Lecture 3: Random Graph Matching: Efficient Algorithms

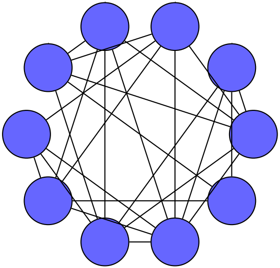
Jiaming Xu

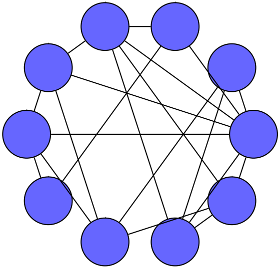
The Fuqua School of Business
Duke University

Joint work with
Jian Ding (PKU), Zhou Fan (Yale), Zongming Ma (Penn)
Cheng Mao (GTech), and Yihong Wu (Yale)

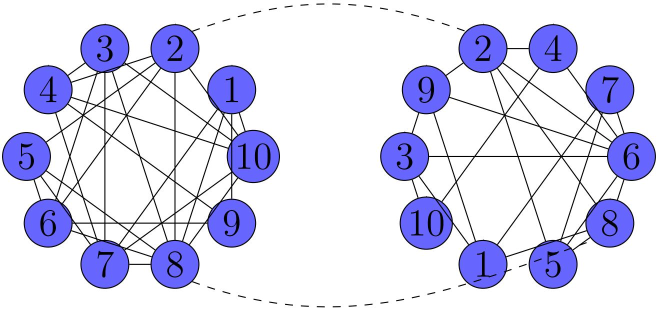
April 20, 2023
AI4OPT Tutorial Lectures

Graph matching (network alignment)



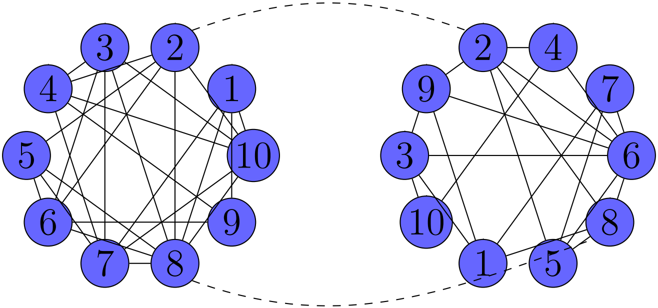


Graph matching (network alignment)



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

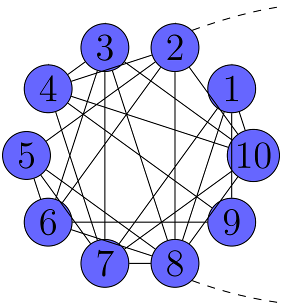
Graph matching (network alignment)



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

Quadratic Assignment Problem (QAP) : $\max_{\Pi \in S_n} \langle A, \Pi B \Pi^\top \rangle$

Graph matching (network alignment)

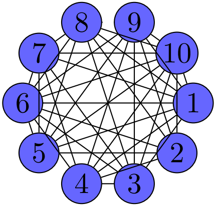


Goal: find a mapping between two node sets that maximally aligns the edges (i.e. minimizes # of adjacency disagreements)

Quadratic Assignment Problem (QAP) : $$\max_{\Pi \in S_n} \langle A, \Pi B \Pi^\top \rangle$$

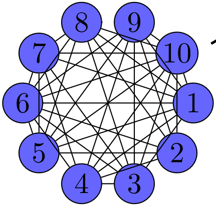
Noiseless case: reduce to graph isomorphism

Correlated Erdős-Rényi graphs model

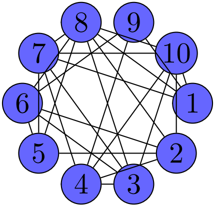


$$G_0 \sim \mathcal{G}(n, p)$$

Correlated Erdős-Rényi graphs model



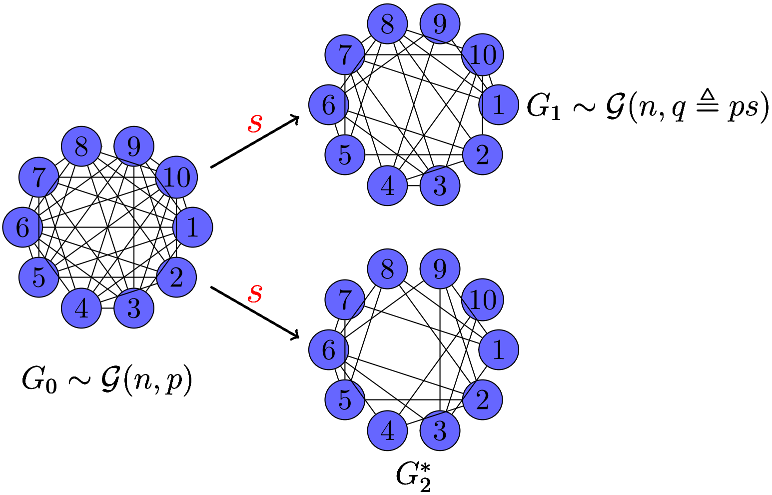
s



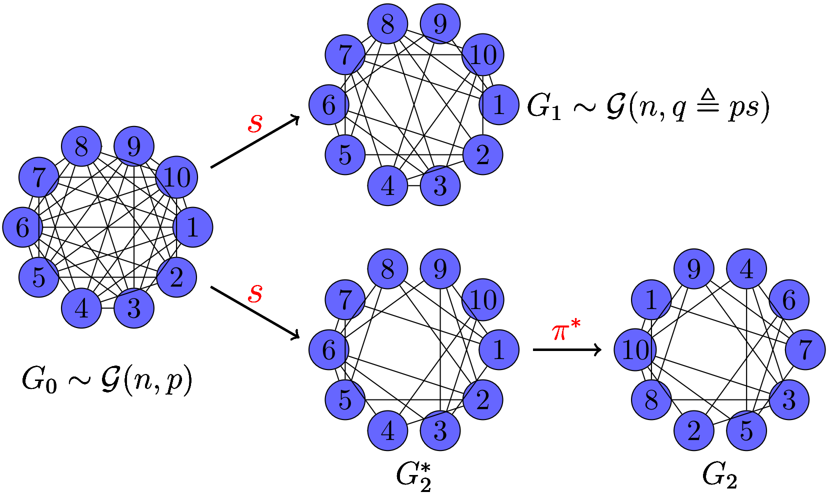
$G_1 \sim \mathcal{G}(n, q \triangleq ps)$

$G_0 \sim \mathcal{G}(n, p)$

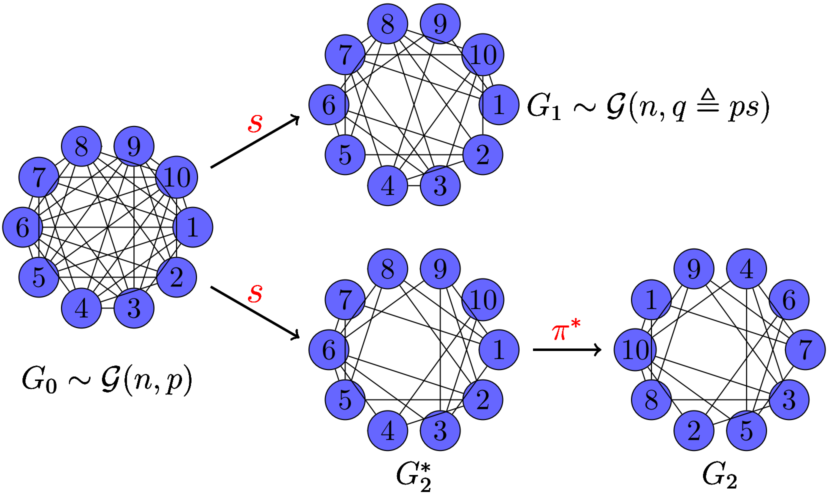
Correlated Erdős-Rényi graphs model



Correlated Erdős-Rényi graphs model



Correlated Erdős-Rényi graphs model



G_1 and G_2 differ by a fraction $\delta \triangleq 1 - s$ of edges, under the correct node mapping

Information-theoretic limits

p: edge probability $\delta = 1 - s$: fraction of errors (differed edges)

Theorem (Cullina-Kiyavash '18, Wu-X.-Yu' 21)

For $p = o(1)$, exact recovery of π^* is information-theoretically possible if and only if

$$nps^2 - \log n \to +\infty$$

Interpretation: Intersection graph $G_1 \wedge G_2^* \sim \mathcal{G}(n, ps^2)$ is connected

Information-theoretic limits

p: edge probability $\delta = 1 - s$: fraction of errors (differed edges)

Theorem (Cullina-Kiyavash '18, Wu-X.-Yu' 21)

For $p = o(1)$, exact recovery of π^* is information-theoretically possible if and only if

$$nps^2 - \log n \to +\infty$$

Interpretation: Intersection graph $G_1 \wedge G_2^* \sim \mathcal{G}(n, ps^2)$ is connected

Computationally:

- Noiseless $s = 1 (\delta = 0)$: optimal condition is attained in linear-time [Bollobás '82, Czajka-Pandurangan '08]
- Noisy case $s < 1 (\delta > 0)$: little is known for efficient algorithms until recently

Main result

p: edge probability $\qquad \delta = 1 - s$: fraction of errors (differed edges)

Theorem (Fan-Mao-Wu-X. '19)

Exact recovery is achieved efficiently by a new spectral method whp if

$$np \gtrsim (\log n)^C \quad \textit{and} \quad \delta \lesssim (\log n)^{-C}$$

for some absolute constant C.

Main result

p: edge probability $\qquad \delta = 1 - s$: fraction of errors (differed edges)

Theorem (Fan-Mao-Wu-X. '19)

Exact recovery is achieved efficiently by a new spectral method whp if

$$np \gtrsim (\log n)^C \qquad \text{and} \qquad \delta \lesssim (\log n)^{-C}$$

for some absolute constant C.

- Classical spectral methods require $\delta \leq n^{-C}$ [Ganassali-Lelarge-Massoulié '19]

Main result

p: edge probability $\qquad \delta = 1 - s$: fraction of errors (differed edges)

Theorem (Fan-Mao-Wu-X. '19)

Exact recovery is achieved efficiently by a new spectral method whp if

$$np \gtrsim (\log n)^C \quad \textit{and} \quad \delta \lesssim (\log n)^{-C}$$

for some absolute constant C.

- Classical spectral methods require $\delta \leq n^{-C}$ [Ganassali-Lelarge-Massoulié '19]
- holds for general correlated Wigner model

Outline

1. A new spectral algorithm
2. Analysis
3. Concluding remarks

Spectral methods and applications

Estimate hidden structure using leading eigenvectors of data matrix A

- Planted clique [Alon-Krivelevich-Sudakov '98]
- Planted partition/Stochastic block model [Mcsherry '98] [Massoulié '13] [Bordenave-Lelarge-Massoulié '15]
- Clustering [von-Luxburg-Bousquet-Belkin '05]
- Graphon estimation [Chatterjee '15]
- Matrix completion [Keshavan-Montanari-Oh '09]
- Ranking [Negahban-Oh-Shah '17]

Spectral methods and applications

Estimate hidden structure using leading eigenvectors of data matrix A

- Planted clique [Alon-Krivelevich-Sudakov '98]
- Planted partition/Stochastic block model [Mcsherry '98] [Massoulié '13] [Bordenave-Lelarge-Massoulié '15]
- Clustering [von-Luxburg-Bousquet-Belkin '05]
- Graphon estimation [Chatterjee '15]
- Matrix completion [Keshavan-Montanari-Oh '09]
- Ranking [Negahban-Oh-Shah '17]

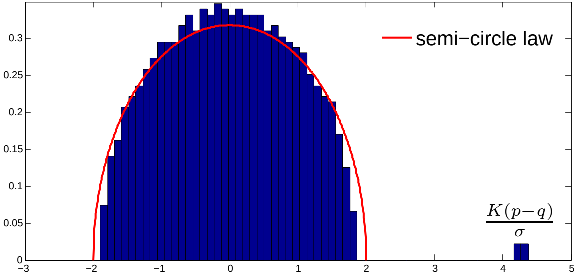
Underlying structure: A is approximately low-rank with large eigen-gap

Analyzing spectral methods: an example

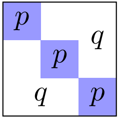
Community detection: $A = \begin{bmatrix} p & & q \\ & p & \\ q & & p \end{bmatrix} + A - \mathbb{E}[A]$

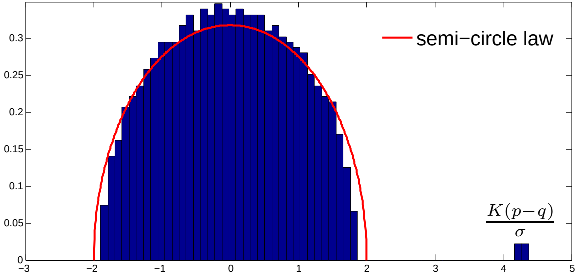
Analyzing spectral methods: an example

Community detection: $A \quad = \quad \begin{bmatrix} p & & q \\ & p & \\ q & & p \end{bmatrix} \quad + A - \mathbb{E}[A]$



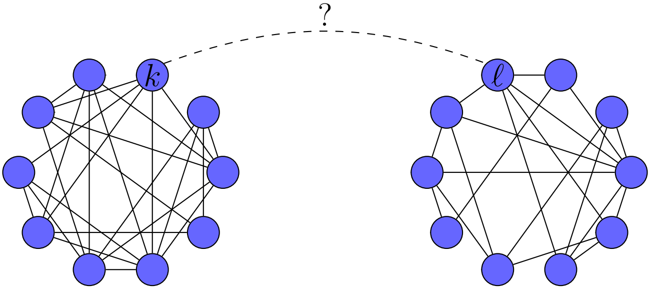
Analyzing spectral methods: an example

Community detection: $A \quad = $  $+ \; A - \mathbb{E}[A]$



- Davis-Kahan and variants: Top eigenvectors of $A \approx$ those of $\mathbb{E}[A]$, if eigen-gap $\gtrsim \|A - \mathbb{E}[A]\|_2$
- However, adjacency matrix of Erdős-Rényi graph has full rank and vanishing eigen-gaps

Spectral graph matching paradigm



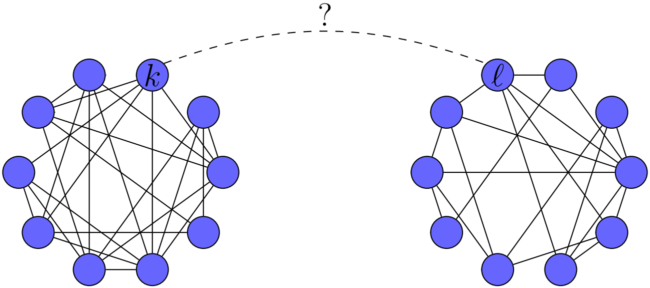
$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^\top$$

$$\lambda_1 \geq \cdots \geq \lambda_n$$

$$B = \sum_{j=1}^{n} \mu_j v_j v_j^\top$$

$$\mu_1 \geq \cdots \geq \mu_n$$

Spectral graph matching paradigm



$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^\top \qquad B = \sum_{j=1}^{n} \mu_j v_j v_j^\top$$

$$\lambda_1 \geq \cdots \geq \lambda_n \qquad \mu_1 \geq \cdots \geq \mu_n$$

1. Construct a similarity matrix X based on (λ_i, u_i) and (μ_j, v_j)
2. Project X to permutation by linear assignment: $\widehat{\Pi} \in \arg\max \langle X, \Pi \rangle$

Failure of previous spectral methods

- Low-rank methods: Aligning the leading eigenvectors

$$X = s_1 u_1 v_1^\top, \qquad s_1 \in \{\pm 1\}$$

Failure of previous spectral methods

- Low-rank methods: Aligning the leading eigenvectors

$$X = s_1 u_1 v_1^\top, \qquad s_1 \in \{\pm 1\}$$

Similar ideas used in IsoRank [Singh-Xu-Berger '08] and EigenAlign [Feizi-Quon-Mendoza-Medard-Kellis-Jadbabaie '19]

Failure of previous spectral methods

- Low-rank methods: Aligning the leading eigenvectors

$$X = s_1 u_1 v_1^\top, \qquad s_1 \in \{\pm 1\}$$

Similar ideas used in IsoRank [Singh-Xu-Berger '08] and EigenAlign [Feizi-Quon-Mendoza-Medard-Kellis-Jadbabaie '19]

- Full-rank methods: [Umeyama '88]

$$X = \sum_{i=1}^{n} s_i u_i v_i^\top, \qquad s_i \in \{\pm 1\}$$

Failure of previous spectral methods

- Low-rank methods: Aligning the leading eigenvectors

$$X = s_1 u_1 v_1^\top, \qquad s_1 \in \{\pm 1\}$$

Similar ideas used in IsoRank [Singh-Xu-Berger '08] and EigenAlign [Feizi-Quon-Mendoza-Medard-Kellis-Jadbabaie '19]

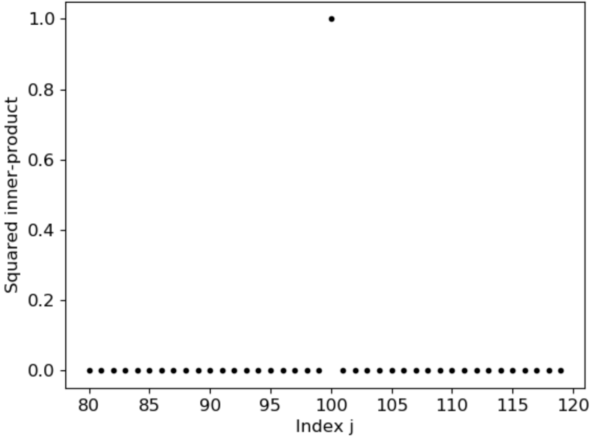
- Full-rank methods: [Umeyama '88]

$$X = \sum_{i=1}^{n} s_i u_i v_i^\top, \qquad s_i \in \{\pm 1\}$$

- All perform well with no noise, but are extremely fragile with noise
- A and B have full rank and vanishing eigen-gaps $\Rightarrow$ decorrelation of u_i and v_i when $\delta = n^{-c}$

Eigenvector correlation decay

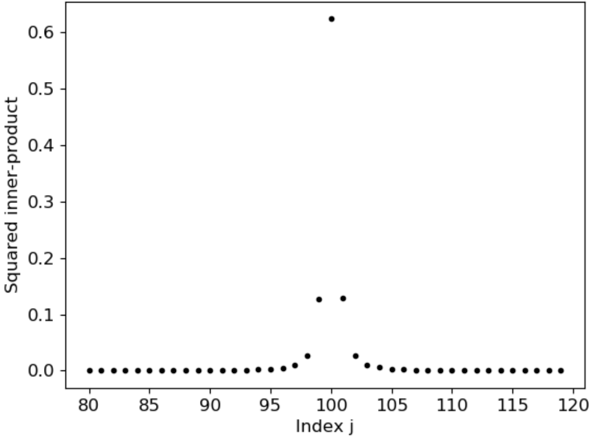
Isomorphic Erdős-Rényi graphs: 500 vertices, edge probability $\frac{1}{2}$



$\langle u_{100}, v_j \rangle^2$ for $j \in \{80, \ldots, 120\}$, averaged across 1000 simulations

Eigenvector correlation decay

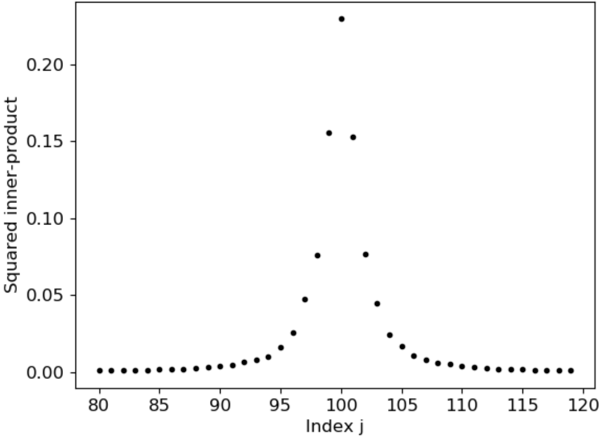
Erdős-Rényi graphs with $\delta = 0.1\%$ differed edges



$\langle u_{100}, v_j \rangle^2$ for $j \in \{80, \ldots, 120\}$, averaged across 1000 simulations

Eigenvector correlation decay

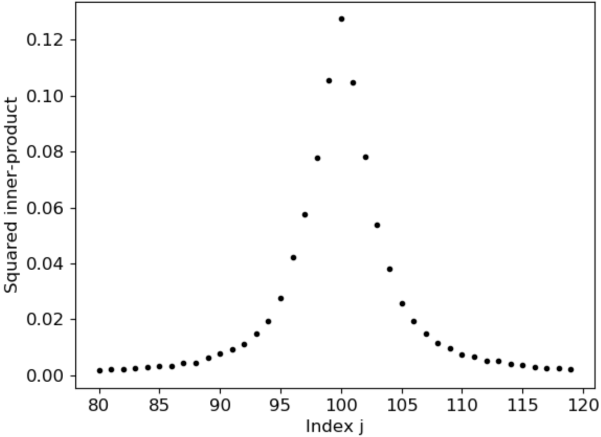
Erdős-Rényi graphs with $\delta = 0.5\%$ differed edges



$\langle u_{100}, v_j \rangle^2$ for $j \in \{80, \ldots, 120\}$, averaged across 1000 simulations

Eigenvector correlation decay

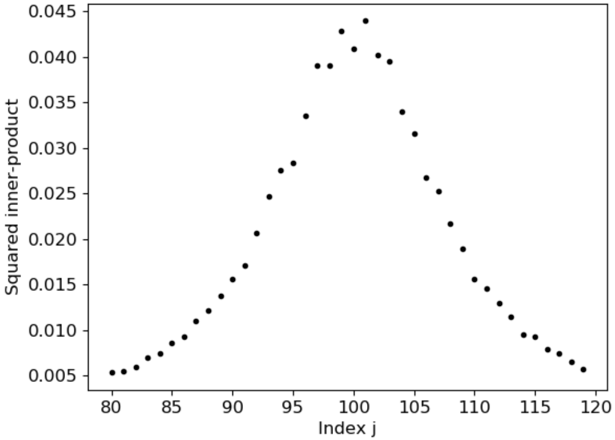
Erdős-Rényi graphs with $\delta = 1\%$ differed edges



$\langle u_{100}, v_j \rangle^2$ for $j \in \{80, \ldots, 120\}$, averaged across 1000 simulations

Eigenvector correlation decay

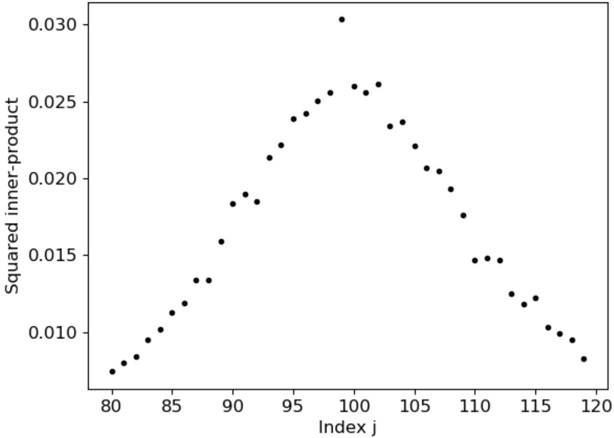
Erdős-Rényi graphs with $\delta = 3\%$ differed edges



$\langle u_{100}, v_j \rangle^2$ for $j \in \{80, \ldots, 120\}$, averaged across 1000 simulations

Eigenvector correlation decay

Erdős-Rényi graphs with $\delta = 5\%$ differed edges



$\langle u_{100}, v_j \rangle^2$ for $j \in \{80, \ldots, 120\}$, averaged across 1000 simulations

A new spectral method: GRAMPA

GRAph Matching by Pairwise eigen-Alignments:

$$X = \sum_{i,j=1}^{n} \underbrace{K\left(\frac{\lambda_i - \mu_j}{\eta}\right)}_{\text{spectral weights}} \times \underbrace{u_i^\top \mathbf{J} v_j \cdot u_i v_j^\top}_{\text{"Alignment" between } u_i \text{ and } v_j}$$

where η = bandwidth parameter, $\mathbf{J}$ = all-one matrix

A new spectral method: GRAMPA

GRAph Matching by Pairwise eigen-Alignments:

$$X = \sum_{i,j=1}^{n} \underbrace{\frac{\eta}{(\lambda_i - \mu_j)^2 + \eta^2}}_{\text{spectral weights}} \times \underbrace{u_i^\top \mathbf{J} v_j \cdot u_i v_j^\top}_{\text{``Alignment'' between } u_i \text{ and } v_j}$$

where η = bandwidth parameter, $\mathbf{J}$ = all-one matrix

A new spectral method: GRAMPA

GRAph Matching by Pairwise eigen-Alignments:

$$X = \sum_{i,j=1}^{n} \underbrace{\frac{\eta}{(\lambda_i - \mu_j)^2 + \eta^2}}_{\text{spectral weights}} \times \underbrace{u_i^\top \mathbf{J} v_j \cdot u_i v_j^\top}_{\text{"Alignment" between } u_i \text{ and } v_j}$$

where η = bandwidth parameter, $\mathbf{J}$ = all-one matrix

- All pairs matter:
  - Spectral weight penalizes pairs whose eigenvalues are far apart
  - Cauchy weight kernel is inspired by the eigenvector correlation decay [Bourgade-Yau '17], [Benigni '17]:

$$n \cdot \mathbb{E}\left[\langle u_i, v_j \rangle^2\right] \approx \frac{\delta}{(\lambda_i - \mu_j)^2 + C\delta^2}$$

- GRAMPA is invariant to the choices of signs for u_i and v_j

GRAMPA as regularized QP relaxation

- Graph matching as a quadratic assignment problem (QAP):

$$\arg\max_{\Pi \in S_n} \langle A, \Pi B \Pi^\top \rangle = \arg\min_{\Pi \in S_n} \|A - \Pi B \Pi^\top\|_F^2$$

GRAMPA as regularized QP relaxation

- Graph matching as a quadratic assignment problem (QAP):

$$\arg\max_{\Pi \in S_n} \langle A, \Pi B \Pi^\top \rangle = \arg\min_{\Pi \in S_n} \|A - \Pi B \Pi^\top\|_F^2 = \arg\min_{\Pi \in S_n} \|A\Pi - \Pi B\|_F^2$$

GRAMPA as regularized QP relaxation

- Graph matching as a quadratic assignment problem (QAP):

$$\arg\max_{\Pi \in S_n} \langle A, \Pi B \Pi^\top \rangle = \arg\min_{\Pi \in S_n} \|A - \Pi B \Pi^\top\|_F^2 = \arg\min_{\Pi \in S_n} \|A\Pi - \Pi B\|_F^2$$

- A popular quadratic programming relaxation [Zaslavskiy-Bach-Vert '09], [Aflalo-Bronstein-Kimmel '15], [Lyzinski-Fishkind-Fiori-Vogelstein-Priebe-Sapiro '15]

$$\arg\min_{X \geq 0:\ X\mathbf{1}=\mathbf{1},\ X^\top \mathbf{1}=\mathbf{1}} \|AX - XB\|_F^2 \qquad \text{(QP-DS)}$$

GRAMPA as regularized QP relaxation

- Graph matching as a quadratic assignment problem (QAP):

$$\arg\max_{\Pi \in S_n} \langle A, \Pi B \Pi^\top \rangle = \arg\min_{\Pi \in S_n} \|A - \Pi B \Pi^\top\|_F^2 = \arg\min_{\Pi \in S_n} \|A\Pi - \Pi B\|_F^2$$

- A popular quadratic programming relaxation [Zaslavskiy-Bach-Vert '09], [Aflalo-Bronstein-Kimmel '15], [Lyzinski-Fishkind-Fiori-Vogelstein-Priebe-Sapiro '15]

$$\arg\min_{X \geq 0:\ X\mathbf{1}=\mathbf{1},\ X^\top \mathbf{1}=\mathbf{1}} \|AX - XB\|_F^2 \qquad \text{(QP-DS)}$$

- The GRAMPA similarity matrix X is (a multiple of)

$$\arg\min_{X:\ \mathbf{1}^\top X \mathbf{1}=n} \|AX - XB\|_F^2 + \eta^2 \|X\|_F^2$$

This further relaxes the DS constraint and adds a ridge regularizer

Analysis of GRAMPA

Diagonal dominance in population version

Question: Is X "close" to true permutation matrix Π^*?

Diagonal dominance in population version

Question: Is X "close" to true permutation matrix Π^*?

Consider the "population version" of the regularized QP:

$$X \propto \arg \min_{X:\, \mathbf{1}^\top X \mathbf{1} = n} \|AX - XB\|_F^2 + \eta^2 \|X\|_F^2$$

Diagonal dominance in population version

Question: Is X "close" to true permutation matrix Π^*?

Consider the "population version" of the regularized QP:

$$X_{\mathsf{pop}} = \arg\min_{X:\ \mathbf{1}^\top X \mathbf{1} = n} \mathbb{E}\left[\|AX - XB\|_F^2\right] + \eta^2 \|X\|_F^2$$

Diagonal dominance in population version

Question: Is X "close" to true permutation matrix Π^*?

Consider the "population version" of the regularized QP:

$$X_{\text{pop}} = \arg \min_{X:\, \mathbf{1}^\top X \mathbf{1} = n} \mathbb{E}\left[\|AX - XB\|_F^2\right] + \eta^2 \|X\|_F^2$$

Assume $\Pi^* = \mathbf{I}$ and $A \leftarrow \frac{A - \mathbb{E}[A]}{\sqrt{nq(1-q)}}$ and $B \leftarrow \frac{B - \mathbb{E}[B]}{\sqrt{nq(1-q)}}$:

$$X_{\text{pop}} = \epsilon \mathbf{I} + (1-\epsilon)\frac{\mathbf{J}}{n}, \qquad \epsilon \approx \frac{2(1-\delta)}{n(2\delta + \eta^2)}$$

- X_{pop} is close to $\frac{\mathbf{J}}{n}$ (center of the Birkhoff polytope)

Diagonal dominance in population version

Question: Is X "close" to true permutation matrix Π^*?

Consider the "population version" of the regularized QP:

$$X_{\mathsf{pop}} = \arg \min_{X:\, \mathbf{1}^\top X \mathbf{1} = n} \mathbb{E}\left[\|AX - XB\|_F^2\right] + \eta^2 \|X\|_F^2$$

Assume $\Pi^* = \mathbf{I}$ and $A \leftarrow \frac{A - \mathbb{E}[A]}{\sqrt{nq(1-q)}}$ and $B \leftarrow \frac{B - \mathbb{E}[B]}{\sqrt{nq(1-q)}}$:

$$X_{\mathsf{pop}} = \epsilon \mathbf{I} + (1-\epsilon)\frac{\mathbf{J}}{n}, \qquad \epsilon \approx \frac{2(1-\delta)}{n(2\delta + \eta^2)}$$

- X_{pop} is close to $\frac{\mathbf{J}}{n}$ (center of the Birkhoff polytope)
- Same analysis holds for tighter QP-DS

Diagonal dominance in population version

Question: Is X "close" to true permutation matrix Π^*?

Consider the "population version" of the regularized QP:

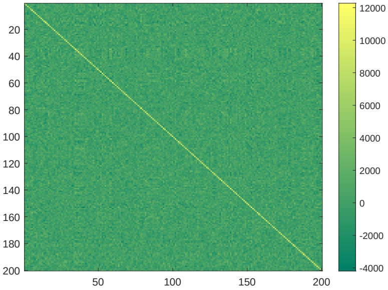
$$X_{\mathsf{pop}} = \arg \min_{X:\, \mathbf{1}^\top X \mathbf{1} = n} \mathbb{E}\left[\|AX - XB\|_F^2\right] + \eta^2 \|X\|_F^2$$

Assume $\Pi^* = \mathbf{I}$ and $A \leftarrow \frac{A - \mathbb{E}[A]}{\sqrt{nq(1-q)}}$ and $B \leftarrow \frac{B - \mathbb{E}[B]}{\sqrt{nq(1-q)}}$:

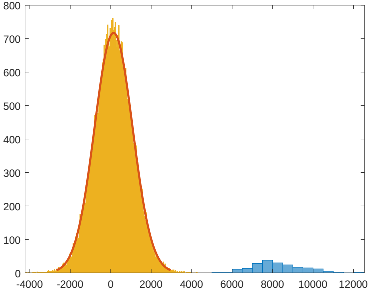
$$X_{\mathsf{pop}} = \epsilon \mathbf{I} + (1-\epsilon)\frac{\mathbf{J}}{n}, \qquad \epsilon \approx \frac{2(1-\delta)}{n(2\delta + \eta^2)}$$

- X_{pop} is close to $\frac{\mathbf{J}}{n}$ (center of the Birkhoff polytope)
- Same analysis holds for tighter QP-DS
- X_{pop} is diagonally dominant: diagonals are $\approx \frac{2}{2\delta + \eta^2}$ times off-diagonals

Diagonal dominance of the similarity matrix



Heatmap of X



Histogram of off-diagonal (orange) and diagonal (blue) entries

When $\Pi^* = \mathbf{I}$, prove **diagonal dominance**

$$\min_k X_{kk} > \max_{k \neq \ell} X_{k\ell}$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

For Gaussian matrices

- eigenvectors are uniform:

$$u_i \sim \text{Uniform}(n\text{-sphere}) \approx \frac{1}{\sqrt{n}} \mathcal{N}(0, I_n)$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

For Gaussian matrices

- eigenvectors are uniform:

$$u_i \sim \text{Uniform}(n\text{-sphere}) \approx \frac{1}{\sqrt{n}} \mathcal{N}(0, I_n)$$

  - Sum: $\langle u_i, \mathbf{1} \rangle \approx 1$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

For Gaussian matrices

- eigenvectors are uniform:

$$u_i \sim \text{Uniform}(n\text{-sphere}) \approx \frac{1}{\sqrt{n}} \mathcal{N}(0, I_n)$$

  - Sum: $\langle u_i, \mathbf{1} \rangle \approx 1$
  - kth entry: $(u_i)_k \approx \mathcal{N}(0, \frac{1}{n})$.

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

For Gaussian matrices

- eigenvectors are uniform:

$$u_i \sim \text{Uniform}(n\text{-sphere}) \approx \frac{1}{\sqrt{n}} \mathcal{N}(0, I_n)$$

  - Sum: $\langle u_i, \mathbf{1} \rangle \approx 1$
  - kth entry: $(u_i)_k \approx \mathcal{N}(0, \frac{1}{n})$.

- eigenvalues satisfies semicircle law:

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i} \approx \rho(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \cdot \mathbf{1}_{\{|x| \le 2\}} \text{ (Wigner's semicircle law)}$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

For Gaussian matrices

- eigenvectors are uniform:

$$u_i \sim \text{Uniform}(n\text{-sphere}) \approx \frac{1}{\sqrt{n}} \mathcal{N}(0, I_n)$$

  - Sum: $\langle u_i, \mathbf{1} \rangle \approx 1$
  - kth entry: $(u_i)_k \approx \mathcal{N}(0, \frac{1}{n})$.
- eigenvalues satisfies semicircle law:

$$\frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i} \approx \rho(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \cdot \mathbf{1}_{\{|x| \le 2\}} \text{ (Wigner's semicircle law)}$$

- eigenvalues and eigenvectors are independent

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i=1}^{n} \frac{1}{\eta} \langle u_i, \mathbf{1} \rangle^2 u_i u_i^\top + \sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i=1}^{n} \frac{1}{\eta} \langle u_i, \mathbf{1} \rangle^2 u_i u_i^\top + \sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

- First term is diagonally dominant:

$$\sum_{i=1}^{n} \frac{1}{\eta} \underbrace{\langle u_i, \mathbf{1} \rangle^2}_{\approx 1} \underbrace{(u_i)_k}_{\mathcal{N}(0, \frac{1}{n})} \underbrace{(u_i)_\ell}_{\mathcal{N}(0, \frac{1}{n})} \approx \begin{cases} \frac{1}{\eta} & \text{if } k = \ell \\ \frac{1}{\eta \sqrt{n}} & \text{if } k \neq \ell \end{cases}$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i=1}^{n} \frac{1}{\eta} \langle u_i, \mathbf{1} \rangle^2 u_i u_i^\top + \sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

- First term is diagonally dominant:

$$\sum_{i=1}^{n} \frac{1}{\eta} \underbrace{\langle u_i, \mathbf{1} \rangle^2}_{\approx 1} \underbrace{(u_i)_k}_{\mathcal{N}(0, \frac{1}{n})} \underbrace{(u_i)_\ell}_{\mathcal{N}(0, \frac{1}{n})} \approx \begin{cases} \frac{1}{\eta} & \text{if } k = \ell \\ \frac{1}{\eta \sqrt{n}} & \text{if } k \neq \ell \end{cases}$$

- Second term is perturbation:

$$\sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle (u_i)_k (u_j)_\ell \approx \sqrt{\frac{1}{n^2} \sum_{i \neq j} \left(\frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \right)^2}$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i=1}^{n} \frac{1}{\eta} \langle u_i, \mathbf{1} \rangle^2 u_i u_i^\top + \sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

- First term is diagonally dominant:

$$\sum_{i=1}^{n} \frac{1}{\eta} \underbrace{\langle u_i, \mathbf{1} \rangle^2}_{\approx 1} \underbrace{(u_i)_k}_{\mathcal{N}(0, \frac{1}{n})} \underbrace{(u_i)_\ell}_{\mathcal{N}(0, \frac{1}{n})} \approx \begin{cases} \frac{1}{\eta} & \text{if } k = \ell \\ \frac{1}{\eta\sqrt{n}} & \text{if } k \neq \ell \end{cases}$$

- Second term is perturbation:

$$\sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle \, (u_i)_k \, (u_j)_\ell \approx \sqrt{\iint \left(\frac{\eta}{(x-y)^2 + \eta^2} \right)^2 d\rho(x) d\rho(y)}$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i=1}^{n} \frac{1}{\eta} \langle u_i, \mathbf{1} \rangle^2 u_i u_i^\top + \sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

- First term is diagonally dominant:

$$\sum_{i=1}^{n} \frac{1}{\eta} \underbrace{\langle u_i, \mathbf{1} \rangle^2}_{\approx 1} \underbrace{(u_i)_k}_{\mathcal{N}(0, \frac{1}{n})} \underbrace{(u_i)_\ell}_{\mathcal{N}(0, \frac{1}{n})} \approx \begin{cases} \frac{1}{\eta} & \text{if } k = \ell \\ \frac{1}{\eta\sqrt{n}} & \text{if } k \neq \ell \end{cases}$$

- Second term is perturbation:

$$\sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle (u_i)_k (u_j)_\ell \approx \sqrt{\iint \left(\frac{\eta}{(x-y)^2 + \eta^2} \right)^2 d\rho(x) d\rho(y)}$$

$$\approx \frac{1}{\sqrt{\eta}} \quad \text{(using semicircle law)}$$

Heuristic argument: noiseless Gaussian case

$$X = \sum_{i=1}^{n} \frac{1}{\eta} \langle u_i, \mathbf{1} \rangle^2 u_i u_i^\top + \sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle u_i u_j^\top$$

- First term is diagonally dominant:

$$\sum_{i=1}^{n} \frac{1}{\eta} \underbrace{\langle u_i, \mathbf{1} \rangle^2}_{\approx 1} \underbrace{(u_i)_k}_{\mathcal{N}(0, \frac{1}{n})} \underbrace{(u_i)_\ell}_{\mathcal{N}(0, \frac{1}{n})} \approx \begin{cases} \frac{1}{\eta} & \text{if } k = \ell \\ \frac{1}{\eta \sqrt{n}} & \text{if } k \neq \ell \end{cases}$$

- Second term is perturbation:

$$\sum_{i \neq j} \frac{\eta}{(\lambda_i - \lambda_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle u_j, \mathbf{1} \rangle (u_i)_k (u_j)_\ell \approx \sqrt{\iint \left(\frac{\eta}{(x-y)^2 + \eta^2} \right)^2 d\rho(x) d\rho(y)}$$

$$\approx \frac{1}{\sqrt{\eta}} \text{ (using semicircle law)}$$

- Made rigorous when A, B are Gaussian since eigenvalues and eigenvectors are independent, but hard to extend to Erdős-Rényi graphs

Universality: Correlated Wigner model

Standardized weighted adjacency matrices A, B where (A_{ij}, B_{ij}) are independent pairs satisfying

$$\mathbb{E}[A_{ij}] = \mathbb{E}[B_{ij}] = 0, \quad \mathbb{E}[A_{ij}^2] = \mathbb{E}[B_{ij}^2] = \frac{1}{n}, \quad \mathbb{E}[A_{ij}B_{ij}] = \frac{1-\delta}{n}$$

Key proof technique: Resolvent and local laws

$$R_A(z) \triangleq (A - z\mathbf{I})^{-1} = \sum_i \frac{1}{\lambda_i - z} u_i u_i^\top, \qquad z \in \mathbb{C} \backslash \mathbb{R}$$

Key proof technique: Resolvent and local laws

$$R_A(z) \triangleq (A - z\mathbf{I})^{-1} = \sum_i \frac{1}{\lambda_i - z} u_i u_i^\top, \qquad z \in \mathbb{C}\backslash\mathbb{R}$$

Denote Wigner's semicircle density and its Stieltjes transform by

$$\rho(x) = \frac{1}{2\pi}\sqrt{4 - x^2} \cdot \mathbf{1}_{\{|x| \le 2\}} \quad \text{and} \quad m(z) = \int \frac{1}{x - z} \rho(x) dx$$

Key proof technique: Resolvent and local laws

$$R_A(z) \triangleq (A - z\mathbf{I})^{-1} = \sum_i \frac{1}{\lambda_i - z} u_i u_i^\top, \qquad z \in \mathbb{C}\backslash\mathbb{R}$$

Denote Wigner's semicircle density and its Stieltjes transform by

$$\rho(x) = \frac{1}{2\pi}\sqrt{4 - x^2} \cdot \mathbf{1}_{\{|x| \le 2\}} \quad \text{and} \quad m(z) = \int \frac{1}{x - z} \rho(x) dx$$

- Empirical eigenvalue distribution converges to ρ

$$\frac{1}{n} \operatorname{Tr} R_A(z) \to m(z)$$

Key proof technique: Resolvent and local laws

$$R_A(z) \triangleq (A - z\mathbf{I})^{-1} = \sum_i \frac{1}{\lambda_i - z} u_i u_i^\top, \qquad z \in \mathbb{C}\backslash\mathbb{R}$$

Denote Wigner's semicircle density and its Stieltjes transform by

$$\rho(x) = \frac{1}{2\pi}\sqrt{4 - x^2} \cdot \mathbf{1}_{\{|x| \le 2\}} \quad \text{and} \quad m(z) = \int \frac{1}{x - z}\rho(x)dx$$

- Empirical eigenvalue distribution converges to ρ

$$\frac{1}{n}\operatorname{Tr} R_A(z) \to m(z)$$

- $R_A(z) \approx m(z)\mathbf{I}$ entrywise [Erdos-Knowles-Yau-Yin '13]:

$$(R_A(z))_{ij} \approx m(z) \cdot \mathbf{1}_{\{i=j\}}$$

Key proof technique: Resolvent and local laws

$$R_A(z) \triangleq (A - z\mathbf{I})^{-1} = \sum_i \frac{1}{\lambda_i - z} u_i u_i^\top, \qquad z \in \mathbb{C}\backslash\mathbb{R}$$

Denote Wigner's semicircle density and its Stieltjes transform by

$$\rho(x) = \frac{1}{2\pi}\sqrt{4 - x^2} \cdot \mathbf{1}_{\{|x| \le 2\}} \quad \text{and} \quad m(z) = \int \frac{1}{x - z}\rho(x)dx$$

- Empirical eigenvalue distribution converges to ρ

$$\frac{1}{n}\operatorname{Tr} R_A(z) \to m(z)$$

- $R_A(z) \approx m(z)\mathbf{I}$ entrywise [Erdos-Knowles-Yau-Yin '13]:

$$(R_A(z))_{ij} \approx m(z) \cdot \mathbf{1}_{\{i=j\}}$$

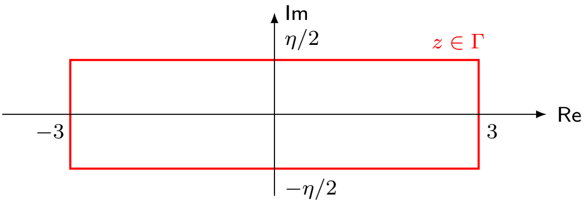
- Similarly, row-sum and total-sum satisfy:

$$\sum_j (R_A(z))_{ij} \lesssim \text{polylog}(n) \qquad \sum_{i,j} (R_A(z))_{ij} \approx n \cdot m(z)$$

Universality proof step 1: Resolvent representation

Lemma (Fan-Mao-Wu-X. '19)

$$X \triangleq \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \mu_j)^2 + \eta^2} \langle u_i, \mathbf{1} \rangle \langle v_j, \mathbf{1} \rangle u_i v_j^\top$$
$$= \frac{1}{2\pi} \operatorname{Re} \oint_\Gamma R_A(z) \mathbf{1}\mathbf{1}^\top R_B(z + \mathbf{i}\eta) dz$$



Γ encloses $\lambda_1, \ldots, \lambda_n$ but not $\mu_1 - \mathbf{i}\eta, \ldots, \mu_n - \mathbf{i}\eta$

Step 2: Leave-one-out relation

$$X_{11} = \frac{1}{2\pi} \mathrm{Re} \oint_{\Gamma} \underbrace{\left[e_1^\top R_A(z)\mathbf{1}\right] \left[\mathbf{1}^\top R_B(z + \mathbf{i}\eta)e_1\right]}_{\text{correlated 1st row sums}} dz$$

Step 2: Leave-one-out relation

$$X_{11} = \frac{1}{2\pi} \mathrm{Re} \oint_{\Gamma} \underbrace{\left[e_1^\top R_A(z)\mathbf{1}\right] \left[\mathbf{1}^\top R_B(z+\mathbf{i}\eta)e_1\right]}_{\text{correlated 1st row sums}} dz$$

-
$$A = \begin{pmatrix} a_{11} & a_1^\top \\ a_1 & A^{(1)} \end{pmatrix} \qquad R_A(z) = \begin{pmatrix} R_{A,11} & R_{A,1*} \\ R_{A,*1} & R_{A,**} \end{pmatrix}$$

Step 2: Leave-one-out relation

$$X_{11} = \frac{1}{2\pi} \operatorname{Re} \oint_\Gamma \underbrace{\left[e_1^\top R_A(z)\mathbf{1}\right] \left[\mathbf{1}^\top R_B(z+\mathbf{i}\eta)e_1\right]}_{\text{correlated 1st row sums}} dz$$

- $$A = \begin{pmatrix} a_{11} & a_1^\top \\ a_1 & A^{(1)} \end{pmatrix} \qquad R_A(z) = \begin{pmatrix} R_{A,11} & R_{A,1*} \\ R_{A,*1} & R_{A,**} \end{pmatrix}$$
- By the Schur-complement formula

$$\begin{aligned} R_{A,1*}(z) &= -R_{A,11}(z) \cdot a_1^\top (A^{(1)} - z\mathbf{I})^{-1} \\ &= -R_{A,11}(z) \cdot a_1^\top R_{A^{(1)}}(z) \end{aligned}$$

Step 2: Leave-one-out relation

$$X_{11} = \frac{1}{2\pi} \operatorname{Re} \oint_\Gamma \underbrace{\left[e_1^\top R_A(z)\mathbf{1}\right] \left[\mathbf{1}^\top R_B(z+\mathbf{i}\eta)e_1\right]}_{\text{correlated 1st row sums}} dz$$

- $$A = \begin{pmatrix} a_{11} & a_1^\top \\ a_1 & A^{(1)} \end{pmatrix} \qquad R_A(z) = \begin{pmatrix} R_{A,11} & R_{A,1*} \\ R_{A,*1} & R_{A,**} \end{pmatrix}$$
- By the Schur-complement formula

$$\begin{aligned} R_{A,1*}(z) &= -R_{A,11}(z) \cdot a_1^\top (A^{(1)} - z\mathbf{I})^{-1} \\ &\approx -m(z) \cdot a_1^\top R_{A^{(1)}}(z) \end{aligned}$$

Step 2: Leave-one-out relation

$$X_{11} = \frac{1}{2\pi} \mathrm{Re} \oint_\Gamma \underbrace{\left[e_1^\top R_A(z)\mathbf{1}\right]\left[\mathbf{1}^\top R_B(z+\mathbf{i}\eta)e_1\right]}_{\text{correlated 1st row sums}} dz$$

- $$A = \begin{pmatrix} a_{11} & a_1^\top \\ a_1 & A^{(1)} \end{pmatrix} \qquad R_A(z) = \begin{pmatrix} R_{A,11} & R_{A,1*} \\ R_{A,*1} & R_{A,**} \end{pmatrix}$$
- By the Schur-complement formula
$$R_{A,1*}(z) = -R_{A,11}(z) \cdot a_1^\top (A^{(1)} - z\mathbf{I})^{-1}$$
$$\approx -m(z) \cdot a_1^\top R_{A^{(1)}}(z)$$
- Writing a similar expression for B, we get
$$X_{11} \approx \frac{1}{2\pi} \mathrm{Re} \oint_\Gamma m(z)m(z+\mathbf{i}\eta) \left[a_1^\top R_{A^{(1)}}(z)\mathbf{1}\mathbf{1}^\top R_{B^{(1)}}(z+\mathbf{i}\eta)b_1\right] dz$$

Step 2: Leave-one-out relation

$$X_{11} = \frac{1}{2\pi} \mathrm{Re} \oint_\Gamma \underbrace{\left[e_1^\top R_A(z)\mathbf{1}\right] \left[\mathbf{1}^\top R_B(z+\mathbf{i}\eta)e_1\right]}_{\text{correlated 1st row sums}} dz$$

- $$A = \begin{pmatrix} a_{11} & a_1^\top \\ a_1 & A^{(1)} \end{pmatrix} \qquad R_A(z) = \begin{pmatrix} R_{A,11} & R_{A,1*} \\ R_{A,*1} & R_{A,**} \end{pmatrix}$$
- By the Schur-complement formula
$$\begin{aligned} R_{A,1*}(z) &= -R_{A,11}(z) \cdot a_1^\top (A^{(1)} - z\mathbf{I})^{-1} \\ &\approx -m(z) \cdot a_1^\top R_{A^{(1)}}(z) \end{aligned}$$
- Writing a similar expression for B, we get
$$X_{11} \approx \frac{1}{2\pi} \mathrm{Re} \oint_\Gamma m(z)m(z+\mathbf{i}\eta) \left[a_1^\top R_{A^{(1)}}(z)\mathbf{1}\mathbf{1}^\top R_{B^{(1)}}(z+\mathbf{i}\eta)b_1\right] dz$$
- The vectors (a_1, b_1) are correlated, and independent of $(A^{(1)}, B^{(1)})$

Step 3: Separating signal from noise

Diagonal entries: Apply concentration of the bilinear form

$$X_{11} \approx \frac{1}{2\pi} \mathrm{Re}\; a_1^\top \left[\oint_\Gamma m(z)m(z+\mathbf{i}\eta)R_{A^{(1)}}(z)\mathbf{1}\mathbf{1}^\top R_{B^{(1)}}(z+\mathbf{i}\eta)dz \right] b_1$$

Step 3: Separating signal from noise

Diagonal entries: Apply concentration of the bilinear form

$$X_{11} \approx \frac{1}{2\pi} \operatorname{Re} \underbrace{a_1^\top \left[\oint_\Gamma m(z) m(z + \mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{1}\mathbf{1}^\top R_{B^{(1)}}(z + \mathbf{i}\eta) dz \right] b_1}_{\approx \frac{1-\delta}{n} \operatorname{Tr}\left[\oint_\Gamma m(z) m(z+\mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{J} R_{B^{(1)}}(z+\mathbf{i}\eta) dz\right]}$$

Step 3: Separating signal from noise

Diagonal entries: Apply concentration of the bilinear form

$$X_{11} \approx \frac{1}{2\pi} \operatorname{Re} \underbrace{a_1^\top \left[\oint_\Gamma m(z)m(z+\mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{1}\mathbf{1}^\top R_{B^{(1)}}(z+\mathbf{i}\eta) dz \right] b_1}_{\approx \frac{1-\delta}{n} \operatorname{Tr}\left[\oint_\Gamma m(z)m(z+\mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{J} R_{B^{(1)}}(z+\mathbf{i}\eta) dz\right]}$$

Off-diagonal entries:

$$X_{12} \approx \frac{1}{2\pi} \operatorname{Re} a_1^\top \left[\oint_\Gamma m(z)m(z+\mathbf{i}\eta) R_{A^{(12)}}(z) \mathbf{1}\mathbf{1}^\top R_{B^{(12)}}(z+\mathbf{i}\eta) dz \right] b_2$$

Here (a_1, b_2) are independent, so the conditional mean is 0

Step 3: Separating signal from noise

Diagonal entries: Apply concentration of the bilinear form

$$X_{11} \approx \frac{1}{2\pi} \mathrm{Re} \underbrace{a_1^\top \left[\oint_\Gamma m(z)m(z+\mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{1}\mathbf{1}^\top R_{B^{(1)}}(z+\mathbf{i}\eta) dz \right] b_1}_{\approx \frac{1-\delta}{n} \mathrm{Tr}\left[\oint_\Gamma m(z)m(z+\mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{J} R_{B^{(1)}}(z+\mathbf{i}\eta) dz\right]}$$

Off-diagonal entries:

$$X_{12} \approx \frac{1}{2\pi} \mathrm{Re} \underbrace{a_1^\top \left[\oint_\Gamma m(z)m(z+\mathbf{i}\eta) R_{A^{(12)}}(z) \mathbf{1}\mathbf{1}^\top R_{B^{(12)}}(z+\mathbf{i}\eta) dz \right] b_2}_{\lesssim \frac{(\log n)^{2+\varepsilon}}{n} \left\| \oint_\Gamma m(z)m(z+\mathbf{i}\eta) R_{A^{(12)}}(z) \mathbf{J} R_{B^{(12)}}(z+\mathbf{i}\eta) dz \right\|_F}$$

Here (a_1, b_2) are independent, so the conditional mean is 0

Step 4: Proof of diagonal dominance

- Diagonal entries:

$$
\begin{aligned}
X_{11} &\approx \frac{1-\delta}{2\pi} \mathrm{Re} \frac{1}{n} \mathrm{Tr} \left[\oint_\Gamma m(z) m(z+\mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{J} R_{B^{(1)}}(z+\mathbf{i}\eta) dz \right] \\
&\approx \frac{1-\delta}{2\pi} \mathrm{Re} \frac{1}{\mathbf{i}\eta} \oint_\Gamma m(z) m(z+\mathbf{i}\eta) \left(m(z+\mathbf{i}\eta) - m(z) \right) dz + \frac{\sqrt{\delta}}{\eta^2} \\
&\approx \frac{1-\delta}{\eta} + \frac{\sqrt{\delta}}{\eta^2}
\end{aligned}
$$

Step 4: Proof of diagonal dominance

- Diagonal entries:

$$X_{11} \approx \frac{1-\delta}{2\pi} \operatorname{Re} \frac{1}{n} \operatorname{Tr} \left[\oint_\Gamma m(z) m(z+\mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{J} R_{B^{(1)}}(z+\mathbf{i}\eta) dz \right]$$

$$\approx \frac{1-\delta}{2\pi} \operatorname{Re} \frac{1}{\mathbf{i}\eta} \oint_\Gamma m(z) m(z+\mathbf{i}\eta) \left(m(z+\mathbf{i}\eta) - m(z) \right) dz + \frac{\sqrt{\delta}}{\eta^2}$$

$$\approx \frac{1-\delta}{\eta} + \frac{\sqrt{\delta}}{\eta^2}$$

- Off-diagonal entries

$$X_{12} \lesssim \frac{(\log n)^{2+\varepsilon}}{n} \left\| \oint_\Gamma m(z) m(z+\mathbf{i}\eta) R_{A^{(12)}}(z) \mathbf{J} R_{B^{(12)}}(z+\mathbf{i}\eta) dz \right\|_F$$

$$\lesssim \frac{(\log n)^{2+\varepsilon}}{\sqrt{\eta}}$$

Step 4: Proof of diagonal dominance

- Diagonal entries:

$$X_{11} \approx \frac{1-\delta}{2\pi} \mathrm{Re} \frac{1}{n} \mathrm{Tr} \left[\oint_\Gamma m(z) m(z+\mathbf{i}\eta) R_{A^{(1)}}(z) \mathbf{J} R_{B^{(1)}}(z+\mathbf{i}\eta) dz \right]$$

$$\approx \frac{1-\delta}{2\pi} \mathrm{Re} \frac{1}{\mathbf{i}\eta} \oint_\Gamma m(z) m(z+\mathbf{i}\eta) \left(m(z+\mathbf{i}\eta) - m(z) \right) dz + \frac{\sqrt{\delta}}{\eta^2}$$

$$\approx \frac{1-\delta}{\eta} + \frac{\sqrt{\delta}}{\eta^2}$$

- Off-diagonal entries

$$X_{12} \lesssim \frac{(\log n)^{2+\varepsilon}}{n} \left\| \oint_\Gamma m(z) m(z+\mathbf{i}\eta) R_{A^{(12)}}(z) \mathbf{J} R_{B^{(12)}}(z+\mathbf{i}\eta) dz \right\|_F$$

$$\lesssim \frac{(\log n)^{2+\varepsilon}}{\sqrt{\eta}}$$

- Applying this and a union bound for every $X_{k\ell}$ shows that X is diagonally dominant when

$$\sqrt{\delta} \lesssim \eta \lesssim (\log n)^{-(4+2\varepsilon)}$$

Extensions

- Dense graphs $q = \Theta(1)$: improvement to $\delta \lesssim (\log n)^{-(4+\epsilon)}$
- Gaussian weighted graphs: improvement to $\delta \lesssim (\log n)^{-2}$ by direct analysis
- Similar result for a tighter QP relaxation (this is not automatic!)

$$\arg \max_{X:\, X\mathbf{1}=\mathbf{1}} \|AX - XB\|_F^2 + \eta^2 \|X\|_F^2$$

- Similar results for matching bipartite graphs

Outline

1. Degree profile matching algorithm
2. Spectral graph matching algorithm
3. Concluding remarks

Concluding remarks

- New spectral graph matching algorithm: "full-rank" spectral method

$$X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \mu_j)^2 + \eta^2} u_i u_i^\top \mathbf{J} v_j v_j^\top$$

- Efficiently matches two graphs with average degree $\geq \mathrm{polylog}(n)$ and fraction of differed edges $\leq 1/\mathrm{polylog}(n)$

Open problems

- Theoretical guarantees for QP-DS

$$\min_{X \text{ doubly stochastic}} \|AX - XB\|_F^2$$

Open problems

- Theoretical guarantees for QP-DS

$$\min_{X \text{ doubly stochastic}} \|AX - XB\|_F^2$$

- Other random graphs ensembles, e.g., geometric graphs

References

- J. Ding, Z. Ma, Y. Wu, J. Xu. *Efficient random graph matching via degree profiles*. Probability Theory & Related Fields, arXiv:1811.07821.
- Z. Fan, C. Mao, Y. Wu, J. Xu. *Spectral graph matching and regularized quadratic relaxations I: Algorithm and Gaussian analysis*, Foundations of Computational Mathematics, arxiv:1907.08880.
- Z. Fan, C. Mao, Y. Wu, J. Xu. *Spectral graph matching and regularized quadratic relaxations II: Erdős-Rényi graphs and universality*, Foundations of Computational Mathematics, arxiv:1907.08883.