Lecture 2: Random Graph Matching: Information-theoretic Limits

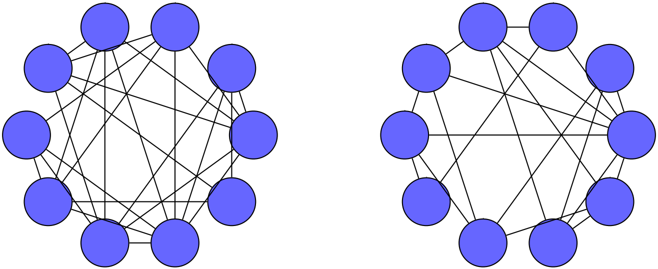
Jiaming Xu

The Fuqua School of Business
Duke University

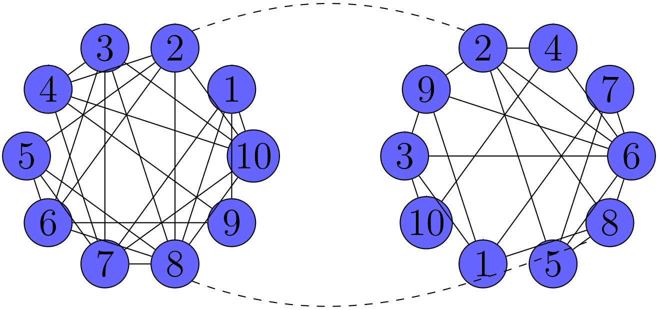
Joint work with
Yihong Wu (Yale) and Sophie H. Yu (Duke)

April 18, 2023
AI4OPT Tutorial Lectures

Graph matching (network alignment)

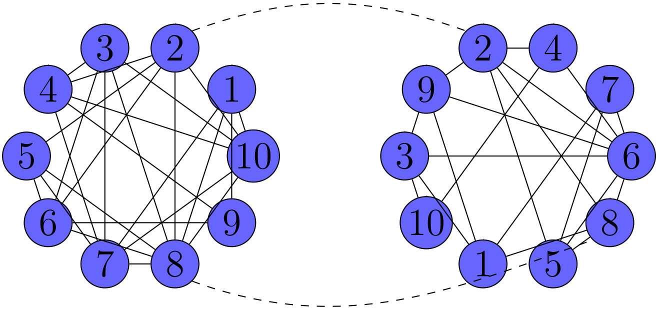


Graph matching (network alignment)



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

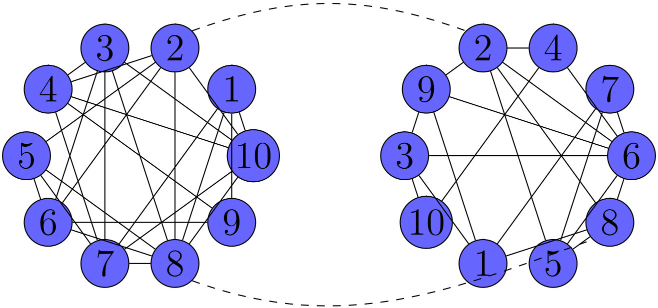
Graph matching (network alignment)



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

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

Graph matching (network alignment)

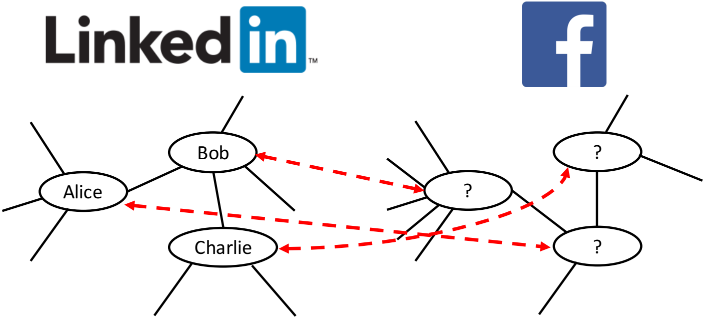


Goal: find a mapping between two node sets that maximally aligns the edges (i.e. minimizes # of adjacency disagreements)

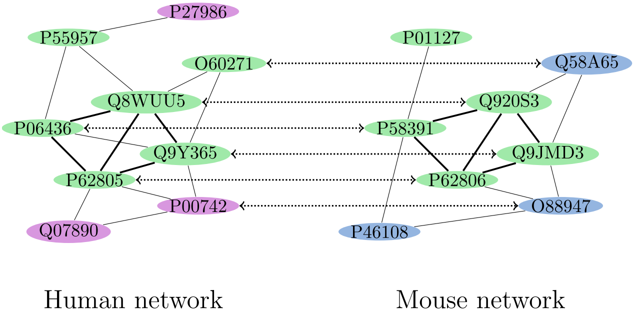
Quadratic Assignment Problem (QAP) : $$\max_{\pi \in S_n} \sum_{i<j} A_{ij} B_{\pi(i)\pi(j)}$$

Noiseless case: reduce to graph isomorphism

Application 1: Network de-anonymization



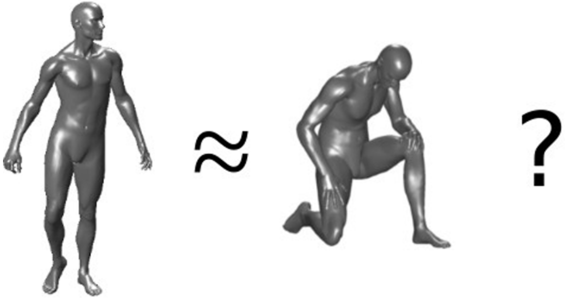
- Successfully de-anonymize Netflix dataset by matching it to IMDB [Narayanan-Shmatikov '08]
- Correctly identify 30.8% of shared users between Twitter and Flickr [Narayanan-Shmatikov '09]



[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]

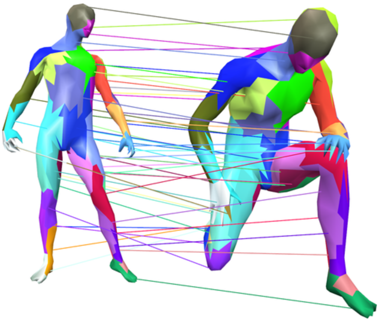
A fundamental problem in computer vision: Detect and match similar objects that undergo different deformations



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

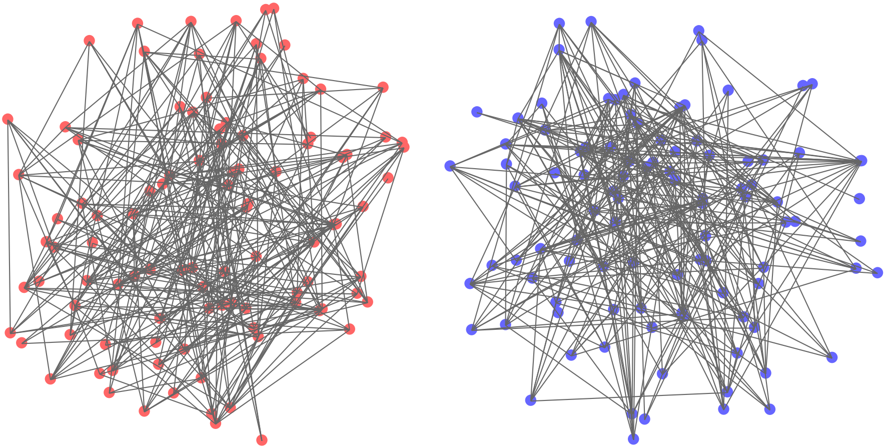


Shape REtrieval Contest (SHREC) dataset [Lähner et al '16]

3-D shapes $\to$ geometric graphs (features $\to$ nodes, distances $\to$ edges)

Two key challenges

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



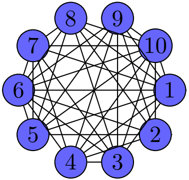
Beyond worst-case intractability

- NP-hard for matching two graphs in worst case
- However, real networks are not arbitrary and have latent structures

Beyond worst-case intractability

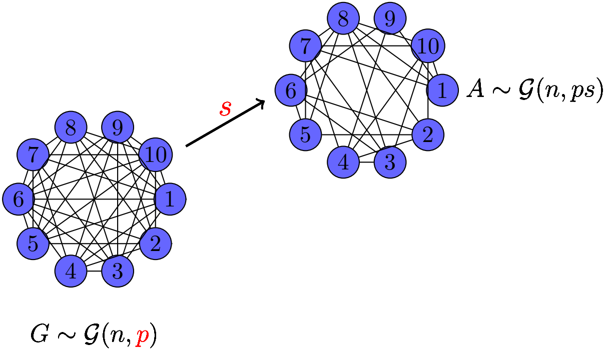
- NP-hard for matching two graphs in worst case
- However, real networks are not arbitrary and have latent structures
- Recent surge of interests on the average-case analysis of matching two correlated random graphs [Feizi at el.'16, Lyzinski at el'16, Cullina-Kiyavash'16,17, Ding-Ma-Wu-Xu'18, Barak-Chou-Lei-Schramm-Sheng'19, Fan-Mao-Wu-Xu'19a,19b, Ganassali-Massoulié'20, Hall-Massoulié'20, ...]
- Focus on correlated Erdős-Rényi graphs model [Pedarsani-Grossglauser '11]

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

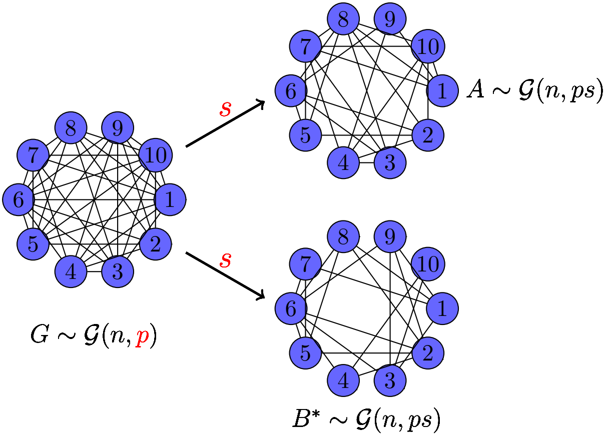


$G \sim \mathcal{G}(n, p)$

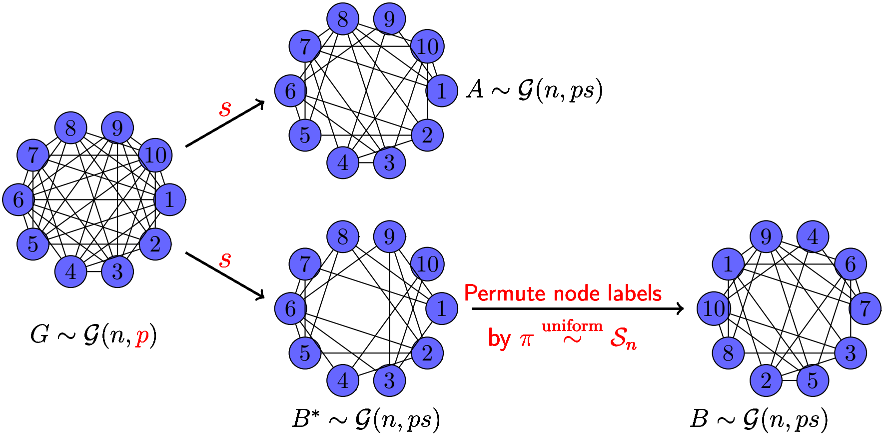
Correlated Erdős-Rényi graphs model $\mathcal{G}(n, p, s)$



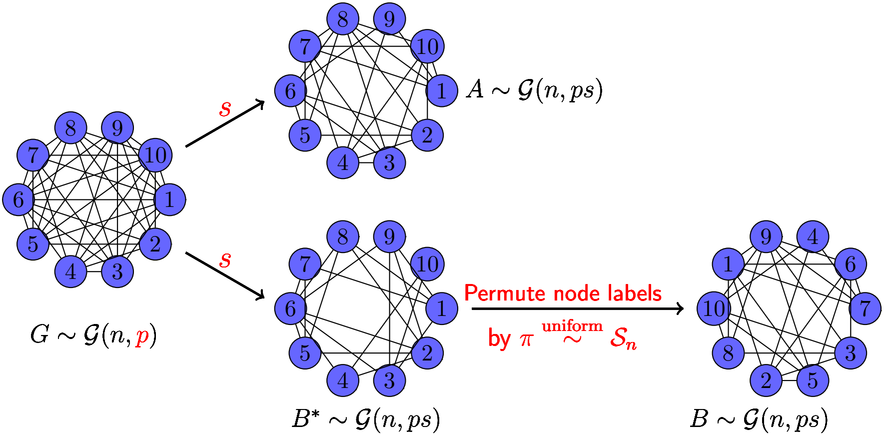
Correlated Erdős-Rényi graphs model $\mathcal{G}(n, p, s)$



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 $i.i.d.$ 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 $i.i.d.$ standard normal entries;
- $A^{\pi} = (A_{\pi(i)\pi(j)})$ denote the relabeled version of A
- Conditional on π, for any $1 \le i < j \le n$,

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

Two statistical tasks: detection and estimation

- Detection:
  - $\mathcal{H}_0$: A and B are two independent Erdős-Rényi graphs $\mathcal{G}(n, ps)$
  - $\mathcal{H}_1$: A and B are two 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

Two statistical tasks: detection and estimation

- Detection:
  - $\mathcal{H}_0$: A and B are two independent Erdős-Rényi graphs $\mathcal{G}(n, ps)$
  - $\mathcal{H}_1$: A and B are two 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

Focus of this lecture

What are the information-theoretic limits of detection and estimation?

Two statistical tasks: detection and estimation

- Detection:
  - $\mathcal{H}_0$: A and B are two independent Erdős-Rényi graphs $\mathcal{G}(n, ps)$
  - $\mathcal{H}_1$: A and B are two 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

Focus of this lecture

What are the information-theoretic limits of detection and estimation?

10+ years of development:

[Pedarsani-Grossglauser '11], [Cullina-Kiyavash '16,17], [Hall-Massoulié '20], [Ganassali '20], [Wu-X.-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):

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

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

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

Sharp threshold for detection: Gaussian

Theorem (Wu-X.-Yu '20)

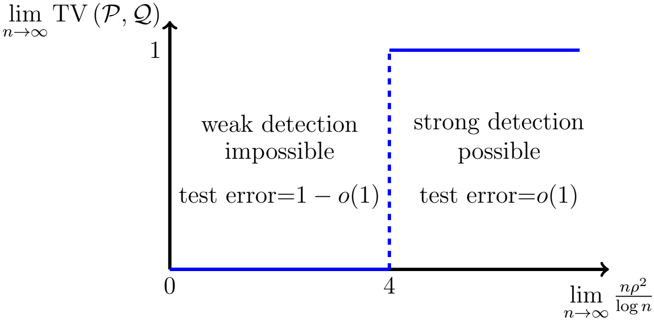
$$n\rho^2 \geq (4+\epsilon)\log n \implies \mathrm{TV}\left(\mathcal{P},\mathcal{Q}\right) = 1 - o(1)\,(\textit{test error}{=}o(1))$$

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

Sharp threshold for detection: Gaussian

Theorem (Wu-X.-Yu '20)

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

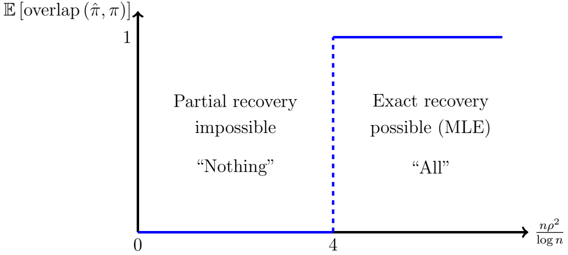


Sharp threshold for recovery: Gaussian model

Theorem (Wu-X.-Yu '21)

$$n\rho^2 \geq (4+\epsilon)\log n \implies \widehat{\pi}_{\mathsf{ML}} = \pi \; \textit{w.h.p}$$

$$n\rho^2 \leq (4-\epsilon)\log n \implies \text{overlap}\,(\widehat{\pi}, \pi) = o(1), \; \textit{w.h.p}, \; \forall \; \textit{estimator} \; \widehat{\pi}$$

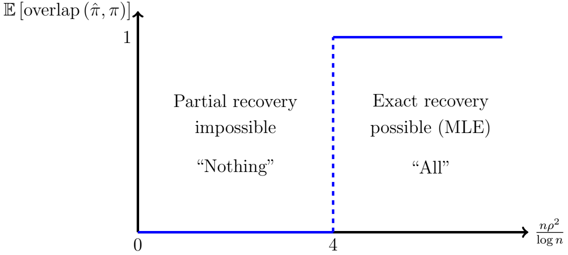


Sharp threshold for recovery: Gaussian model

Theorem (Wu-X.-Yu '21)

$$n\rho^2 \geq (4+\epsilon)\log n \implies \widehat{\pi}_{\mathsf{ML}} = \pi \;\; \textit{w.h.p}$$

$$n\rho^2 \leq (4-\epsilon)\log n \implies \text{overlap}\,(\widehat{\pi}, \pi) = o(1), \;\; \textit{w.h.p}, \; \forall \; \textit{estimator } \widehat{\pi}$$



- Exact recovery threshold is derived in [Ganassali '20]
- Exhibits a stronger form of "all or nothing" phenomenon

Sharp detection threshold: dense Erdős-Rényi graphs

Theorem (Wu-X.-Yu '20)

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

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

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

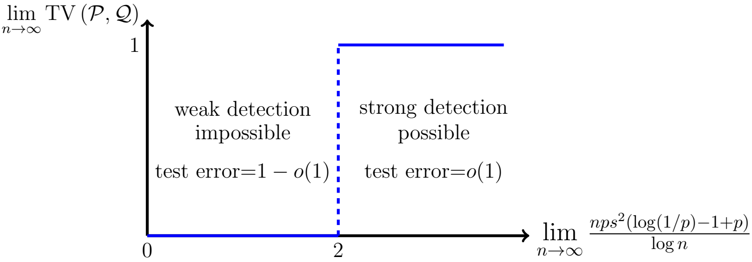
Sharp detection threshold: dense Erdős-Rényi graphs

Theorem (Wu-X.-Yu '20)

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

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

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



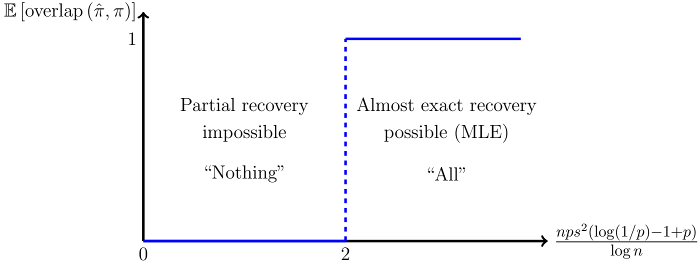
Sharp recovery threshold: dense Erdős-Rényi

Theorem (Wu-X.-Yu '21)

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

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

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



Sharp recovery threshold: dense Erdős-Rényi

Theorem (Wu-X.-Yu '21)

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

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

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

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 $I(\pi; A, B) \approx H(\pi)$

Sharp detection threshold: sparse Erdős-Rényi

Theorem (Ding-Du '22a)

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

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

- Sharpened our threshold at $nps^2 = \Theta(1)$ [Wu-X.-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-Hu '22b)

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

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

Sharp recovery threshold: sparse Erdős-Rényi

Theorem (Ding-Hu '22b)

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

$$nps^2 \geq \lambda^* + \epsilon \implies \text{overlap}\,(\widehat{\pi}_{\text{ML}}, \pi) \geq \Omega(1) \text{ w.h.p.}$$
$$nps^2 \leq \lambda^* - \epsilon \implies \text{overlap}\,(\widehat{\pi}, \pi) = o(1) \text{ w.h.p. } \forall \widehat{\pi}$$

- The negative result of $\alpha = 1$ is proved in [Ganassali-Lelarge-Massoulié '21]
- Sharpen our partial recovery threshold at $nps^2 = \Theta(1)$ [Wu-X.-Yu '21]
- "All-or-nothing" phenomenon does not exist, as almost exact recovery (overlap $= 1 - o(1)$) requires $nps^2 \to \infty$ [Cullina-Kiyavash-Mittal-Poor '19]

Exact recovery threshold

Theorem (Wu-X.-Yu '21)

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

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

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

Exact recovery threshold

Theorem (Wu-X.-Yu '21)

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

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

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

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

Exact recovery threshold

Theorem (Wu-X.-Yu '21)

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

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

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

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

Analysis

- Proof of detection thresholds
- Proof of partial recovery thresholds
- Proof of exact recovery thresholds

Proof of detection thresholds: Positive results

- Gaussian or dense Erdős-Rényi: analyzing QAP statistic

$$\max_{\pi \in \mathcal{S}_n} \sum_{i<j} A_{ij} B_{\pi(i)\pi(j)} \quad (\text{\# of common edges})$$

- 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 $A^\pi \wedge B$

- Standard first-moment computation

Proof of detection thresholds: Negative results

Second-moment method

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

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

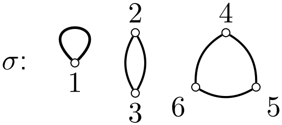
Cycle (orbit) decomposition

- **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 $\binom{[n]}{2}$: $\sigma^{\mathsf{E}}((i,j)) \triangleq (\sigma(i), \sigma(j))$

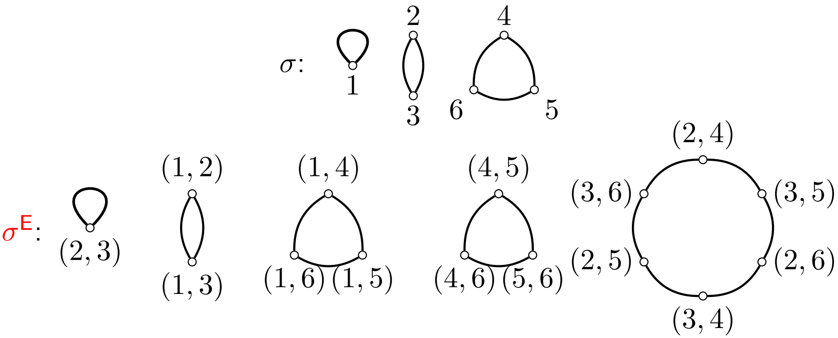
Example: $n = 6$ and $\sigma = (1)(23)(456)$:



Cycle (orbit) decomposition

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

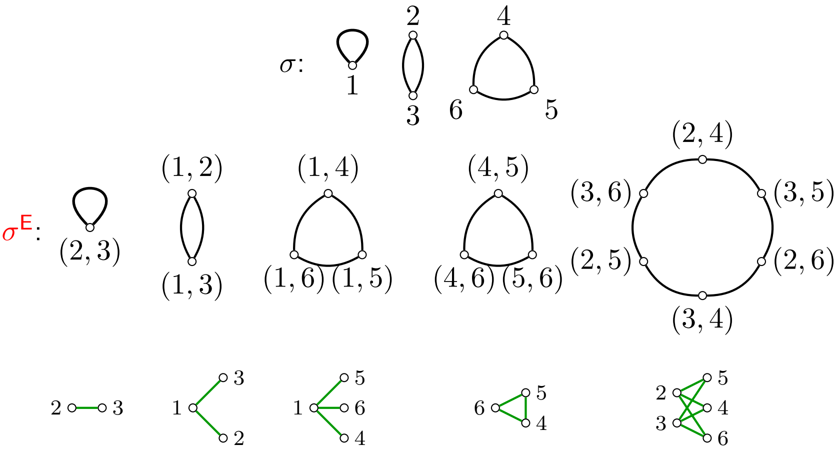
Example: $n = 6$ and $\sigma = (1)(23)(456)$:



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 calculation via orbit decomposition

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

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

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

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

Second-moment calculation via orbit decomposition

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

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

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

$\mathcal{O}$: 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 calculation via orbit decomposition

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

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

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

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

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

$$\mathbb{E}_{\mathcal{Q}}[X_O] = \begin{cases} \frac{1}{1-\rho^{2|O|}} & \text{Gaussian} \\ 1+\rho^{2|O|} & \text{Erdős-Rényi} \end{cases}$$

Failure of second-moment

We show

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = \begin{cases} 1+o(1) & \text{if } \rho^2 \le \frac{(2-\epsilon)\log n}{n} \\ +\infty & \text{if } \rho^2 \ge \frac{(2+\epsilon)\log n}{n} \end{cases}$$

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

Failure of second-moment

We show

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = \begin{cases} 1+o(1) & \text{if } \rho^2 \le \frac{(2-\epsilon)\log n}{n} \\ +\infty & \text{if } \rho^2 \ge \frac{(2+\epsilon)\log n}{n} \end{cases}$$

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

Obstruction from short orbits

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

Failure of second-moment

We show

$$\mathbb{E}_{\mathcal{Q}}\left[\left(\frac{\mathcal{P}(A,B)}{\mathcal{Q}(A,B)}\right)^2\right] = \begin{cases} 1+o(1) & \text{if } \rho^2 \le \frac{(2-\epsilon)\log n}{n} \\ +\infty & \text{if } \rho^2 \ge \frac{(2+\epsilon)\log n}{n} \end{cases}$$

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

Obstruction from short orbits

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

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

Let $\mathcal{E}$ denote an event that holds whp under $\mathcal{P}$:

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

Strong detection is impossible

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

Weak detection is impossible

Truncated second-moment: Gaussian model

It suffices to consider $k = 1$:

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

- 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

It suffices to consider $k = 1$:

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

- 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 $\rho\binom{|S|}{2}$ uniformly over all S when $|S|$ is large

Truncated second-moment: Gaussian model

It suffices to consider $k = 1$:

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

- 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 $\rho\binom{|S|}{2}$ uniformly over all S when $|S|$ is large
- On this typical event $\mathcal{E}$ under $\mathcal{P}$, when $|F|$ is large,

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

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

Truncated second-moment: dense Erdős-Rényi

Again it suffices to consider $k = 1$:

$$Y \triangleq \prod_{O \in \mathcal{O}:|O|=1} X_O \approx \left(\frac{1}{p}\right)^{2e_{A^\pi \wedge B}(F)}$$

- Under $\mathcal{P}$: $e_{A^\pi \wedge B}(S)$ concentrates on $\binom{|S|}{2} ps^2$ uniformly over all S when $|S|$ is large
- On this typical event $\mathcal{E}$ under $\mathcal{P}$, when $|F|$ is large,

$$\mathbb{E}_{\mathcal{Q}}\left[Y \mathbf{1}_{\{\mathcal{E}\}}\right] \lesssim \mathbb{E}_{\mathcal{Q}}\left[\left(\frac{1}{p}\right)^{2e_{A^\pi \wedge B}(F)} \mathbf{1}_{\left\{e_{A^\pi \wedge B}(F) \le \binom{|F|}{2} ps^2\right\}}\right]$$

Truncated second-moment: sparse Erdős-Rényi

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

- Long orbits:

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

- Short incomplete orbits:

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

- 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 $H_k \triangleq \cup_{O:|O|\le k, O\subset E(A\wedge B^\pi)} O$

Truncated second-moment: sparse Erdős-Rényi

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

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

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

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

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

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

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

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

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 partial recovery thresholds
- Proof of exact recovery thresholds

Proof of partial recovery thresholds: Positive results

- Gaussian or dense Erdős-Rényi: analyzing MLE (QAP)

$$\widehat{\pi}_{\mathrm{ML}} \in \arg\max_{\pi \in \mathcal{S}_n} \sum_{i<j} A_{ij} B_{\pi(i)\pi(j)}$$

- Sparse Erdős-Rényi: Analyze densest subgraph in intersection graph $A^{\pi} \wedge B$

Proof of negative results: Gaussian model

1. Characterization of mutual info by truncated 2nd moment method:

$$I(\pi^*; A, B) = \binom{n}{2} I(\rho) - \mathsf{KL}(\mathcal{P} \| \mathcal{Q}) \approx \binom{n}{2} I(\rho)$$

where $I(\rho) = \frac{1}{2} \log \frac{1}{1-\rho^2}$ is the mutual info between two ρ-correlated standard Gaussians

Proof of negative results: Gaussian model

1. Characterization of mutual info by truncated 2nd moment method:

$$I(\pi^*; A, B) = \binom{n}{2} I(\rho) - \mathsf{KL}(\mathcal{P} \| \mathcal{Q}) \approx \binom{n}{2} I(\rho)$$

where $I(\rho) = \frac{1}{2} \log \frac{1}{1-\rho^2}$ is the mutual info between two ρ-correlated standard Gaussians

2. An area theorem based on I-MMSE formula [Guo-Shamai-Verdú '05]

$$I(\pi^*; A, B) = \frac{1}{2} \int_0^{\rho^2} \frac{\mathsf{mmse}_\theta(A^{\pi^*} | A, B)}{(1-\theta)^2} d\theta$$

Proof of negative results: Gaussian model

1. Characterization of mutual info by truncated 2nd moment method:

$$I(\pi^*; A, B) = \binom{n}{2} I(\rho) - \mathsf{KL}(\mathcal{P} \| \mathcal{Q}) \approx \binom{n}{2} I(\rho)$$

where $I(\rho) = \frac{1}{2} \log \frac{1}{1-\rho^2}$ is the mutual info between two ρ-correlated standard Gaussians

2. An area theorem based on I-MMSE formula [Guo-Shamai-Verdú '05]

$$I(\pi^*; A, B) = \frac{1}{2} \int_0^{\rho^2} \frac{\mathsf{mmse}_\theta(A^{\pi^*} | A, B)}{(1-\theta)^2} d\theta$$

3. Impossibility of estimating A^{π^*} in squared error
$\Longrightarrow$ Impossibility of estimating π^* in overlap

Proof of negative results: Erdős-Rényi

1. Characterization of mutual info by truncated 2nd moment method:

$$I(\pi^*; A, B) = \binom{n}{2} I(p, s) - \mathsf{KL}(\mathcal{P} \| \mathcal{Q}) \approx \binom{n}{2} I(p, s)$$

Proof of negative results: Erdős-Rényi

1. Characterization of mutual info by truncated 2nd moment method:

$$I(\pi^*; A, B) = \binom{n}{2} I(p, s) - \mathsf{KL}(\mathcal{P} \| \mathcal{Q}) \approx \binom{n}{2} I(p, s)$$

2. An "approximate" area theorem: Find an interpolating model P_θ such that $P_0 = Q$ and $P_1 = P$, and

$$I(\pi^*; A, B) \approx \binom{n}{2} I(p, s) - s \int_0^1 \theta \cdot \mathsf{mmse}_\theta(A^{\pi^*} | A, B) \mathrm{d}\theta$$

Proof of negative results: Erdős-Rényi

1. Characterization of mutual info by truncated 2nd moment method:

$$I(\pi^*; A, B) = \binom{n}{2} I(p, s) - \mathsf{KL}(\mathcal{P} \| \mathcal{Q}) \approx \binom{n}{2} I(p, s)$$

2. An "approximate" area theorem: Find an interpolating model P_θ such that $P_0 = Q$ and $P_1 = P$, and

$$I(\pi^*; A, B) \approx \binom{n}{2} I(p, s) - s \int_0^1 \theta \cdot \mathsf{mmse}_\theta(A^{\pi^*} | A, B) \mathrm{d}\theta$$

3. Impossibility of estimating A^{π^*} in squared error
$\implies$ Impossibility of estimating π^* in overlap

Analysis

- Proof of detection thresholds
- Proof of partial recovery thresholds
- Proof of exact recovery thresholds

Proof of exact recovery thresholds: Positive results

- Decompose the likelihood difference via edge orbits

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

- 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

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

Proof of exact recovery thresholds: Negative results

- Suffices to show the failure of MLE
- The bottleneck happens when π' differs from π by a 2-node orbit (i,j), for which the likelihood difference simplifies to

$$\left\langle A^{\pi} - A^{\pi^*}, B \right\rangle = -\sum_{k \in [n]\setminus\{i,j\}} \left(A_{\pi^*(i)\pi^*(k)} - A_{\pi^*(j)\pi^*(k)} \right) \left(B_{ik} - B_{jk} \right)$$

- Prove the existence of (i,j) for which $\langle A^{\pi} - A^{\pi^*}, B \rangle \geq 0$ whp

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.