)} \leq p \leq 1 - \Omega(1)$. *Then,*

$$nps^2 \geq \frac{(2+\epsilon)\log n}{\log\frac{1}{p} - 1 + p} \implies \mathrm{TV}\left(\mathcal{P}, \mathcal{Q}\right) = 1 - o\left(1\right)\ (\textit{test error}{=}o(1))$$

$$nps^2 \leq \frac{(2-\epsilon)\log n}{\log\frac{1}{p} - 1 + p} \implies \mathrm{TV}\left(\mathcal{P}, \mathcal{Q}\right) = o(1)\,(\textit{test error}{=}1 - o(1))$$

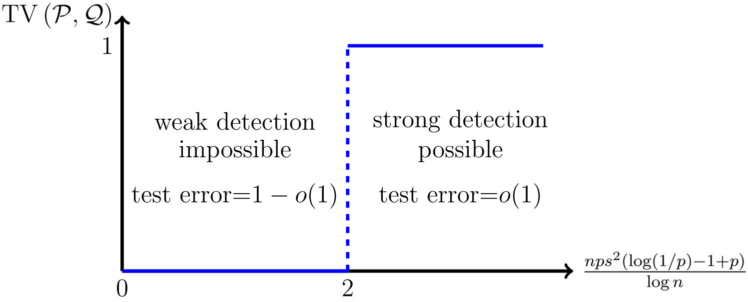
Sharp detection threshold: dense Erdős-Rényi graphs

Theorem (W-Xu-Yu '20)

Suppose $n^{-o(1)} \le p \le 1 - \Omega(1)$. *Then,*

$$nps^2 \ge \frac{(2+\epsilon)\log n}{\log \frac{1}{p} - 1 + p} \implies \mathrm{TV}(\mathcal{P}, \mathcal{Q}) = 1 - o(1) \ (\textit{test error}=o(1))$$

$$nps^2 \le \frac{(2-\epsilon)\log n}{\log \frac{1}{p} - 1 + p} \implies \mathrm{TV}(\mathcal{P}, \mathcal{Q}) = o(1) \ (\textit{test error}=1-o(1))$$



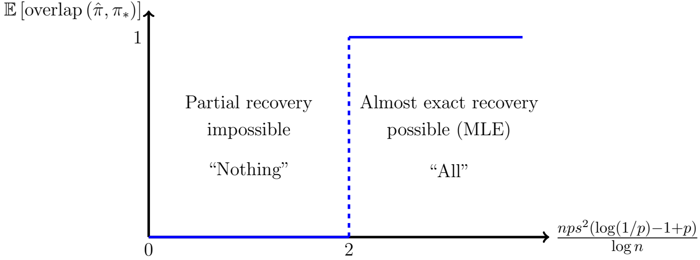
Sharp recovery threshold: dense Erdős-Rényi

Theorem (W-Xu-Yu '21)

Suppose $n^{-o(1)} \le p \le 1 - \Omega(1)$. *Then,*

$$nps^2 \ge \frac{(2+\epsilon)\log n}{\log\frac{1}{p} - 1 + p} \implies \mathrm{overlap}\left(\widehat{\pi}_{\mathrm{ML}}, \pi_*\right) = 1 - o(1) \text{ whp}$$

$$nps^2 \le \frac{(2-\epsilon)\log n}{\log\frac{1}{p} - 1 + p} \implies \mathrm{overlap}\left(\widehat{\pi}, \pi_*\right) = o(1), \text{ whp, } \forall \text{ estimator } \widehat{\pi}$$



Sharp recovery threshold: dense Erdős-Rényi

Theorem (W-Xu-Yu '21)

Suppose $n^{-o(1)} \le p \le 1 - \Omega(1)$. *Then,*

$$nps^2 \ge \frac{(2+\epsilon)\log n}{\log\frac{1}{p} - 1 + p} \implies \text{overlap}\left(\widehat{\pi}_{\text{ML}}, \pi_*\right) = 1 - o(1) \text{ whp}$$

$$nps^2 \le \frac{(2-\epsilon)\log n}{\log\frac{1}{p} - 1 + p} \implies \text{overlap}\left(\widehat{\pi}, \pi_*\right) = o(1), \text{ whp, } \forall \text{ estimator } \widehat{\pi}$$

Interpretation of threshold:

- $I(\pi_*; A, B) \approx \binom{n}{2} \times \underbrace{ps^2\left(\log\frac{1}{p} - 1 + p\right)}_{\text{mutual info btw two correlated edges}}$
- $H(\pi_*) \approx n \log n$
- Threshold is at $I(\pi; A, B) \approx H(\pi_*)$
- Only a vanishing amount of correlation allows detection and recovery

Sharp detection threshold: sparse Erdős-Rényi

Theorem (Ding-Du '22a)

Suppose $p = n^{-\alpha}$ *for* $\alpha \in (0,1)$ *and* $\lambda^* = \gamma^{-1}(1/\alpha)$.

$$nps^2 \geq \lambda^* + \epsilon \implies \mathrm{TV}(\mathcal{P}, \mathcal{Q}) = 1 - o(1) \, (\textit{test error}{=}o(1))$$
$$nps^2 \leq \lambda^* - \epsilon \implies \mathrm{TV}(\mathcal{P}, \mathcal{Q}) = o(1) \, (\textit{test error}{=}1 - o(1))$$

- Sharpens the earlier threshold of $nps^2 = \Theta(1)$ [W-Xu-Yu '20]
- $\gamma : [1, \infty) \to [1, \infty)$ is given by the densest subgraph problem in Erdős-Rényi $\mathcal{G}(n, \frac{\lambda}{n})$ [Hajek '90, Anantharam-Salez' 16]

$$\max_{\emptyset \neq U \subset [n]} \frac{|\mathcal{E}(U)|}{|U|} \to \gamma(\lambda)$$

- When $np = \Theta(1)$, there is no zero-one phase transition.

Sharp recovery threshold: sparse Erdős-Rényi

Theorem (Ding-Du '22b)

Suppose $p = n^{-\alpha}$ *for* $\alpha \in (0,1]$ *and* $\lambda^* = \gamma^{-1}(1/\alpha)$.

$$nps^2 \geq \lambda^* + \epsilon \implies \mathrm{overlap}\left(\widehat{\pi}_{\mathrm{ML}}, \pi_*\right) \geq \Omega(1) \text{ whp.}$$
$$nps^2 \leq \lambda^* - \epsilon \implies \mathrm{overlap}\left(\widehat{\pi}, \pi_*\right) = o(1) \text{ whp. } \forall \widehat{\pi}$$

- The case of $\alpha = 1$ is proved in [Ganassali-Lelarge-Massoulié '21]
- Sharpen the partial recovery threshold at $nps^2 = \Theta(1)$ [W-Xu-Yu '20]
- "All-or-nothing" phenomenon does not exist, as almost exact recovery $(\mathrm{overlap} = 1 - o(1))$ requires $nps^2 \to \infty$ [Cullina-Kiyavash-Mittal-Poor '19]

Exact recovery threshold

Theorem (W-Xu-Yu '21)

Suppose $p \le 1 - \Omega(1)$. *Then*

$$nps^2 \ge \frac{(1+\epsilon)\log n}{\left(1-\sqrt{p}\right)^2} \implies \mathrm{overlap}\left(\widehat{\pi}_{\mathrm{ML}}, \pi_*\right) = 1 \textit{ whp.}$$

$$nps^2 \le \frac{(1-\epsilon)\log n}{\left(1-\sqrt{p}\right)^2} \implies \mathrm{overlap}\left(\widehat{\pi}, \pi_*\right) \neq 1 \textit{ whp. } \forall\, \widehat{\pi}\,.$$

Exact recovery threshold

Theorem (W-Xu-Yu '21)

Suppose $p \le 1 - \Omega(1)$. *Then*

$$nps^2 \ge \frac{(1+\epsilon)\log n}{\left(1-\sqrt{p}\right)^2} \implies \text{overlap}\left(\widehat{\pi}_{\mathrm{ML}}, \pi_*\right) = 1 \textit{ whp.}$$

$$nps^2 \le \frac{(1-\epsilon)\log n}{\left(1-\sqrt{p}\right)^2} \implies \text{overlap}\left(\widehat{\pi}, \pi_*\right) \ne 1 \textit{ whp. } \forall\, \widehat{\pi}\,.$$

- $p = o(1)$: reduces to the connectivity threshold of the intersection graph $A \wedge B^* \sim \mathcal{G}(n, ps^2)$ [Cullina-Kiyavash'16,17].

  Fact about Erdős-Rényi graph: For $G \sim \mathcal{G}(n, q)$,
  - If $q \ge \frac{(1+\epsilon)\log n}{n}$, G is connected.
  - If $q \le \frac{(1-\epsilon)\log n}{n}$, G has many isolated vertices.

Exact recovery threshold

Theorem (W-Xu-Yu '21)

Suppose $p \leq 1 - \Omega(1)$. *Then*

$$nps^2 \geq \frac{(1+\epsilon)\log n}{\left(1-\sqrt{p}\right)^2} \implies \text{overlap}\left(\widehat{\pi}_{\mathrm{ML}}, \pi_*\right) = 1 \textit{ whp.}$$

$$nps^2 \leq \frac{(1-\epsilon)\log n}{\left(1-\sqrt{p}\right)^2} \implies \text{overlap}\left(\widehat{\pi}, \pi_*\right) \neq 1 \textit{ whp. } \forall\, \widehat{\pi}\,.$$

- $p = o(1)$: reduces to the connectivity threshold of the intersection graph $A \wedge B^* \sim \mathcal{G}(n, ps^2)$ [Cullina-Kiyavash'16,17].

  Fact about Erdős-Rényi graph: For $G \sim \mathcal{G}(n, q)$,
  - If $q \geq \frac{(1+\epsilon)\log n}{n}$, G is connected.
  - If $q \leq \frac{(1-\epsilon)\log n}{n}$, G has many isolated vertices.
- $p = \Omega(1)$: strictly higher than the connectivity threshold

Analysis

- Proof of detection thresholds
- Proof of exact recovery thresholds

Proof of detection thresholds: Positive results

- Gaussian or dense Erdős-Rényi: analyzing QAP statistic

$$T_{\mathsf{QAP}} = \max_{\pi \in \mathcal{S}_n} \sum_{i<j} A_{\pi(i)\pi(j)} B_{ij}$$

  In Erdős-Rényi model: $T_{\mathsf{QAP}} =$ size of maximal common subgraph
- Analysis: standard first-moment computation (next page)

Proof of detection thresholds: Positive results

- Gaussian or dense Erdős-Rényi: analyzing QAP statistic

$$T_{\mathsf{QAP}} = \max_{\pi \in \mathcal{S}_n} \sum_{i<j} A_{\pi(i)\pi(j)} B_{ij}$$

  In Erdős-Rényi model: $T_{\mathsf{QAP}} =$ size of maximal common subgraph
- Analysis: standard first-moment computation (next page)
- Sparse Erdős-Rényi: analyzing densest subgraph statistic

$$\max_{\pi \in \mathcal{S}_n} \max_{U \subset [n]: |U| \geq n/\log n} \frac{\mathcal{E}_\pi(U)}{|U|},$$

  where $\mathcal{E}_\pi(U)$ is the set of edges induced by vertices in U in intersection graph $A^\pi \wedge B$

Proof of detection thresholds: Positive results

Gaussian analysis:

$$T_{\mathsf{QAP}} = \max_{\pi \in \mathcal{S}_n} \sum_{i<j} A_{\pi(i)\pi(j)} B_{ij}.$$

- Under $\mathcal{P}$ (ρ-correlated):

$$T_{\mathsf{QAP}} \geq \sum_{i<j} A_{\pi_*(i)\pi_*(j)} B_{ij} \approx \rho \binom{n}{2}$$

- Under $\mathcal{Q}$ (independent):

$$\mathcal{Q}\left(T_{\mathsf{QAP}} \leq \rho \binom{n}{2}\right) \lesssim n! \exp\left(-\frac{(\rho\binom{n}{2})^2}{2\binom{n}{2}}\right) \approx \exp\left(\rho^2 n^2/4 - n \log n\right)$$

- $\rho^2 = \frac{(4+\epsilon)\log n}{n} \implies$ success

Proof of detection thresholds: Negative results

Second-moment method (Chap 7):

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = O(1) \qquad \Longrightarrow \; \mathrm{TV}(\mathcal{P},\mathcal{Q}) \le 1-\Omega(1)$$

Strong detection is impossible

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = 1+o(1) \qquad \Longrightarrow \; \mathrm{TV}(\mathcal{P},\mathcal{Q}) = o(1)$$

Weak detection is impossible

Proof of detection thresholds: Negative results

Second-moment method (Chap 7):

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = O(1) \qquad \Longrightarrow \mathrm{TV}(\mathcal{P},\mathcal{Q}) \le 1-\Omega(1)$$

Strong detection is impossible

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = 1+o(1) \qquad \Longrightarrow \mathrm{TV}(\mathcal{P},\mathcal{Q}) = o(1)$$

Weak detection is impossible

Here

$$\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)} = \frac{1}{n!}\sum_{\pi_*\in S_n}\frac{\mathcal{P}(A,B|\pi_*)}{\mathcal{Q}(A,B)}.$$

As usual, second moment computation involves two iid replicas π_* and $\tilde{\pi}$

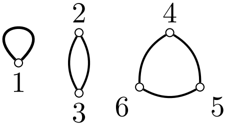
Cycle (orbit) decomposition

- **Node permutation** σ acts on $[n]$
- **Edge permutation** σ^{E} acts on $\binom{[n]}{2}$: $\sigma^{\mathsf{E}}((i,j)) \triangleq (\sigma(i), \sigma(j))$

Cycle (orbit) decomposition

- **Node permutation** σ acts on $[n]$
- **Edge permutation** σ^{E} acts on $\binom{[n]}{2}$: $\sigma^{\mathsf{E}}((i,j)) \triangleq (\sigma(i), \sigma(j))$

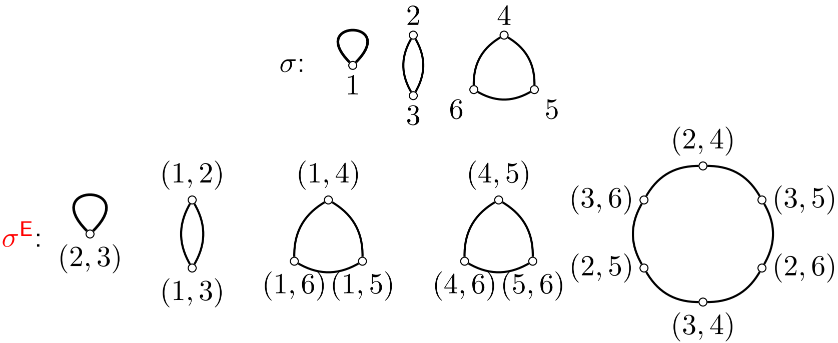
Example: $n = 6$ and $\sigma = (1)(23)(456)$:

σ: 

Cycle (orbit) decomposition

- **Node permutation** σ acts on $[n]$
- **Edge permutation** σ^{E} acts on $\binom{[n]}{2}$: $\sigma^{\mathsf{E}}((i,j)) \triangleq (\sigma(i), \sigma(j))$

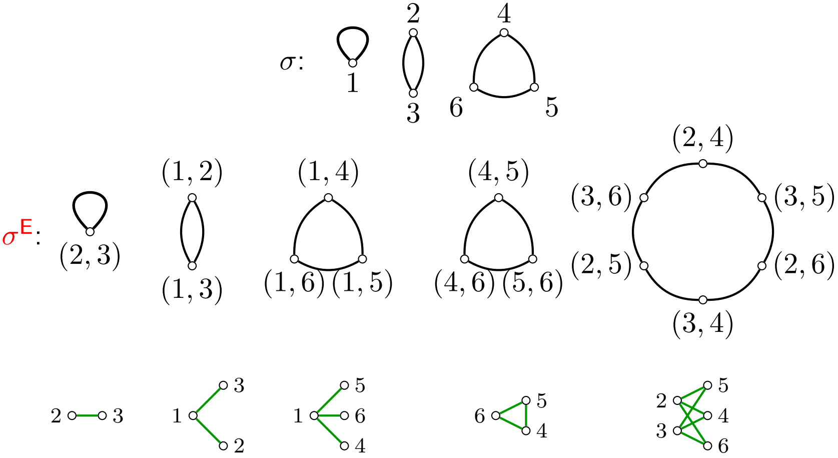
Example: $n = 6$ and $\sigma = (1)(23)(456)$:



Cycle (orbit) decomposition

- **Node permutation** σ acts on $[n]$
- **Edge permutation** σ^{E} acts on $\binom{[n]}{2}$: $\sigma^{\mathsf{E}}((i,j)) \triangleq (\sigma(i),\sigma(j))$

Example: $n = 6$ and $\sigma = (1)(23)(456)$:



Second moment via orbit decomposition (1)

$$\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2 = \left(\mathbb{E}_{\pi_*}\left[\frac{\mathcal{P}(A,B|\pi_*)}{\mathcal{Q}(A,B)}\right]\right)^2$$

$$= \mathbb{E}_{\widetilde{\pi} \perp\!\!\!\perp \pi_*} \prod_{i<j} X_{ij} \quad X_{ij} \triangleq \frac{\mathcal{P}(B_{ij}|A_{\pi_*(i)\pi_*(j)})\mathcal{P}(B_{ij}|A_{\widetilde{\pi}(i)\widetilde{\pi}(j)})}{\mathcal{Q}(B_{ij})^2}$$

$$= \mathbb{E}_{\widetilde{\pi} \perp\!\!\!\perp \pi_*} \prod_{O \in \mathcal{O}} X_O \quad X_O \triangleq \prod_{(i,j)\in O} X_{ij}$$

$\mathcal{O}$: disjoint orbits of edge permutation σ^{E} with $\sigma \triangleq \pi_*^{-1} \circ \widetilde{\pi}$

Second moment via orbit decomposition (1)

$$\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2 = \left(\mathbb{E}_{\pi_*}\left[\frac{\mathcal{P}(A,B|\pi_*)}{\mathcal{Q}(A,B)}\right]\right)^2$$

$$= \mathbb{E}_{\widetilde{\pi}\perp\!\!\!\perp\pi_*}\prod_{i<j} X_{ij} \quad X_{ij} \triangleq \frac{\mathcal{P}(B_{ij}|A_{\pi_*(i)\pi_*(j)})\mathcal{P}(B_{ij}|A_{\widetilde{\pi}(i)\widetilde{\pi}(j)})}{\mathcal{Q}(B_{ij})^2}$$

$$= \mathbb{E}_{\widetilde{\pi}\perp\!\!\!\perp\pi_*}\prod_{O\in\mathcal{O}} X_O \quad X_O \triangleq \prod_{(i,j)\in O} X_{ij}$$

$\mathcal{O}$: disjoint orbits of edge permutation σ^{E} with $\sigma \triangleq \pi_*^{-1}\circ\widetilde{\pi}$

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = \mathbb{E}_{\widetilde{\pi}\perp\!\!\!\perp\pi_*}\mathbb{E}_{\mathcal{Q}}\prod_{O\in\mathcal{O}} X_O = \mathbb{E}_{\widetilde{\pi}\perp\!\!\!\perp\pi_*}\prod_{O\in\mathcal{O}}\mathbb{E}_{\mathcal{Q}}\left[X_O\right]$$

Second moment via orbit decomposition (1)

$$\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2 = \left(\mathbb{E}_{\pi_*}\left[\frac{\mathcal{P}(A,B|\pi_*)}{\mathcal{Q}(A,B)}\right]\right)^2$$
$$= \mathbb{E}_{\widetilde{\pi} \perp\!\!\!\perp \pi_*} \prod_{i<j} X_{ij} \quad X_{ij} \triangleq \frac{\mathcal{P}(B_{ij}|A_{\pi_*(i)\pi_*(j)})\mathcal{P}(B_{ij}|A_{\widetilde{\pi}(i)\widetilde{\pi}(j)})}{\mathcal{Q}(B_{ij})^2}$$
$$= \mathbb{E}_{\widetilde{\pi} \perp\!\!\!\perp \pi_*} \prod_{O \in \mathcal{O}} X_O \quad X_O \triangleq \prod_{(i,j) \in O} X_{ij}$$

$\mathcal{O}$: disjoint orbits of edge permutation σ^{E} with $\sigma \triangleq \pi_*^{-1} \circ \widetilde{\pi}$

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = \mathbb{E}_{\widetilde{\pi} \perp\!\!\!\perp \pi_*}\mathbb{E}_{\mathcal{Q}} \prod_{O \in \mathcal{O}} X_O = \mathbb{E}_{\widetilde{\pi} \perp\!\!\!\perp \pi_*} \prod_{O \in \mathcal{O}} \mathbb{E}_{\mathcal{Q}}[X_O]$$

We will show

$$\mathbb{E}_{\mathcal{Q}}[X_O] = \frac{1}{1-\rho^{2|O|}} \tag{1}$$

Proof of (1)

$$X_{ij} \triangleq L\left(A_{\pi_*(i)\pi_*(j)}, B_{ij}\right) L\left(A_{\widetilde{\pi}(i)\widetilde{\pi}(j)}, B_{ij}\right).$$

where for Gaussian model

$$L(a,b) = \frac{P(a,b)}{Q(a,b)} = \frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{-\rho^2\left(b^2+a^2\right)+2\rho ab}{2\left(1-\rho^2\right)}\right).$$

Proof of (1)

$$X_{ij} \triangleq L\left(A_{\pi_*(i)\pi_*(j)}, B_{ij}\right) L\left(A_{\widetilde{\pi}(i)\widetilde{\pi}(j)}, B_{ij}\right).$$

where for Gaussian model

$$L(a,b) = \frac{P(a,b)}{Q(a,b)} = \frac{1}{\sqrt{1-\rho^2}} \exp\left(\frac{-\rho^2\left(b^2+a^2\right)+2\rho ab}{2\left(1-\rho^2\right)}\right).$$

Example: $\pi_* = \text{id}$, $\tilde{\pi} = \sigma$ as previously. Consider $O = \{14, 15, 16\}$:

$$X_O = \underbrace{L(A_{14}, B_{14})L(A_{15}, B_{14})}\underbrace{L(A_{15}, B_{15})L(A_{16}, B_{15})}\underbrace{L(A_{16}, B_{16})L(A_{14}, B_{16})}$$

For an edge orbit $|O| = k$, computing $\mathbb{E}_{\mathcal{Q}}[X_O]$ boils down to

$$\mathbb{E}_{\mathcal{Q}}[X_O] = \mathbb{E}\left[\prod_{\ell=1}^{k} L\left(a_\ell, b_\ell\right) L\left(a_\ell, b_{(\ell+1) \bmod k}\right)\right], \quad a_\ell, b_\ell \overset{\text{iid}}{\sim} N(0,1)$$

Proof of (1)

Two ways:

1. Write $\mathbb{E}[\exp(x^\top Cx)]$, where $x = (a_1, \ldots, a_k, b_1, \ldots, b_k) \sim N(0, I_{2k})$. Find MGF of Gaussian quadratic form determined by eigenvalues of C.
2. Slicker way: view L as a kernel

$$(Lf)(x) \triangleq \mathbb{E}_{Y \sim Q}\left[L(x,Y)f(Y)\right] = \mathbb{E}_{(X,Y)\sim P}\left[f(Y) \mid X = x\right].$$

and $L^2 \equiv L \circ L$. Then

$$\mathbb{E}\left[\prod_{\ell=1}^{k} L\left(a_\ell, b_\ell\right) L\left(a_\ell, b_{(\ell+1) \bmod k}\right)\right] = \mathbb{E}\left[\prod_{\ell=1}^{k} L^2\left(a_\ell, a_{(\ell+1) \bmod k}\right)\right]$$
$$= \operatorname{tr}\left(L^{2k}\right) = \sum \lambda_i^{2k}$$

where $\lambda_i = \rho^i$ (L is Mehler kernel, diagonalized by Hermite polynomials).

Second moment via orbit decomposition (2)

Overall, we get

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = \mathbb{E}_{\sigma}\left[\prod_{k=1}^{\binom{n}{2}}\left(\frac{1}{1-\rho^{2k}}\right)^{N_k}\right]$$

where

- $\sigma = \pi_*^{-1} \circ \tilde{\pi}$ is a uniform random permutation on $[n]$
- Cycle length of σ: $n_1, n_2, \ldots$
- Cycle length of σ^{E}: $N_1, N_2, \ldots$

$$N_1 = \binom{n_1}{2} + n_2, \quad N_2 = \binom{n_2}{2} \times 2 + n_1 n_2 + n_4$$

Second moment via orbit decomposition (2)

Overall, we get

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = \mathbb{E}_{\sigma}\left[\prod_{k=1}^{\binom{n}{2}}\left(\frac{1}{1-\rho^{2k}}\right)^{N_k}\right]$$

where

- $\sigma = \pi_*^{-1} \circ \tilde{\pi}$ is a uniform random permutation on $[n]$
- Cycle length of σ: $n_1, n_2, \ldots$
- Cycle length of σ^{E}: $N_1, N_2, \ldots$

$$N_1 = \binom{n_1}{2} + n_2, \quad N_2 = \binom{n_2}{2} \times 2 + n_1 n_2 + n_4$$

- Poisson approximation [Arratia-Tavaré '92]: n_k's are approximated independent Poi$(\frac{1}{k})$ (we will need their joint MGF)

Second moment via orbit decomposition (3)

Let $\tau = \log \frac{1}{1-\rho^2} = \rho^2(1+o(1))$. We get

$$\mathbb{E}_\sigma \left[\prod_{k=1}^{\binom{n}{2}} \left(\frac{1}{1-\rho^{2k}} \right)^{N_k} \right] \approx \mathbb{E}_\sigma \left[\exp(\tau N_1) \right] \approx \mathbb{E}_\sigma \left[\exp(\tau n_1^2/2) \right] \tag{2}$$

$$\approx \mathbb{E} \left[\exp(\tau \mathsf{Poi}(1)^2/2) \mathbf{1}_{\{\mathsf{Poi}(1) \le n\}} \right]$$

$$= \sum_{\ell=0}^{n} \frac{\exp(\tau \ell^2/2)}{\ell!} = 1 + o(1)$$

if $\tau = \frac{(2-\epsilon)\log n}{n}$.

Second moment via orbit decomposition (3)

Let $\tau = \log \frac{1}{1-\rho^2} = \rho^2(1+o(1))$. We get

$$\mathbb{E}_\sigma \left[\prod_{k=1}^{\binom{n}{2}} \left(\frac{1}{1-\rho^{2k}} \right)^{N_k} \right] \approx \mathbb{E}_\sigma \left[\exp(\tau N_1) \right] \approx \mathbb{E}_\sigma \left[\exp(\tau n_1^2/2) \right] \qquad (2)$$

$$\approx \mathbb{E}\left[\exp(\tau \mathsf{Poi}(1)^2/2) \mathbf{1}_{\{\mathsf{Poi}(1) \le n\}} \right]$$

$$= \sum_{\ell=0}^{n} \frac{\exp(\tau \ell^2/2)}{\ell!} = 1 + o(1)$$

if $\tau = \frac{(2-\epsilon)\log n}{n}$.

Summary: We have shown

$$\rho^2 \le \frac{(\mathbf{2} - \epsilon)\log n}{n} \implies \mathbb{E}_{\mathcal{Q}} \left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)} \right)^2 \right] = 1 + o(1).$$

But we want $\rho^2 \le \frac{(\mathbf{4}-\epsilon)\log n}{n}$...

Limitation of vanilla second-moment method

It turns out that

$$\rho^2 \geq \frac{(\mathbf{2} + \epsilon)\log n}{n} \implies \mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] \to \infty$$

- Gaussian: suboptimal by a factor of 2
- ER graphs: suboptimal by an unbounded factor when $p = o(1)$

Limitation of vanilla second-moment method

It turns out that

$$\rho^2 \geq \frac{(\mathbf{2} + \epsilon)\log n}{n} \implies \mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] \to \infty$$

- Gaussian: suboptimal by a factor of 2
- ER graphs: suboptimal by an unbounded factor when $p = o(1)$

Obstruction from short orbits

$$\mathbb{E}_{(A,B)\sim\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = \mathbb{E}_{\pi \perp\!\!\!\perp \widetilde{\pi}}\left[\prod_{O\in\mathcal{O}} \mathbb{E}_{\mathcal{Q}}\left[X_O\right]\right] \overset{\widetilde{\pi}=\pi}{\geq} \frac{1}{n!}\left(1+\rho^2\right)^{\binom{n}{2}}$$

Limitation of vanilla second-moment method

It turns out that

$$\rho^2 \geq \frac{(\mathbf{2} + \epsilon) \log n}{n} \implies \mathbb{E}_{\mathcal{Q}} \left[\left(\frac{\mathcal{P}(A, B)}{\mathcal{Q}(A, B)} \right)^2 \right] \to \infty$$

- Gaussian: suboptimal by a factor of 2
- ER graphs: suboptimal by an unbounded factor when $p = o(1)$

Obstruction from short orbits

$$\mathbb{E}_{(A,B)\sim\mathcal{Q}} \left[\left(\frac{\mathcal{P}(A, B)}{\mathcal{Q}(A, B)} \right)^2 \right] = \mathbb{E}_{\pi \perp\!\!\!\perp \widetilde{\pi}} \left[\prod_{O \in \mathcal{O}} \mathbb{E}_{\mathcal{Q}} [X_O] \right] \overset{\widetilde{\pi} = \pi}{\geq} \frac{1}{n!} \left(1 + \rho^2\right)^{\binom{n}{2}}$$

Atypically large magnitude of $\prod_{O \in \mathcal{O}: |O| = k} X_O$ for short orbits of length $k \lesssim \log n \Rightarrow$ second-moment blows up

Truncated second-moment method

Let $\mathcal{E}$ denote a typical event under $\mathcal{P}$, i.e., $\mathcal{P}((A, B, \pi_*) \in \mathcal{E}) = 1 - o(1)$.

$$\text{Truncated 2nd moment} = \mathbb{E}_{\pi_* \perp\!\!\!\perp \widetilde{\pi}} \left[\mathbb{E}_Q \left[\prod_{O \in \mathcal{O}} X_O \mathbf{1}_{\{(A,B,\pi) \in \mathcal{E}\}} \mathbf{1}_{\{(A,B,\widetilde{\pi}) \in \mathcal{E}\}} \right] \right]$$

Then

$$\text{Truncated 2nd moment} = O(1) \implies \mathsf{TV}(\mathcal{P}(A,B), \mathcal{Q}(A,B)) \le 1 - \Omega(1)$$

$$\text{Truncated 2nd moment} = 1 + o(1) \implies \mathsf{TV}(\mathcal{P}(A,B), \mathcal{Q}(A,B)) = o(1)$$

Truncated second-moment method

Let $\mathcal{E}$ denote a typical event under $\mathcal{P}$, i.e., $\mathcal{P}((A, B, \pi_*) \in \mathcal{E}) = 1 - o(1)$.

$$\text{Truncated 2nd moment} = \mathbb{E}_{\pi_* \perp\!\!\!\perp \widetilde{\pi}}\left[\mathbb{E}_Q\left[\prod_{O \in \mathcal{O}} X_O \mathbf{1}_{\{(A,B,\pi)\in\mathcal{E}\}} \mathbf{1}_{\{(A,B,\widetilde{\pi})\in\mathcal{E}\}}\right]\right]$$

Then

$$\text{Truncated 2nd moment} = O(1) \implies \text{TV}(\mathcal{P}(A,B), \mathcal{Q}(A,B)) \le 1 - \Omega(1)$$
$$\text{Truncated 2nd moment} = 1 + o(1) \implies \text{TV}(\mathcal{P}(A,B), \mathcal{Q}(A,B)) = o(1)$$

Caveat:

- The event $\mathcal{E}$ must be measurable wrt (A, B, π_*).
- Although $\pi_* = \tilde{\pi}$ is a rare event, we cannot truncate on anything involving the interaction between two replicas $(\pi_*, \tilde{\pi})$.

Truncated second-moment method

Let $\mathcal{E}$ denote a typical event under $\mathcal{P}$, i.e., $\mathcal{P}((A, B, \pi_*) \in \mathcal{E}) = 1 - o(1)$.

$$\text{Truncated 2nd moment} = \mathbb{E}_{\pi_* \perp\!\!\!\perp \widetilde{\pi}} \left[\mathbb{E}_Q \left[\prod_{O \in \mathcal{O}} X_O \mathbf{1}_{\{(A,B,\pi) \in \mathcal{E}\}} \mathbf{1}_{\{(A,B,\widetilde{\pi}) \in \mathcal{E}\}} \right] \right]$$

Then

$$\text{Truncated 2nd moment} = O(1) \implies \mathsf{TV}(\mathcal{P}(A,B), \mathcal{Q}(A,B)) \le 1 - \Omega(1)$$
$$\text{Truncated 2nd moment} = 1 + o(1) \implies \mathsf{TV}(\mathcal{P}(A,B), \mathcal{Q}(A,B)) = o(1)$$

Caveat:

- The event $\mathcal{E}$ must be measurable wrt (A, B, π_*).
- Although $\pi_* = \tilde{\pi}$ is a rare event, we cannot truncate on anything involving the interaction between two replicas $(\pi_*, \tilde{\pi})$.

Let's see why.

Details of truncated second-moment

Goal: bound $\mathsf{TV}(\mathcal{P}(A,B),\mathcal{Q}(A,B))$ from above.

- Introduce conditional planted model:

$$\mathcal{P}'\left(A,B,\pi\right) \triangleq \frac{\mathcal{P}\left(A,B,\pi\right)\mathbf{1}_{\{(A,B,\pi)\in\mathcal{E}\}}}{\mathcal{P}\left(\mathcal{E}\right)} = \left(1+o\left(1\right)\right)\mathcal{P}\left(A,B,\pi\right)\mathbf{1}_{\{(A,B,\pi)\in\mathcal{E}\}},$$

Details of truncated second-moment

Goal: bound $\mathsf{TV}(\mathcal{P}(A,B),\mathcal{Q}(A,B))$ from above.

- Introduce conditional planted model:

$$\mathcal{P}'(A,B,\pi) \triangleq \frac{\mathcal{P}(A,B,\pi)\,\mathbf{1}_{\{(A,B,\pi)\in\mathcal{E}\}}}{\mathcal{P}(\mathcal{E})} = (1+o(1))\,\mathcal{P}(A,B,\pi)\,\mathbf{1}_{\{(A,B,\pi)\in\mathcal{E}\}},$$

- Triangle inequality of TV

$$\mathsf{TV}(\mathcal{P}(A,B),\mathcal{Q}(A,B)) \le \mathsf{TV}(\mathcal{P}'(A,B),\mathcal{Q}(A,B)) + \underbrace{\mathsf{TV}(\mathcal{P}(A,B),\mathcal{P}'(A,B))}_{\le \mathcal{P}((A,B,\pi_*)\notin\mathcal{E})=o(1)}$$

Details of truncated second-moment

Goal: bound $\mathsf{TV}(\mathcal{P}(A,B), \mathcal{Q}(A,B))$ from above.

- Introduce conditional planted model:

$$\mathcal{P}'(A,B,\pi) \triangleq \frac{\mathcal{P}(A,B,\pi)\,\mathbf{1}_{\{(A,B,\pi)\in\mathcal{E}\}}}{\mathcal{P}(\mathcal{E})} = (1+o(1))\,\mathcal{P}(A,B,\pi)\,\mathbf{1}_{\{(A,B,\pi)\in\mathcal{E}\}},$$

- Triangle inequality of TV

$$\mathsf{TV}(\mathcal{P}(A,B),\mathcal{Q}(A,B)) \le \mathsf{TV}(\mathcal{P}'(A,B),\mathcal{Q}(A,B)) + \underbrace{\mathsf{TV}(\mathcal{P}(A,B),\mathcal{P}'(A,B))}_{\le \mathcal{P}((A,B,\pi_*)\notin\mathcal{E})=o(1)}$$

- Apply second-moment method

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}'(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = (1+o(1))\,\mathbb{E}_{\pi_*\perp\!\!\!\perp\widetilde{\pi}}\left[\mathbb{E}_{Q}\left[\underbrace{\frac{\mathcal{P}(A,B\mid\pi)}{\mathcal{Q}(A,B)}\frac{\mathcal{P}(A,B\mid\widetilde{\pi})}{\mathcal{Q}(A,B)}}_{\prod_{O\in\mathcal{O}} X_O}\mathbf{1}_{\{(A,B,\pi)\in\mathcal{E}\}}\mathbf{1}_{\{(A,B,\widetilde{\pi})\in\mathcal{E}\}}\right]\right]$$

Truncated second-moment: Gaussian model

Major contribution comes from $k = 1$ (fixed points):

$$Y \triangleq \prod_{O \in \mathcal{O}:|O|=1} X_O \approx \exp\left(-\rho^2 \binom{n_1}{2} + 2\rho e_{A^{\pi_*} \wedge B}(F)\right)$$

- F is the set of fixed points of $\sigma \triangleq \pi_*^{-1} \circ \widetilde{\pi}$ and $n_1 = |F|$
- $e_{A^{\pi_*} \wedge B}(F) \triangleq \sum_{(i,j) \in F} A_{\pi_*(i)\pi_*(j)} B_{ij}$

Truncated second-moment: Gaussian model

Major contribution comes from $k = 1$ (fixed points):

$$Y \triangleq \prod_{O \in \mathcal{O}:|O|=1} X_O \approx \exp\left(-\rho^2 \binom{n_1}{2} + 2\rho e_{A^{\pi_*} \wedge B}(F)\right)$$

- F is the set of fixed points of $\sigma \triangleq \pi_*^{-1} \circ \widetilde{\pi}$ and $n_1 = |F|$
- $e_{A^{\pi_*} \wedge B}(F) \triangleq \sum_{(i,j) \in F} A_{\pi_*(i)\pi_*(j)} B_{ij}$
- Under $\mathcal{P}$: $e_{A^{\pi_*} \wedge B}(S)$ concentrates on its mean $\rho\binom{|S|}{2}$ uniformly over all S with large $|S|$ (Hanson-Wright)

Truncated second-moment: Gaussian model

Major contribution comes from $k = 1$ (fixed points):

$$Y \triangleq \prod_{O \in \mathcal{O}:|O|=1} X_O \approx \exp\left(-\rho^2 \binom{n_1}{2} + 2\rho e_{A^{\pi_*} \wedge B}(F)\right)$$

- F is the set of fixed points of $\sigma \triangleq \pi_*^{-1} \circ \widetilde{\pi}$ and $n_1 = |F|$
- $e_{A^{\pi_*} \wedge B}(F) \triangleq \sum_{(i,j) \in F} A_{\pi_*(i)\pi_*(j)} B_{ij}$
- Under $\mathcal{P}$: $e_{A^{\pi_*} \wedge B}(S)$ concentrates on its mean $\rho\binom{|S|}{2}$ uniformly over all S with large $|S|$ (Hanson-Wright)
- On this typical (under $\mathcal{P}$) event $\mathcal{E}$, when $|F|$ is large,

$$\mathbb{E}_{\mathcal{Q}}\left[Y \mathbf{1}_{\mathcal{E}}\right] \lesssim e^{-\rho^2 \binom{n_1}{2}} \mathbb{E}_{\mathcal{Q}}\left[e^{2\rho e_{A^{\pi_*} \wedge B}(F)} \mathbf{1}_{\left\{e_{A \wedge B_\pi}(F) \le \rho\binom{n_1}{2}\right\}}\right]$$

$$\approx \exp\left(\frac{\rho^2}{2}\binom{n_1}{2}\right) \quad \text{(Gain a factor of 2 over (2))}$$

by truncated MGF

Truncated second-moment: sparse Erdős-Rényi

Need to consider $k = \Theta(\log n)$. It can be shown

- Long orbits:

$$\mathbb{E}_{\mathcal{Q}}\left[\prod_{|O|>k} X_O\right] \leq \left(1+\rho^k\right)^{\frac{n^2}{k}} = 1 + o(1)$$

- Short incomplete orbits:

$$\mathbb{E}_{\mathcal{Q}}\left[X_O \mid O \not\subset E\left(A \wedge B^\pi\right)\right] \leq 1$$

- Short complete orbits:

$$X_O = \left(\frac{1}{p}\right)^{2|O|}, \quad \forall O \subset E\left(A \wedge B^\pi\right)$$

Suffices to consider subgraph $H_k \triangleq \cup_{O:|O|\leq k, O\subset E(A\wedge B^\pi)} O$

Truncated second-moment: sparse Erdős-Rényi

- If $nps^2 \le 1 - \omega(n^{-1/3})$:

$$\mathcal{E} \triangleq \{A^\pi \wedge B \text{ is a pseudoforest}\}$$

- If $nps^2 \le \lambda^* - \epsilon$:

$$\mathcal{E} \triangleq \{\text{The subgraph density of } A^\pi \wedge B \text{ is smaller than } \gamma(\lambda^*)\}$$

Then

$$\mathbb{E}_\mathcal{Q}\left[\prod_{O \in \mathcal{O}} X_O \mathbf{1}_\mathcal{E}\right] \le (1+o(1))\,\mathbb{E}_\mathcal{Q}\left[\left(\frac{1}{p}\right)^{2e(H_k)} \mathbf{1}_{\{H_k \text{ is admissible}\}}\right]$$

$$= (1+o(1)) \sum_{H \in \mathcal{H}_k} s^{2e(H)} \quad \text{(generating function)}$$

$\mathcal{H}_k$: The set of all admissible H_k

Truncated second-moment: sparse Erdős-Rényi

- If $nps^2 \le 1 - \omega(n^{-1/3})$:

$$\mathcal{E} \triangleq \{A^\pi \wedge B \text{ is a pseudoforest}\}$$

- If $nps^2 \le \lambda^* - \epsilon$:

$$\mathcal{E} \triangleq \{\text{The subgraph density of } A^\pi \wedge B \text{ is smaller than } \gamma(\lambda^*)\}$$

Then

$$\mathbb{E}_{\mathcal{Q}}\left[\prod_{O \in \mathcal{O}} X_O \mathbf{1}_{\mathcal{E}}\right] \le (1+o(1))\,\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{1}{p}\right)^{2e(H_k)} \mathbf{1}_{\{H_k \text{ is admissible}\}}\right]$$
$$= (1+o(1)) \sum_{H \in \mathcal{H}_k} s^{2e(H)} \quad \text{(generating function)}$$

$\mathcal{H}_k$: The set of all admissible H_k

Key remaining challenge: enumerate $\mathcal{H}_k$ using orbit structure

Analysis

- Proof of detection thresholds
- Proof of exact recovery thresholds

Exact recovery: Positive results

- Decompose the difference of objectives via edge orbits

$$\langle A^{\pi} - A^{\pi_*}, B\rangle = \sum_{O \in \mathcal{O} \setminus \mathcal{O}_1} \underbrace{\sum_{(i,j)\in O} A_{\pi(i)\pi(j)} B_{ij}}_{X_O} - \sum_{O \in \mathcal{O} \setminus \mathcal{O}_1} \underbrace{\sum_{(i,j)\in O} A_{\pi^*(i)\pi^*(j)} B_{ij}}_{Y_O}$$

- Apply large deviation analysis:
  - For π far away from π^*: bound $\sum_O X_O$ and $\sum_O Y_O$ separately
  - For π close to π^*: bound $\sum_O (X_O - Y_O)$ directly
- The contribution of longer edge orbits can be effectively bounded by that of the 2-edge orbits

$$M_{|O|} \triangleq \mathbb{E}\left[\exp(tX_O)\right] \leq M_2^{|O|/2}, \quad \forall |O| \geq 2$$

Computation of $M_{|O|}$ is similar to (1)

Exact recovery: Negative results

- Suffices to show MLE fails (WLOG $\pi_* = \text{id}$)
- Bottleneck: π is a transposition swapping i and j, for which

$$\Delta_{ij} \equiv \langle A^{\pi} - A^{\pi_*}, B \rangle = -\sum_{k \neq i,j} (A_{ik} - A_{jk})(B_{ik} - B_{jk})$$

- Prove the existence of (i, j) for which $\Delta_{ij} > 0$ whp

Exact recovery: Negative results

- Suffices to show MLE fails (WLOG $\pi_* = \text{id}$)
- Bottleneck: π is a transposition swapping i and j, for which

$$\Delta_{ij} \equiv \langle A^{\pi} - A^{\pi_*}, B \rangle = -\sum_{k \neq i,j} (A_{ik} - A_{jk})(B_{ik} - B_{jk})$$

- Prove the existence of (i,j) for which $\Delta_{ij} > 0$ whp
- Since $B = \rho A + \sqrt{1-\rho^2} Z$, conditioned on variance parameter $v_{ij} \equiv \sum_{k \neq i,j} (A_{ik} - A_{jk})^2$,

$$\Delta_{ij} \sim N(-\rho v_{ij}, 2(1-\rho^2) v_{ij})$$

- Whp, all v_{ij} concentrates on $\mathbb{E}[v_{ij}] \approx 2n$. So $\mathbb{P}\{\Delta_{ij} > 0\} \approx \exp(-\frac{\rho^2 n}{2}$.
- Total number of transpositions: $\binom{n}{2}$. So $\rho^2 \leq \frac{(4-\epsilon)\log n}{n} \implies \mathbb{E}[\sum \mathbf{1}_{\{\Delta_{ij} > 0\}}] \to \infty$.

Exact recovery: Negative results

- Suffices to show MLE fails (WLOG $\pi_* = \mathrm{id}$)
- Bottleneck: π is a transposition swapping i and j, for which

$$\Delta_{ij} \equiv \langle A^{\pi} - A^{\pi_*}, B \rangle = -\sum_{k \neq i,j} (A_{ik} - A_{jk})(B_{ik} - B_{jk})$$

- Prove the existence of (i, j) for which $\Delta_{ij} > 0$ whp
- Since $B = \rho A + \sqrt{1-\rho^2} Z$, conditioned on variance parameter $v_{ij} \equiv \sum_{k \neq i,j} (A_{ik} - A_{jk})^2$,

$$\Delta_{ij} \sim N(-\rho v_{ij}, 2(1-\rho^2) v_{ij})$$

- Whp, all v_{ij} concentrates on $\mathbb{E}[v_{ij}] \approx 2n$. So $\mathbb{P}\{\Delta_{ij} > 0\} \approx \exp(-\frac{\rho^2 n}{2}$.
- Total number of transpositions: $\binom{n}{2}$. So $\rho^2 \leq \frac{(4-\epsilon)\log n}{n} \implies \mathbb{E}[\sum \mathbf{1}_{\{\Delta_{ij} > 0\}}] \to \infty$.
- Since Δ_{ij} are not independent, need to compute 2nd moment applying Paley-Zymund (Chap 1)

Concluding remarks

		Partial recovery & detection	Almost exact recovery	Exact recovery
p	$n^{-o(1)}$	$nps^2 = \frac{2\log n}{\log(1/p)-1+p}$		$\frac{nps^2}{(1-\sqrt{p})^2\log n} = 1$
	$n^{-\alpha}$	$nps^2 = \lambda^*$	$nps^2 = \omega(1)$	
Gaussian		$\frac{n\rho^2}{\log n} = 4$		

Concluding remarks

		Partial recovery & detection	Almost exact recovery	Exact recovery
p	$n^{-o(1)}$	$nps^2 = \frac{2\log n}{\log(1/p)-1+p}$		$\frac{nps^2}{(1-\sqrt{p})^2 \log n} = 1$
	$n^{-\alpha}$	$nps^2 = \lambda^*$	$nps^2 = \omega(1)$	
Gaussian		$\frac{n\rho^2}{\log n} = 4$		

Reference

- Y. Wu, J. Xu, & S. H. Yu, *Testing correlation of unlabeled random graphs*, Annals of Applied Probability, arXiv:2008.10097.
- Y. Wu, J. Xu, & S. H. Yu, *Settling the sharp reconstruction thresholds of random graph matching*, IEEE Transactions on Information Theory, arXiv:2102.00082.
- J. Ding & H. Du, *Detection threshold for correlated Erdős-Rényi graphs via densest subgraphs*. arXiv:2203.14573.
- J. Ding & H. Du, *Matching recovery threshold for correlated random graphs*. arXiv:2205.14650.