Spectral methods for planted graph matching

Jiaming Xu

The Fuqua School of Business
Duke University

Joint work with
Zhou Fan, Cheng Mao, Yihong Wu
Department of Statistics and Data Science, Yale

Discussion of Markov Lecture of Laurent Massoulié
October 21, 2019
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): \[
\max_{\Pi \in S_n} \langle A, \Pi B \Pi^T \rangle
\]

Noiseless case: reduce to graph isomorphism
Goal: find a mapping between two node sets that maximally aligns the edges (i.e. minimizes # of adjacency disagreements)
Goal: find a mapping between two node sets that maximally aligns the edges (i.e. minimizes $\#$ of adjacency disagreements)

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

Noiseless case: reduce to graph isomorphisms
Application 1: Network de-anonymization

- Successfully de-anonymized Netflix by matching it to IMDB [Narayanan-Shmatikov '08]
- Correctly identified 30.8% of node pairings between Twitter and Flickr [Narayanan-Shmatikov '09]
Application 1: Network de-anonymization

- Successfully de-anonymize Netflix by matching it to IMDB [Narayanan-Shmatikov '08]
- Correctly identify 30.8% of node pairings between Twitter and Flickr [Narayanan-Shmatikov '09]
Application 2: Protein-protein interaction network

Human network

Mouse network

[Kazemi-Hassani-Grossglauser-Modarres '16]

Ontology: Discover proteins with similar functions across different species based on interaction network topology
Two key challenges

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

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

- NP-hard for matching two general graphs
- Even approximately solving graph matching is NP-hard
  
  [Makarychev-Manokaran-Sviridenko ’10]
- However, real networks are not arbitrary and have latent structures
Beyond computational intractability

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

- **NP-hard** for matching two general graphs
- Even approximately solving graph matching is **NP-hard**
  [Makarychev-Manokaran-Sviridenko '10]
- However, real networks are not arbitrary and have latent structures

Focus of this talk

Planted models for graph matching: **correlated random graphs**
Beyond computational intractability

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

- NP-hard for matching two general graphs
- Even approximately solving graph matching is NP-hard
  [Makarychev-Manokaran-Sviridenko ’10]
- However, real networks are not arbitrary and have latent structures

Focus of this talk

Planted models for graph matching: correlated random graphs

- Focus on correlated Erdős-Rényi graphs model [Pedarsani-Grossglauser ’11]
- Results can be extended to more general correlated models
Correlated Erdős-Rényi graph model

- \((A_{ij}, B^*_{ij}) \in \{0, 1\}\) are independent across different pairs \(\{i, j\}\)
- Marginally \(A_{ij}, B^*_{ij} \sim \text{Bern}(q)\), and

\[
\mathbb{P}[A_{ij} = B^*_{ij} = 1] = (1 - \delta)q
\]
Correlated Erdős-Rényi graph model

- $(A_{ij}, B^*_{ij}) \in \{0, 1\}$ are independent across different pairs $\{i, j\}$
- Marginally $A_{ij}, B^*_{ij} \sim \text{Bern}(q)$, and
  \[
  \mathbb{P}[A_{ij} = B^*_{ij} = 1] = (1 - \delta)q
  \]
- We observe $A$ and $B$, and seek to recover $\Pi^*$
Main result

$q$: edge probability \hspace{1em} \delta$: fraction of errors (differed edges)

Theorem (Fan-Mao-Wu-X. ’19)

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

\[ nq \gtrsim (\log n)^{48+\epsilon} \hspace{1em} \text{and} \hspace{1em} \delta \lesssim (\log n)^{-(8+\epsilon)} \]
Main result

$q$: edge probability  \quad \delta$: fraction of errors (differed edges)

**Theorem (Fan-Mao-Wu-X. ’19)**

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

\[ nq \gtrsim (\log n)^{48+\epsilon} \quad \text{and} \quad \delta \lesssim (\log n)^{-8-\epsilon} \]

- Previous spectral methods require $\delta \leq n^{-\Omega(1)}$
- Match the best known guarantee for polynomial-time algorithms
  [Ding-Ma-Wu-X. ’18]
- Information-theoretically possible if and only if $nq(1-\delta) - \log n \to \infty$
  [Cullina-Kiyavash ’18]
- Polynomial-time algorithm for constant $\delta$ is open
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

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

\[ \lambda_1 \geq \cdots \geq \lambda_n \]

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

\[ \mu_1 \geq \cdots \geq \mu_n \]

1. Construct a similarity matrix \( X \) based on \((\lambda_i, u_i)\) and \((\mu_j, v_j)\)

2. Project \( X \) to permutation by linear assignment: \( \hat{\Pi} \in \text{arg max} \langle X, \Pi \rangle \)
Low-rank methods: Aligning the leading eigenvectors

\[ X = s_1 u_1 v_1^\top, \quad s_1 \in \{\pm 1\} \]
Failure of previous spectral methods

- **Low-rank methods**: Aligning the leading eigenvectors

  \[ X = s_1 u_1 v_1^\top, \quad s_1 \in \{ \pm 1 \} \]

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

- **Full-rank methods**: [Umeyama '88]

  \[ X = \sum_{i=1}^{n} s_i u_i v_i^\top, \quad s_i \in \{ \pm 1 \} \]

  \( A \) and \( B \) have full rank and vanishing eigen-gaps \( \Rightarrow \) decorrelation of \( u_i \) and \( v_i \) when \( \delta = n - c \)
Failure of previous spectral methods

- **Low-rank methods**: Aligning the leading eigenvectors

\[
X = s_1 u_1 v_1^\top, \quad s_1 \in \{\pm 1\}
\]

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

- **Full-rank methods**: [Umeyama '88]

\[
X = \sum_{i=1}^{n} s_i u_i v_i^\top, \quad s_i \in \{\pm 1\}
\]
Failure of previous spectral methods

- **Low-rank methods:** Aligning the leading eigenvectors
  
  \[ X = s_1 u_1 v_1^\top, \quad s_1 \in \{ \pm 1 \} \]

  Similar ideas used in IsoRank \cite{Singh-Xu-Berger '08} and EigenAlign \cite{Feizi-Quon-Mendoza-Medard-Kellis-Jadbabaie '19}

- **Full-rank methods:** \cite{Umeyama '88}
  
  \[ X = \sum_{i=1}^{n} s_i u_i v_i^\top, \quad s_i \in \{ \pm 1 \} \]

  - 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 $\frac{1}{2}$

\[ \langle u_{100}, v_j \rangle^2 \text{ for } j \in \{80, \ldots, 120\}, \text{ averaged across 1000 simulations} \]
Eigenvector correlation decay

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

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

\[
\langle u_{100}, v_j \rangle^2 \text{ for } j \in \{80, \ldots, 120\}, \text{ averaged across 1000 simulations}
\]
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:**

\[
X = \sum_{i,j=1}^{n} K \left( \frac{\lambda_i - \mu_j}{\eta} \right) \times \left\{ \begin{array}{c}
\text{spectral weights} \\
\text{"Alignment" between } u_i \text{ and } v_j
\end{array} \right\}
\]

where \( \eta = \) bandwidth parameter, \( J = \) all-one matrix
**A new spectral method: GRAMPA**

**GRAph Matching by Pairwise eigen-Alignments:**

\[
X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \mu_j)^2 + \eta^2} \times \begin{pmatrix} u_i u_i^\top J v_j v_j^\top \end{pmatrix} \]

where \( \eta = \) bandwidth parameter, \( J = \) all-one matrix

"Alignment" between \( u_i \) and \( v_j \)
A new spectral method: GRAMPA

**GRAph Matching by Pairwise eigen-Alignments:**

\[
X = \sum_{i,j=1}^{n} \frac{\eta}{(\lambda_i - \mu_j)^2 + \eta^2} \times \begin{pmatrix} u_i u_i^\top & J v_j v_j^\top \end{pmatrix} \text{ “Alignment” between } u_i \text{ and } v_j
\]

where \( \eta = \) bandwidth parameter, \( J = \) all-one matrix

- **All pairs matter:** Cauchy weight kernel is inspired by the eigenvector correlation decay [Bourgade-Yau ’17], [Benigni ’17]:

\[
n \cdot \mathbb{E} [\langle u_i, v_j \rangle^2] \approx \frac{\delta}{(\lambda_i - \mu_j)^2 + C\delta^2}
\]
Analysis of GRAMPA: Diagonal dominance

Heatmap of $X$

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

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

$$\min_k X_{kk} > \max_{k \neq \ell} X_{k\ell}$$
Analysis of GRAMPA: Resolvent and local laws

\[ R_A(z) \triangleq (A - zI)^{-1} = \sum_i \frac{1}{\lambda_i - z} u_i u_i^\top, \quad z \in \mathbb{C} \setminus \mathbb{R} \]
Analysis of GRAMPA: Resolvent and local laws

\[ R_A(z) \triangleq (A - zI)^{-1} = \sum_i \frac{1}{\lambda_i - z} u_i u_i^\top, \quad z \in \mathbb{C}\setminus\mathbb{R} \]

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

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

\[ = \frac{1}{2\pi} \text{Re} \oint\limits_{\Gamma} R_A(z) J R_B(z + i\eta) dz \]
Analysis of GRAMPA: Resolvent and local laws

\[ R_A(z) \triangleq (A - zI)^{-1} = \sum_i \frac{1}{\lambda_i - z} u_i u_i^\top, \quad z \in \mathbb{C} \setminus \mathbb{R} \]

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

= \frac{1}{2\pi} \text{Re} \oint_{\Gamma} R_A(z) J R_B(z + i\eta) \, dz

Show \( X \) is diagonal dominant using \( R_A(z) \approx m(z)I \) entrywise, where \( m(z) \) is the Stieltjes transform of Wigner’s semicircle law
Concluding remarks

- Develop a new spectral graph matching algorithm

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

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

- Still far away from the information-theoretic limits

- Conjecture that a large information-computation gap exists
Concluding remarks

- Develop a new spectral graph matching algorithm

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

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

- Still far away from the information-theoretic limits

- Conjecture that a large information-computation gap exists

- Other planted structures beyond low-rankness:
  - Traveling salespeople (Hamiltonian cycle) [Bagaria-Ding-Tse-Wu-X. '18]
  - Bipartite graph matching (Assignment) [Moharrami-Moore-X. '19]
  - Spanning trees ....