Robust Random Graph Matching in Dense Graphs via Vector Approximate Message Passing

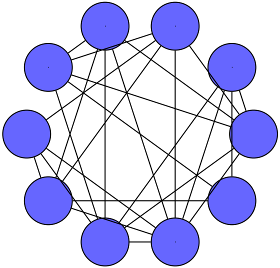
Zhangsong Li

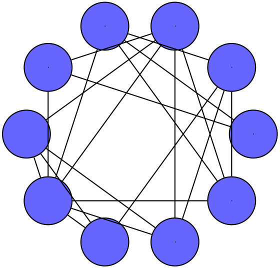
School of Mathematical Sciences, Peking University

June 3, 2025

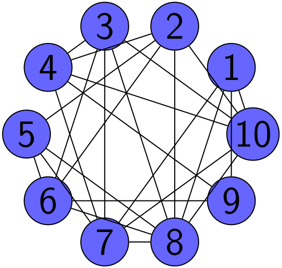
International Conference on Applied Probability

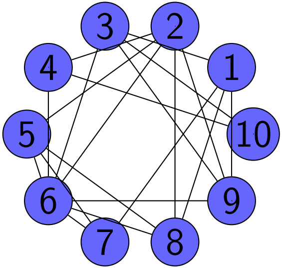
Graph matching (network alignment)





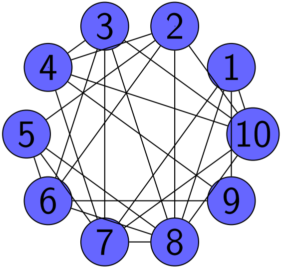
Graph matching (network alignment)

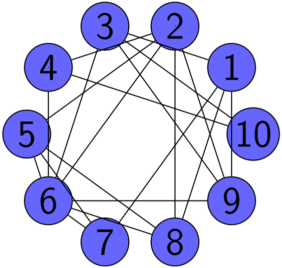




- Goal: find a mapping between two node sets that maximally aligns the edges.

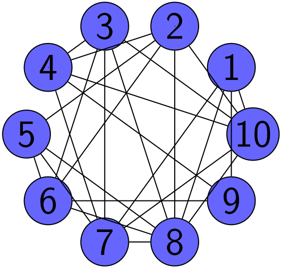
Graph matching (network alignment)

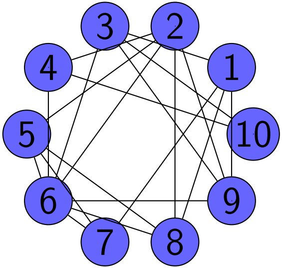




- Goal: find a mapping between two node sets that maximally aligns the edges.
- Quadratic Assignment Problem (QAP): $\max_{\Pi \in \mathfrak{S}_n} \langle A, \Pi B \Pi^\top \rangle$.

Graph matching (network alignment)





- Goal: find a mapping between two node sets that maximally aligns the edges.
- Quadratic Assignment Problem (QAP): $\max_{\Pi \in \mathfrak{S}_n} \langle A, \Pi B \Pi^\top \rangle$.
- NP-hard to solve/approximate in worst case.

An idealistic model: correlated (weighted) random graphs

- π_*: latent permutation on $[n] = \{1, \ldots, n\}$.

An idealistic model: correlated (weighted) random graphs

- π_*: latent permutation on $[n] = \{1, \ldots, n\}$.
- Observation: two *weighted* random graphs A and B, s.t. $(A_{i,j}, B_{\pi_*(i),\pi_*(j)}) \sim \mathbf{F}$.

An idealistic model: correlated (weighted) random graphs

- π_*: latent permutation on $[n] = \{1, \ldots, n\}$.
- Observation: two *weighted* random graphs A and B, s.t. $(A_{i,j}, B_{\pi_*(i),\pi_*(j)}) \sim \mathbf{F}$.
- Two special cases:

An idealistic model: correlated (weighted) random graphs

- π_*: latent permutation on $[n] = \{1, \ldots, n\}$.
- Observation: two *weighted* random graphs A and B, s.t. $(A_{i,j}, B_{\pi_*(i),\pi_*(j)}) \sim \mathbf{F}$.
- Two special cases:
  - *Correlated Gaussian Wigner model.* $\mathbf{F} = \mathcal{N}(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$.

An idealistic model: correlated (weighted) random graphs

- π_*: latent permutation on $[n] = \{1, \dots, n\}$.
- Observation: two *weighted* random graphs A and B, s.t. $(A_{i,j}, B_{\pi_*(i),\pi_*(j)}) \sim \mathbf{F}$.
- Two special cases:
  - *Correlated Gaussian Wigner model.* $\mathbf{F} = \mathcal{N}(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$.
  - *Correlated Erdős-Rényi model.* $\mathbf{F} =$ law of two $\mathbf{Ber}(q)$ with covariance ρ.

An idealistic model: correlated (weighted) random graphs

- π_*: latent permutation on $[n] = \{1, \ldots, n\}$.
- Observation: two *weighted* random graphs A and B, s.t. $(A_{i,j}, B_{\pi_*(i),\pi_*(j)}) \sim \mathbf{F}$.
- Two special cases:
  - *Correlated Gaussian Wigner model.* $\mathbf{F} = \mathcal{N}(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$.
  - *Correlated Erdős-Rényi model.* $\mathbf{F}$ = law of two $\mathbf{Ber}(q)$ with covariance ρ.
- Goal: recover π_* (exactly/partially) using efficient algorithms

An idealistic model: correlated (weighted) random graphs

- π_*: latent permutation on $[n] = \{1, \ldots, n\}$.
- Observation: two *weighted* random graphs A and B, s.t. $(A_{i,j}, B_{\pi_*(i),\pi_*(j)}) \sim \mathbf{F}$.
- Two special cases:
  - *Correlated Gaussian Wigner model.* $\mathbf{F} = \mathcal{N}(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$.
  - *Correlated Erdős-Rényi model.* $\mathbf{F}$ = law of two $\mathbf{Ber}(q)$ with covariance ρ.
- Goal: recover π_* (exactly/partially) using efficient algorithms
  - Noiseless case ($\rho = 1$): optimal condition is attained in linear-time [Bollobás'82, Czajka-Pandurangan'08].

An idealistic model: correlated (weighted) random graphs

- π_*: latent permutation on $[n] = \{1, \ldots, n\}$.
- Observation: two *weighted* random graphs A and B, s.t. $(A_{i,j}, B_{\pi_*(i),\pi_*(j)}) \sim \mathbf{F}$.
- Two special cases:
    - *Correlated Gaussian Wigner model.* $\mathbf{F} = \mathcal{N}(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$.
    - *Correlated Erdős-Rényi model.* $\mathbf{F} =$ law of two $\mathbf{Ber}(q)$ with covariance ρ.
- Goal: recover π_* (exactly/partially) using efficient algorithms
    - Noiseless case ($\rho = 1$): optimal condition is attained in linear-time [Bollobás'82, Czajka-Pandurangan'08].
    - Noisy case ($\rho < 1$): little is known for efficient algorithms until recently.

Recent progress in efficient algorithms

In the case of correlated Erdős-Rényi model with edge-density q and correlation ρ:

Recent progress in efficient algorithms

In the case of correlated Erdős-Rényi model with edge-density q and correlation ρ:

- Progressively improved algorithms have been obtained by community (e.g. [Dai-Cullina-Kiyavash'18, Barak-Chou-Lei-Schramm-Sheng'19, Ding-Ma-Wu-Xu'21, Mao-Rudelson-Tikhomirov'21, Ganassali-Massoulié-Lelarge'22], etc.)

Recent progress in efficient algorithms

In the case of correlated Erdős-Rényi model with edge-density q and correlation ρ:

- Progressively improved algorithms have been obtained by community (e.g. [Dai-Cullina-Kiyavash'18, Barak-Chou-Lei-Schramm-Sheng'19, Ding-Ma-Wu-Xu'21, Mao-Rudelson-Tikhomirov'21, Ganassali-Massoulié-Lelarge'22], etc.)
- The state-of-the-art algorithm:

Recent progress in efficient algorithms

In the case of correlated Erdős-Rényi model with edge-density q and correlation ρ:

- Progressively improved algorithms have been obtained by community (e.g. [Dai-Cullina-Kiyavash'18, Barak-Chou-Lei-Schramm-Sheng'19, Ding-Ma-Wu-Xu'21, Mao-Rudelson-Tikhomirov'21, Ganassali-Massoulié-Lelarge'22], etc.)
- The state-of-the-art algorithm:
  - [Mao-Wu-Xu-Yu'23,24]: polynomial-time algorithm for exact matching when $q > \frac{\log n}{n}$ and correlation $\rho > \sqrt{\alpha}$ where $\alpha \approx 0.338$ is the Otter's constant, based on counting carefully curated family of rooted trees.

Recent progress in efficient algorithms

In the case of correlated Erdős-Rényi model with edge-density q and correlation ρ:

- Progressively improved algorithms have been obtained by community (e.g. [Dai-Cullina-Kiyavash'18, Barak-Chou-Lei-Schramm-Sheng'19, Ding-Ma-Wu-Xu'21, Mao-Rudelson-Tikhomirov'21, Ganassali-Massoulié-Lelarge'22], etc.)
- The state-of-the-art algorithm:
  - [Mao-Wu-Xu-Yu'23,24]: polynomial-time algorithm for exact matching when $q > \frac{\log n}{n}$ and correlation $\rho > \sqrt{\alpha}$ where $\alpha \approx 0.338$ is the Otter's constant, based on counting carefully curated family of rooted trees.
  - [Ding-L.'22+,23+] polynomial-time iterative algorithm for exact matching when $q \geq n^{-1+\delta}$ and correlation $\rho = \Omega(1)$.

Recent progress in efficient algorithms

In the case of correlated Erdős-Rényi model with edge-density q and correlation ρ:

- Progressively improved algorithms have been obtained by community (e.g. [Dai-Cullina-Kiyavash'18, Barak-Chou-Lei-Schramm-Sheng'19, Ding-Ma-Wu-Xu'21, Mao-Rudelson-Tikhomirov'21, Ganassali-Massoulié-Lelarge'22], etc.)
- The state-of-the-art algorithm:
  - [Mao-Wu-Xu-Yu'23,24]: polynomial-time algorithm for exact matching when $q > \frac{\log n}{n}$ and correlation $\rho > \sqrt{\alpha}$ where $\alpha \approx 0.338$ is the Otter's constant, based on counting carefully curated family of rooted trees.
  - [Ding-L.'22+,23+] polynomial-time iterative algorithm for exact matching when $q \geq n^{-1+\delta}$ and correlation $\rho = \Omega(1)$.
- Evidence in [Ding-Du-L.'23+] suggests that the state-of-the-art algorithms have nearly reached the limit of efficient algorithms.

The robustness limitations

The robustness limitations

- [Ameen-Hajek'24]: many efficient random graph matching algorithms will break down if one adversarially modify a small fraction of edges.

The robustness limitations

- [Ameen-Hajek'24]: many efficient random graph matching algorithms will break down if one adversarially modify a small fraction of edges.
- Reason: algorithms based on sophisticated subgraph structures/ delicate spectral properties are not robust under adversarial perturbations (e.g., planting a $\Theta(\sqrt{n})$ clique or other "undesired" subgraphs).

The robustness limitations

- [Ameen-Hajek'24]: many efficient random graph matching algorithms will break down if one adversarially modify a small fraction of edges.
- Reason: algorithms based on sophisticated subgraph structures/ delicate spectral properties are not robust under adversarial perturbations (e.g., planting a $\Theta(\sqrt{n})$ clique or other "undesired" subgraphs).
- **Motivation from application**: somewhat more "practical" graph matching algorithm for real networks?

The robustness limitations

- [Ameen-Hajek'24]: many efficient random graph matching algorithms will break down if one adversarially modify a small fraction of edges.
- Reason: algorithms based on sophisticated subgraph structures/ delicate spectral properties are not robust under adversarial perturbations (e.g., planting a $\Theta(\sqrt{n})$ clique or other "undesired" subgraphs).
- **Motivation from application**: somewhat more "practical" graph matching algorithm for real networks?
- **Motivation from theory**: can we find efficient graph matching algorithms for semi-random models?

The corrupted correlated Gaussian model

The corrupted correlated Gaussian model

- First sample (A, B) from the law of correlated Gaussian Wigner model.

The corrupted correlated Gaussian model

- First sample (A, B) from the law of correlated Gaussian Wigner model.
- Corruption: An adversary can arbitrary choose two subsets $Q, R \subset [n]$ with $|Q|, |R| = \epsilon n$ and arbitrary revise the entries $\{A_{i,j} : i, j \in Q\}$ and $\{B_{i,j} : i, j \in R\}$ according to A, B.

The corrupted correlated Gaussian model

- First sample (A, B) from the law of correlated Gaussian Wigner model.
- Corruption: An adversary can arbitrary choose two subsets $Q, R \subset [n]$ with $|Q|, |R| = \epsilon n$ and arbitrary revise the entries $\{A_{i,j} : i, j \in Q\}$ and $\{B_{i,j} : i, j \in R\}$ according to A, B.
- Observation: the revised matrices $(A', B') = (A + E, B + F)$, where E, F supported on an unknown $\epsilon n * \epsilon n$ principle minor of (A, B).

Our result: a robust Gaussian matching algorithm

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

Theorem (L.'25)

Exact recovery is achieved efficiently by an approximate message passing algorithm w.h.p. if

$$\rho = \Omega(1) \text{ and } \epsilon = o\left(\frac{1}{(\log n)^{20}}\right).$$

Our result: a robust Gaussian matching algorithm

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

Theorem (L.'25)

Exact recovery is achieved efficiently by an approximate message passing algorithm w.h.p. if

$$\rho = \Omega(1) \text{ and } \epsilon = o\left(\frac{1}{(\log n)^{20}}\right).$$

- The first graph matching algorithm that is robust under $n^{1-o(1)}$ size of perturbations.

Our result: a robust Gaussian matching algorithm

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

Theorem (L.'25)

Exact recovery is achieved efficiently by an approximate message passing algorithm w.h.p. if

$$\rho = \Omega(1) \text{ and } \epsilon = o\left(\frac{1}{(\log n)^{20}}\right).$$

- The first graph matching algorithm that is robust under $n^{1-o(1)}$ size of perturbations.
- Extends to the case of correlated Erdős-Rényi models when the edge-density q is a constant.

Approximate message passing (AMP) and applications

A general framework for estimating hidden structures given data matrix A.

Approximate message passing (AMP) and applications

A general framework for estimating hidden structures given data matrix A.

- Compress sensing [Donoho-Maleki-Montanari'09]

Approximate message passing (AMP) and applications

A general framework for estimating hidden structures given data matrix A.

- Compress sensing [Donoho-Maleki-Montanari'09]
- Sparse Principle Component Analysis (PCA) [Deshpande-Montanari'14]

Approximate message passing (AMP) and applications

A general framework for estimating hidden structures given data matrix A.

- Compress sensing [Donoho-Maleki-Montanari'09]
- Sparse Principle Component Analysis (PCA) [Deshpande-Montanari'14]
- Non-negative PCA [Montanari-Richard'14]

Approximate message passing (AMP) and applications

A general framework for estimating hidden structures given data matrix A.

- Compress sensing [Donoho-Maleki-Montanari'09]
- Sparse Principle Component Analysis (PCA) [Deshpande-Montanari'14]
- Non-negative PCA [Montanari-Richard'14]
- Linear regression [Krzakala-Mézard-Sausset-Sun-Zdeborová'12]

Approximate message passing (AMP) and applications

A general framework for estimating hidden structures given data matrix A.

- Compress sensing [Donoho-Maleki-Montanari'09]
- Sparse Principle Component Analysis (PCA) [Deshpande-Montanari'14]
- Non-negative PCA [Montanari-Richard'14]
- Linear regression [Krzakala-Mézard-Sausset-Sun-Zdeborová'12]
- Perceptron models [Ding-Sun'18]

Approximate message passing (AMP) and applications

A general framework for estimating hidden structures given data matrix A.

- Compress sensing [Donoho-Maleki-Montanari'09]
- Sparse Principle Component Analysis (PCA) [Deshpande-Montanari'14]
- Non-negative PCA [Montanari-Richard'14]
- Linear regression [Krzakala-Mézard-Sausset-Sun-Zdeborová'12]
- Perceptron models [Ding-Sun'18]

Usually in the form of the following iteration:

$$\underset{\substack{\uparrow \\ \text{estimator for the hidden signal}}}{f^{(t+1)}} = \underset{\substack{\uparrow \\ \text{entrywise transform by a suitable denoiser}}}{\varphi} \circ \left(\tfrac{1}{\sqrt{n}} A f^{(t)}\right)$$

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Usually, in an AMP algorithm we hope our estimator $f^{(t)}$ converges to the hidden signal (e.g., the hidden matching).

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Usually, in an AMP algorithm we hope our estimator $f^{(t)}$ converges to the hidden signal (e.g., the hidden matching).
- Our strategy: iteratively construct "signatures" using AMP iteration

$$f^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} A' f^{(t)}\right), \quad f^{(t)} = \left(f_1^{(t)}, \ldots, f_n^{(t)}\right)^\top,$$
$$g^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} B' g^{(t)}\right), \quad g^{(t)} = \left(g_1^{(t)}, \ldots, g_n^{(t)}\right)^\top.$$

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Usually, in an AMP algorithm we hope our estimator $f^{(t)}$ converges to the hidden signal (e.g., the hidden matching).
- Our strategy: iteratively construct "signatures" using AMP iteration

$$f^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} A' f^{(t)}\right), \quad f^{(t)} = \left(f_1^{(t)}, \ldots, f_n^{(t)}\right)^\top,$$
$$g^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} B' g^{(t)}\right), \quad g^{(t)} = \left(g_1^{(t)}, \ldots, g_n^{(t)}\right)^\top.$$

- Hope: if we choose a suitable denoiser function φ, then at a large time t^* we will have

$$\Pi_* = \arg\max_{\Pi \in \mathfrak{S}_n} \langle f^{(t^*)}, \Pi g^{(t^*)} \rangle,$$

then we can find Π_* by solving a linear assignment problem.

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Initialization: starts with a constant-size seed $U \xrightarrow{\pi_*} V$ (obtained by brutal-force searching) and let

$$f^{(0)} = \varphi \circ \left(A'_{[n]\times U}\right), \quad g^{(0)} = \varphi \circ \left(B'_{[n]\times V}\right).$$

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Initialization: starts with a constant-size seed $U \xrightarrow{\pi_*} V$ (obtained by brutal-force searching) and let

$$f^{(0)} = \varphi \circ \left(A'_{[n]\times U}\right), \quad g^{(0)} = \varphi \circ \left(B'_{[n]\times V}\right).$$

- Problem in iteration: for any $n * K$ matrix $f^{(t)}, g^{(t)}$

$$f^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} A' f^{(t)}\right), \quad g^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} B' g^{(t)}\right),$$

one can check that the covariance between $f_i^{(t)}$ and $g_{\pi_*(i)}^{(t)}$ must decreases in t (i.e., entrywise signal is decreasing).

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Strategy in the iteration (motivated by [Ding-L.'22,23]): increase the dimension of the signature

$$f^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} A' f^{(t)} \Xi^{(t)}\right), \quad g^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} B' g^{(t)} \Xi^{(t)}\right).$$

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Strategy in the iteration (motivated by [Ding-L.'22,23]): increase the dimension of the signature

$$f^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} A' f^{(t)} \Xi^{(t)}\right), \quad g^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} B' g^{(t)} \Xi^{(t)}\right).$$

- By choosing an appropriate $K_t * K_{t+1}$ matrix $\Xi^{(t)}$ we can guarantee that

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Strategy in the iteration (motivated by [Ding-L.'22,23]): increase the dimension of the signature

$$f^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} A' f^{(t)} \Xi^{(t)}\right), \quad g^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} B' g^{(t)} \Xi^{(t)}\right).$$

- By choosing an appropriate $K_t * K_{t+1}$ matrix $\Xi^{(t)}$ we can guarantee that
  - $\langle f_i^{(t)}, g_{\pi_*(i)}^{(t)} \rangle$ have expectation $\epsilon_t K_t$ and variance K_t (i.e., in true pairs entrywise signal is ϵ_t and signals in each entry are near independent);

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Strategy in the iteration (motivated by [Ding-L.'22,23]): increase the dimension of the signature

$$f^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} A' f^{(t)} \Xi^{(t)}\right), \quad g^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} B' g^{(t)} \Xi^{(t)}\right).$$

- By choosing an appropriate $K_t * K_{t+1}$ matrix $\Xi^{(t)}$ we can guarantee that
  - $\langle f_i^{(t)}, g_{\pi_*(i)}^{(t)} \rangle$ have expectation $\epsilon_t K_t$ and variance K_t (i.e., in true pairs entrywise signal is ϵ_t and signals in each entry are near independent);
  - $\langle f_i^{(t)}, g_j^{(t)} \rangle$ have expectation 0 and variance K_t (i.e., no signal in fake pairs).

Our algorithmic approach

ρ: edge correlation; ϵn: size of corruption; π_*: hidden matching.

- Strategy in the iteration (motivated by [Ding-L.'22,23]): increase the dimension of the signature

$$f^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} A' f^{(t)} \Xi^{(t)}\right), \quad g^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} B' g^{(t)} \Xi^{(t)}\right).$$

- By choosing an appropriate $K_t * K_{t+1}$ matrix $\Xi^{(t)}$ we can guarantee that
  - $\langle f_i^{(t)}, g_{\pi_*(i)}^{(t)} \rangle$ have expectation $\epsilon_t K_t$ and variance K_t (i.e., in true pairs entrywise signal is ϵ_t and signals in each entry are near independent);
  - $\langle f_i^{(t)}, g_j^{(t)} \rangle$ have expectation 0 and variance K_t (i.e., no signal in fake pairs).
  - The signal-to-noise ratio $\frac{(\epsilon_t K_t)^2}{K_t} = \epsilon_t^2 K_t$ grows rapidly in t.

Dealing with adversarial corruption

Input: a spectral cleaning procedure proposed in [Ivkov-Schramm'24].

Theorem (Ivkov-Schramm'24)

Given a matrix $M' = M + E$ with $\|M\|_{\mathrm{op}} = O(\sqrt{n})$ and E supported on an $\epsilon n * \epsilon n$ minor of M, there exists a polynomial-time algorithm that zeros-out $O(\epsilon n)$ rows and columns of M' such that the "cleaned" matrix $\widehat{M}$ satisfies $\|\widehat{M}\|_{\mathrm{op}} = O(\sqrt{n})$.

Dealing with adversarial corruption

Input: a spectral cleaning procedure proposed in [Ivkov-Schramm'24].

Theorem (Ivkov-Schramm'24)

Given a matrix $M' = M + E$ with $\|M\|_{\mathrm{op}} = O(\sqrt{n})$ and E supported on an $\epsilon n * \epsilon n$ minor of M, there exists a polynomial-time algorithm that zeros-out $O(\epsilon n)$ rows and columns of M' such that the "cleaned" matrix $\widehat{M}$ satisfies $\|\widehat{M}\|_{\mathrm{op}} = O(\sqrt{n})$.

- Our method: apply the spectral cleaning procedure to A', B' respectively to obtain $\widehat{A}, \widehat{B}$. Then run the iteration w.r.t. $(\widehat{A}, \widehat{B})$:

$$\widehat{f}^{(0)} = \varphi \circ \left(\widehat{A}_{[n]\times U}\right), \quad \widehat{f}^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} \widehat{A}\widehat{f}^{(t)} \Xi^{(t)}\right),$$
$$\widehat{g}^{(0)} = \varphi \circ \left(\widehat{B}_{[n]\times V}\right), \quad \widehat{g}^{(t+1)} = \varphi \circ \left(\tfrac{1}{\sqrt{n}} \widehat{B}\widehat{g}^{(t)} \Xi^{(t)}\right),$$

Intuitions behind the spectral cleaning

$\widehat{f}^{(t)}$: iterations w.r.t. $\widehat{A}$; $f^{(t)}$: iterations w.r.t. A;
S: index are zeroed-out; E: adversarial corruption; Q: support of E;

Intuitions behind the spectral cleaning

$\widehat{f}^{(t)}$: iterations w.r.t. $\widehat{A}$; $f^{(t)}$: iterations w.r.t. A;
S: index are zeroed-out; E: adversarial corruption; Q: support of E;

- Initialization: $\widehat{f}^{(0)} = \varphi \circ \left(\widehat{A}_{[n]\times U}\right) \approx \varphi \circ \left(A_{[n]\times U}\right) = f^{(0)}$.

Intuitions behind the spectral cleaning

$\widehat{f}^{(t)}$: iterations w.r.t. $\widehat{A}$; $f^{(t)}$: iterations w.r.t. A;

S: index are zeroed-out; E: adversarial corruption; Q: support of E;

- Initialization: $\widehat{f}^{(0)} = \varphi \circ \left(\widehat{A}_{[n]\times U}\right) \approx \varphi \circ \left(A_{[n]\times U}\right) = f^{(0)}$.
- Iteration: suppose $\widehat{f}^{(t)} \approx f^{(t)}$, then

$$\widehat{A}\widehat{f}^{(t)}\Xi^{(t)} - Af^{(t)}\Xi^{(t)} = \widehat{A}(\widehat{f}^{(t)} - f^{(t)})\Xi^{(t)} + (\underbrace{\widehat{A} - A}_{\uparrow \atop E_{Q\setminus S\times Q\setminus S} + A_{[n]\times Q\setminus S} + A_{Q\setminus S\times[n]}})f^{(t)}\Xi^{(t)}$$

Intuitions behind the spectral cleaning

$\widehat{f}^{(t)}$: iterations w.r.t. $\widehat{A}$; $f^{(t)}$: iterations w.r.t. A;
S: index are zeroed-out; E: adversarial corruption; Q: support of E;

- Initialization: $\widehat{f}^{(0)} = \varphi \circ \left(\widehat{A}_{[n]\times U}\right) \approx \varphi \circ \left(A_{[n]\times U}\right) = f^{(0)}$.
- Iteration: suppose $\widehat{f}^{(t)} \approx f^{(t)}$, then

$$\widehat{A}\widehat{f}^{(t)}\Xi^{(t)} - Af^{(t)}\Xi^{(t)} = \widehat{A}(\widehat{f}^{(t)} - f^{(t)})\Xi^{(t)} + (\widehat{A} - A)f^{(t)}\Xi^{(t)}$$

$$\uparrow$$

$$E_{Q\setminus S\times Q\setminus S} + A_{[n]\times Q\setminus S} + A_{Q\setminus S\times[n]}$$

Induction hypothesis $\Longrightarrow (*) \approx 0$;

Intuitions behind the spectral cleaning

$\widehat{f}^{(t)}$: iterations w.r.t. $\widehat{A}$; $f^{(t)}$: iterations w.r.t. A;

S: index are zeroed-out; E: adversarial corruption; Q: support of E;

- Initialization: $\widehat{f}^{(0)} = \varphi \circ \left(\widehat{A}_{[n]\times U}\right) \approx \varphi \circ \left(A_{[n]\times U}\right) = f^{(0)}$.
- Iteration: suppose $\widehat{f}^{(t)} \approx f^{(t)}$, then

$$\widehat{A}\widehat{f}^{(t)}\Xi^{(t)} - Af^{(t)}\Xi^{(t)} = \widehat{A}(\widehat{f}^{(t)} - f^{(t)})\Xi^{(t)} + \underbrace{(\widehat{A} - A)}_{E_{Q\setminus S\times Q\setminus S} + A_{[n]\times Q\setminus S} + A_{Q\setminus S\times [n]}} f^{(t)}\Xi^{(t)}$$

Induction hypothesis $\Longrightarrow (*) \approx 0$;

$|Q \setminus S| \leq \epsilon n \Longrightarrow (*) \approx 0$;

Intuitions behind the spectral cleaning

$\widehat{f}^{(t)}$: iterations w.r.t. $\widehat{A}$; $f^{(t)}$: iterations w.r.t. A;
S: index are zeroed-out; E: adversarial corruption; Q: support of E;

- Initialization: $\widehat{f}^{(0)} = \varphi \circ \left(\widehat{A}_{[n]\times U}\right) \approx \varphi \circ \left(A_{[n]\times U}\right) = f^{(0)}$.
- Iteration: suppose $\widehat{f}^{(t)} \approx f^{(t)}$, then

$$\widehat{A}\widehat{f}^{(t)}\Xi^{(t)} - Af^{(t)}\Xi^{(t)} = \widehat{A}(\widehat{f}^{(t)} - f^{(t)})\Xi^{(t)} + (\widehat{A} - A)f^{(t)}\Xi^{(t)}$$

$$\uparrow$$

$$E_{Q\setminus S\times Q\setminus S} + A_{[n]\times Q\setminus S} + A_{Q\setminus S\times [n]}$$

Induction hypothesis $\Longrightarrow (*) \approx 0$;
$|Q \setminus S| \leq \epsilon n \Longrightarrow (*) \approx 0$;
$\|E_{Q\setminus S\times Q\setminus S}\|_{\mathrm{op}} \leq \|\widehat{A} - A\|_{\mathrm{op}} = O(1) \Longrightarrow (*) \leq O(1)\cdot \|f^{(t)}_{Q\setminus S}\|_{\mathrm{F}} \approx 0$.

Intuitions behind the spectral cleaning

$\widehat{f}^{(t)}$: iterations w.r.t. $\widehat{A}$; $f^{(t)}$: iterations w.r.t. A;

S: index are zeroed-out; E: adversarial corruption; Q: support of E;

- Initialization: $\widehat{f}^{(0)} = \varphi \circ \left(\widehat{A}_{[n]\times U}\right) \approx \varphi \circ \left(A_{[n]\times U}\right) = f^{(0)}$.
- Iteration: suppose $\widehat{f}^{(t)} \approx f^{(t)}$, then

$$\widehat{A\widehat{f}^{(t)}}\Xi^{(t)} - Af^{(t)}\Xi^{(t)} = \widehat{A}(\widehat{f}^{(t)} - f^{(t)})\Xi^{(t)} + \underbrace{(\widehat{A} - A)}_{\uparrow \atop E_{Q\setminus S\times Q\setminus S} + A_{[n]\times Q\setminus S} + A_{Q\setminus S\times[n]}} f^{(t)}\Xi^{(t)}$$

Induction hypothesis $\Longrightarrow (*) \approx 0$;
$|Q \setminus S| \leq \epsilon n \Longrightarrow (*) \approx 0$;
$\|E_{Q\setminus S\times Q\setminus S}\|_{\mathrm{op}} \leq \|\widehat{A} - A\|_{\mathrm{op}} = O(1) \Longrightarrow (*) \leq O(1)\cdot\|f^{(t)}_{Q\setminus S}\|_{\mathrm{F}} \approx 0$.

- Then $\widehat{f}^{(t+1)} = \varphi \circ (\frac{1}{\sqrt{n}}\widehat{A\widehat{f}^{(t)}}\Xi^{(t)}) \approx \varphi \circ (\frac{1}{\sqrt{n}}A\widehat{f}^{(t)}\Xi^{(t)}) \approx f^{(t+1)}$.

Conclusions and open problems

- We found a poly-time algorithm that matches two correlated Gaussian matrices with constant correlation even when two $\frac{n}{\mathrm{poly}(\log n)}$ size submatrices are adversarially corrupted.
- Our method: construct "signatures" by iteratively running an vector AMP on two matrices.
- A few open problems:
    - Other ways of corruption (e.g., corruption on arbitrary small edge set).
    - Robust algorithm for sparse graphs (edge density $q = n^{-\alpha+o(1)}$ when $\alpha > 0$)?

References

- Jian Ding and Zhangsong Li. A polynomial-time iterative algorithm for matching Gaussian matrices with non-vanishing correlation. *Foundations of Computational Mathematics*, to appear.
- Misha Ivkov and Tselil Schramm. Fast, robust approximate message passing. *STOC 2025*, to appear.
- Zhangsong Li. Robust random graph matching in dense graphs via vector approximate message passing. *COLT 2025*, to appear.

Thank you!