Random perturbation of low-rank matrices

Ke Wang

Department of Mathematics
Hong Kong University of Science and Technology

May 7, 2019

Joint work with Zhigang Bao (HKUST) & Xiucai Ding (Duke)

Matrix denoising model

Example: In image processing, we only observe $\widetilde{S}$. What can be said about the true data matrix S?

$$\underbrace{\tilde{S}} \quad = \quad \underbrace{S} \quad + \quad \underbrace{X}$$

Other applications: principal component analysis, matrix completion, community detection, etc.

Matrix denoising model

Signal: a real deterministic $m \times n$ matrix S with rank $r \ll \min\{m, n\}$.
Singular values: $d_1 \geq d_2 \geq \cdots \geq d_r > 0$.
Right and left singular vectors: v_i and u_i.

Matrix denoising model:

$$\widetilde{S} = \underbrace{S}_{\text{low-rank signal}} + \underbrace{X}_{\text{random noise}}$$

The (ordered) s.v.'s and s. vectors of $\widetilde{S}$ are $\widetilde{d}_i$, $\widetilde{v}_i$ and $\widetilde{u}_i$.

▷**Question**: How does the noise X affect the key parameters of S?

Key parameters: $\|S\|_2 = d_1$, top s.v.'s, top singular vectors, singular subspaces.

Matrix denoising model

Signal: a real deterministic $m \times n$ matrix S with rank $r \ll \min\{m, n\}$.
Singular values: $d_1 \geq d_2 \geq \cdots \geq d_r > 0$.
Right and left singular vectors: v_i and u_i.

Matrix denoising model:

$$\widetilde{S} = \underbrace{S}_{\text{low-rank signal}} + \underbrace{X}_{\text{random noise}}$$

The (ordered) s.v.'s and s. vectors of $\widetilde{S}$ are $\widetilde{d}_i$, $\widetilde{v}_i$ and $\widetilde{u}_i$.

▷**Question**: How does the noise X affect the key parameters of S?

Key parameters: $\|S\|_2 = d_1$, top s.v.'s, top singular vectors, singular subspaces.

Matrix denoising model

Signal: a real deterministic $m \times n$ matrix S with rank $r \ll \min\{m, n\}$.
Singular values: $d_1 \geq d_2 \geq \cdots \geq d_r > 0$.
Right and left singular vectors: v_i and u_i.

Matrix denoising model:

$$\widetilde{S} = \underbrace{S}_{\text{low-rank signal}} + \underbrace{X}_{\text{random noise}}$$

The (ordered) s.v.'s and s. vectors of $\widetilde{S}$ are $\widetilde{d}_i$, $\widetilde{v}_i$ and $\widetilde{u}_i$.

▷**Question**: How does the noise X affect the key parameters of S?

Key parameters: $\|S\|_2 = d_1$, top s.v.'s, top singular vectors, singular subspaces.

$$\widetilde{S} = S + X.$$

Classical matrix perturbation bounds:

- Weyl's inequality (1912): $|d_i - \widetilde{d}_i| \leq \|X\|_2$.
- Wedin sine theorem (1972): $\sin \angle(v_1, \widetilde{v}_1) \leq \frac{2\|X\|_2}{d_1 - d_2}$.

Quite wasteful when S is low-rank and X is random!

Key observation: When S is low rank, what really matters is r, the rank of S, rather than its size m or n.

Previous results: matrix perturbation theory

$$\widetilde{S} = S + X.$$

Classical matrix perturbation bounds:

- Weyl's inequality (1912): $|d_i - \widetilde{d}_i| \leq \|X\|_2$.
- Wedin sine theorem (1972): $\sin \angle(v_1, \widetilde{v}_1) \leq \frac{2\|X\|_2}{d_1 - d_2}$.

Quite wasteful when S is low-rank and X is random!

Key observation: When S is low rank, what really matters is r, the rank of S, rather than its size m or n.

Non-asymptotic perturbation bounds

▷ Vu '10, O'Rourke-Vu-W. '13 & '18: verify $u_i^T X v_i$ is small and $\|X\|_2 \ll d_r$ w.h.p. Improved classical bounds w.h.p.

E.g. X of size $n \times n$, entries i.i.d. random signs ± 1. Assume $d_1 \geq n$.

- Weyl's: $|d_i - \widetilde{d}_i| \leq C\sqrt{n}$ w.h.p. $\Rightarrow |d_i - \widetilde{d}_i| \leq C\sqrt{r}$ w.h.p.
- Wedin: $\sin \angle(v_1, \widetilde{v}_1) \leq \frac{C\sqrt{n}}{d_1 - d_2}$ w.h.p. $\Rightarrow \sin \angle(v_1, \widetilde{v}_1) \leq \frac{C\sqrt{r}}{d_1 - d_2}$ w.h.p.

Improve classical perturbation bounds for "low-rank + random noise":
▷ Wang '12: non-asymptotic dist. of s. vectors when entries of X i.i.d. $\mathcal{N}(0,1)$.
▷ Allez-Bouchaud '13: eigenvector dynamics (symmetric case).
▷ Cai-Zhang '16: dimensions m, n differ significantly.
▷ Fan-Wang-Zhong '16: improve Wedin's for S incoherent.
▷ Zhong '17: entries of X are i.i.d. sub-gaussian.

Non-asymptotic perturbation bounds

▷ Vu '10, O'Rourke-Vu-W. '13 & '18: verify $u_i^T X v_i$ is small and $\|X\|_2 \ll d_r$ w.h.p. Improved classical bounds w.h.p.

E.g. X of size $n \times n$, entries i.i.d. random signs ± 1. Assume $d_1 \geq n$.

- Weyl's: $|d_i - \widetilde{d}_i| \leq C\sqrt{n}$ w.h.p. $\Rightarrow |d_i - \widetilde{d}_i| \leq C\sqrt{r}$ w.h.p.
- Wedin: $\sin \angle(v_1, \widetilde{v}_1) \leq \frac{C\sqrt{n}}{d_1 - d_2}$ w.h.p. $\Rightarrow \sin \angle(v_1, \widetilde{v}_1) \leq \frac{C\sqrt{r}}{d_1 - d_2}$ w.h.p.

Improve classical perturbation bounds for "low-rank + random noise":
▷ Wang '12: non-asymptotic dist. of s. vectors when entries of X i.i.d. $\mathcal{N}(0,1)$.
▷ Allez-Bouchaud '13: eigenvector dynamics (symmetric case).
▷ Cai-Zhang '16: dimensions m, n differ significantly.
▷ Fan-Wang-Zhong '16: improve Wedin's for S incoherent.
▷ Zhong '17: entries of X are i.i.d. sub-gaussian.

Non-asymptotic perturbation bounds

▷ Vu '10, O'Rourke-Vu-W. '13 & '18: verify $u_i^T X v_i$ is small and $\|X\|_2 \ll d_r$ w.h.p. Improved classical bounds w.h.p.

E.g. X of size $n \times n$, entries i.i.d. random signs ± 1. Assume $d_1 \geq n$.

- Weyl's: $|d_i - \widetilde{d}_i| \leq C\sqrt{n}$ w.h.p. $\Rightarrow$ $|d_i - \widetilde{d}_i| \leq C\sqrt{r}$ w.h.p.
- Wedin: $\sin\angle(v_1, \widetilde{v}_1) \leq \frac{C\sqrt{n}}{d_1 - d_2}$ w.h.p. $\Rightarrow$ $\sin\angle(v_1, \widetilde{v}_1) \leq \frac{C\sqrt{r}}{d_1 - d_2}$ w.h.p.

Improve classical perturbation bounds for "low-rank + random noise":
▷ Wang '12: non-asymptotic dist. of s. vectors when entries of X i.i.d. $\mathcal{N}(0,1)$.
▷ Allez-Bouchaud '13: eigenvector dynamics (symmetric case).
▷ Cai-Zhang '16: dimensions m, n differ significantly.
▷ Fan-Wang-Zhong '16: improve Wedin's for S incoherent.
▷ Zhong '17: entries of X are i.i.d. sub-gaussian.

Deformation models:

▷ Low-rank deformed Wigner: P_n fixed-rank Hermitian

$$W_n + P_n \quad \text{(additive)}.$$

▷ Spiked covariance matrix: Non-negative $T = I + R$; R fixed-rank

$$T^{1/2} X X^* T^{1/2} \quad \text{(multiplicative)}.$$

▷ Matrix denoising model: $X + S$ with S fixed-rank

$$(X+S)^*(X+S) \quad \text{(additive \& multiplicative)}.$$

Deformation and random matrix are of the same order.

Previous results: low-rank deformation of random matrices

Deformation models:

▷ Low-rank deformed Wigner: P_n fixed-rank Hermitian

$$W_n + P_n \quad \text{(additive)}.$$

▷ Spiked covariance matrix: Non-negative $T = I + R$; R fixed-rank

$$T^{1/2} X X^* T^{1/2} \quad \text{(multiplicative)}.$$

▷ Matrix denoising model: $X + S$ with S fixed-rank

$$(X+S)^*(X+S) \quad \text{(additive \& multiplicative)}.$$

Deformation and random matrix are of the same order.

Previous results: low-rank deformation of random matrices

Deformation models:

▷ Low-rank deformed Wigner: P_n fixed-rank Hermitian

$$W_n + P_n \quad \text{(additive)}.$$

▷ Spiked covariance matrix: Non-negative $T = I + R$; R fixed-rank

$$T^{1/2} X X^* T^{1/2} \quad \text{(multiplicative)}.$$

▷ Matrix denoising model: $X + S$ with S fixed-rank

$$(X + S)^*(X + S) \quad \text{(additive \& multiplicative)}.$$

Deformation and random matrix are of the same order.

Previous results: low-rank deformation of random matrices

Deformation models:

▷ Low-rank deformed Wigner: P_n fixed-rank Hermitian

$$W_n + P_n \quad \text{(additive)}.$$

▷ Spiked covariance matrix: Non-negative $T = I + R$; R fixed-rank

$$T^{1/2} X X^* T^{1/2} \quad \text{(multiplicative)}.$$

▷ Matrix denoising model: $X + S$ with S fixed-rank

$$(X + S)^*(X + S) \quad \text{(additive \& multiplicative)}.$$

Deformation and random matrix are of the same order.

Previous results: low-rank deformation of random matrices

BBP phase transition by Baik-Ben Arous-Péché '05 for extreme eigenvalues of spiked complex Gaussian covariance matrix. Phase transition for e.vector: Paul '07, Benaych-Georges-Nadakuditi '11 & '12.

General picture:

- (Supercritical): deformation strength bigger than a critical value c.

- (Subcritical): deformation strength less than a critical value c.

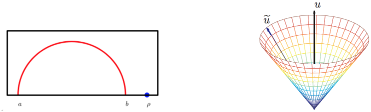
Picture source: *"The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices"*, Benaych-Georges-Nadakuditi, Advances in Mathematics, 227(1):494-521,2011.

Previous results: low-rank deformation of random matrices

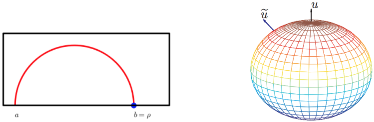
BBP phase transition by Baik-Ben Arous-Péché '05 for extreme eigenvalues of spiked complex Gaussian covariance matrix. Phase transition for e.vector: Paul '07, Benaych-Georges-Nadakuditi '11 & '12.

General picture: $\widetilde{W}_n = W_n + \lambda u u^T$.

- (Supercritical): deformation strength bigger than a critical value c.



- (Subcritical): deformation strength less than a critical value c.



Picture source: *"The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices"*, Benaych-Georges-Nadakuditi, Advances in Mathematics, 227(1):494-521,2011.

Asymptotic behavior

▷ Extreme eigenvalues

- Convergent limit: Bai-Yao '12, Baik-Silverstein '06, Benaych-Georges- Nadakuditi '11 & '12, Capitaine-Donati-Martin '16, Ding '17, Knowles-Yin '13, Paul '07, etc.
- Fluctuation: Bai-Yao '08, Bao-Pan-Zhou '15, Benaych-Georges-Guionnet-Maida '11, Bloemendal-Knowles-Yau-Yin '16, Bloemendal-Virág '13 & '16, Capitaine-Donati-Martin-Féral '09, Knowles-Yin '13, Renfrew-Soshnikov '12, etc.

▷ Extreme eigenvectors

- Convergent limit: Benaych-Georges-Nadakuditi '11 & '12, Capitaine '17, Ding '17, Paul '07.
- Fluctuation: Paul '07, Bloemendal-Knowles-Yau-Yin '16, Capitaine-Donati-Martin '18. Not fully studied for matrix denoising model!

New results: matrix denoising model

Consider the $m \times n$ matrix

$$\widetilde{S} = S + X = \sum_{i=1}^{r} d_i u_i v_i^T + X = \sum_{i=1}^{\min\{m,n\}} \widetilde{d}_i \widetilde{u}_i \widetilde{v}_i^T .$$

Assumptions:

- $m/n \to y \in (0, +\infty)$ as $n \to \infty$.
- Entries of $\sqrt{n}X$ i.i.d. with mean 0, var. 1, bounded high moments.
- (Supercritical condition) There exists $\delta > 0$ such that

$$d_1 > d_2 > \cdots > d_r \geq y^{1/4} + \delta \text{ and } \min_{1 \leq j \neq i \leq r} |d_i - d_j| \geq \delta.$$

New results: Fluctuations of $|\langle v_i, \widetilde{v}_i \rangle|^2$ and singular subspace statistics.

New results: matrix denoising model

Consider the $m \times n$ matrix

$$\widetilde{S} = S + X = \sum_{i=1}^{r} d_i u_i v_i^T + X = \sum_{i=1}^{\min\{m,n\}} \widetilde{d}_i \widetilde{u}_i \widetilde{v}_i^T.$$

Assumptions:

- $m/n \to y \in (0, +\infty)$ as $n \to \infty$.
- Entries of $\sqrt{n}X$ i.i.d. with mean 0, var. 1, bounded high moments.
- (Supercritical condition) There exists $\delta > 0$ such that

$$d_1 > d_2 > \cdots > d_r \geq y^{1/4} + \delta \text{ and } \min_{1 \leq j \neq i \leq r} |d_i - d_j| \geq \delta.$$

New results: Fluctuations of $|\langle v_i, \widetilde{v}_i \rangle|^2$ and singular subspace statistics.

New results: right singular vectors

$$a(d) := \frac{d^4 - y}{d^2(d^2+1)}, \quad \theta(d) := \frac{d^4 + 2yd^2 + y}{d^3(d^2+1)^2}.$$

Theorem (Bao, Ding and W., 2018)

For each $1 \le i \le r$,

$$\sqrt{n}\left(|\langle v_i, \widetilde{v}_i\rangle|^2 - a(d_i)\right) \simeq \Delta_i + \mathcal{Z}_i.$$

- Δ_i and $\mathcal{Z}_i$ are independent.
- $\Delta_i := -2\sqrt{n}\,\theta(d_i) u_i^T X v_i$ and $\mathcal{Z}_i \sim \mathcal{N}(\mathcal{M}_i, \mathcal{V}_i)$ ($\mathcal{M}_i, \mathcal{V}_i$ depend on s.vectors u_i, v_i and cumulants κ_3, κ_4 of $\sqrt{n}X_{11}$).

For a vector w, denote $\mathbf{s}_k(w) := \sum_i w(i)^k$.

$$\psi(d) := \frac{d^6 - 3yd^2 - 2y}{d^3(d^2+1)^2},$$

$$\mathcal{V}^E(d) := \frac{2}{d^4 - y}\Big[2y(y+1)\theta(d)^2 - \frac{y(y-1)(5y+1)}{d(d^2+1)^2}\theta(d) + \frac{(d^4+y)(d^2+y)^2}{d^3(d^2+1)^2}\psi(d) + \frac{2y^2(y-1)^2}{d^2(d^2+1)^4}\Big].$$

$\mathcal{Z}_i \sim \mathcal{N}(\mathcal{M}_i, \mathcal{V}_i)$, where

$$\mathcal{M}_i := -\frac{2\psi(d_i)}{d_i^2}\Big(\frac{\kappa_3}{n}\mathbf{s}_1(u_i)\mathbf{s}_1(v_i)\Big),$$

$$\mathcal{V}_i := \mathcal{V}^E(d_i) - \frac{4}{d_i}\theta(d_i)\psi(d_i)\frac{\kappa_3}{\sqrt{n}}\mathbf{s}_3(u_i)\mathbf{s}_1(v_i) + \frac{4}{d_i}\theta(d_i)^2\frac{\kappa_3}{\sqrt{n}}\mathbf{s}_1(u_i)\mathbf{s}_3(v_i) + \frac{\psi(d_i)^2}{d_i^2}\kappa_4\mathbf{s}_4(u_i) + \frac{y\theta(d_i)^2}{d_i^2}\kappa_4\mathbf{s}_4(v_i).$$

Special examples

Non-universality: fluctuation depends on "structure of signal" and "distribution of noise".

Assume $r = 1$. $\widetilde{S} = duv^T + X$.

- The entries of $\sqrt{n}X$ are i.i.d. $\mathcal{N}(0, 1)$.

$$\Delta + \mathcal{Z} \simeq \mathcal{N}\left(0, V(d)\right).$$

- $\|u\|_\infty = o(1)$ and $\|v\|_\infty = o(1)$.

$$\Delta + \mathcal{Z} \simeq \mathcal{N}\left(\mathcal{M}(d, u, v), V(d)\right).$$

$$\mathcal{M}(d, u, v) := -\tfrac{2\psi(d)}{d^2}\left(\tfrac{\kappa_3}{n}[\textstyle\sum_i u(i)] \cdot [\sum_i v(i)]\right).$$

Special examples

Non-universality: fluctuation depends on "structure of signal" and "distribution of noise".

Assume $r = 1$. $\widetilde{S} = duv^T + X$.

- The entries of $\sqrt{n}X$ are i.i.d. $\mathcal{N}(0,1)$.

$$\Delta + \mathcal{Z} \simeq \mathcal{N}\left(0, V(d)\right).$$

- $\|u\|_\infty = o(1)$ and $\|v\|_\infty = o(1)$.

$$\Delta + \mathcal{Z} \simeq \mathcal{N}\left(\mathcal{M}(d,u,v), V(d)\right).$$

$$\mathcal{M}(d,u,v) := -\tfrac{2\psi(d)}{d^2}\left(\tfrac{\kappa_3}{n}[\textstyle\sum_i u(i)] \cdot [\textstyle\sum_i v(i)]\right).$$

Non-universality: fluctuation depends on "structure of signal" and "distribution of noise".

Assume $r = 1$. $\widetilde{S} = duv^T + X$.

- The entries of $\sqrt{n}X$ are i.i.d. $\mathcal{N}(0, 1)$.

$$\Delta + \mathcal{Z} \simeq \mathcal{N}\left(0, V(d)\right).$$

- $\|u\|_\infty = o(1)$ and $\|v\|_\infty = o(1)$.

$$\Delta + \mathcal{Z} \simeq \mathcal{N}\left(\mathcal{M}(d, u, v), V(d)\right).$$

$$\mathcal{M}(d, u, v) := -\tfrac{2\psi(d)}{d^2}\left(\tfrac{\kappa_3}{n}[\textstyle\sum_i u(i)] \cdot [\textstyle\sum_i v(i)]\right).$$

Special examples

- Only $\|u\|_\infty = o(1)$.

$$\Delta + \mathcal{Z} \simeq \mathcal{N}\left(\mathcal{M}(d,u,v), V(d) + V_1(d,u,v)\right)$$

$$V_1(d,u,v) := \tfrac{4\theta(d)^2}{d}\left(\tfrac{\kappa_3}{\sqrt{n}}[\textstyle\sum_i u(i)] \cdot [\sum_i v(i)^3]\right) + y\tfrac{\theta(d)^2}{d^2}\kappa_4[\textstyle\sum_i v(i)^4].$$

- $u = e_1$ and $v = e_1$.

$$\Delta + \mathcal{Z} \simeq -2\theta(d)\sqrt{n}X_{11} + \mathcal{N}\left(0, \mathcal{V}^E(d) + \kappa_4\frac{\psi(d)^2 + y\theta(d)^2}{d^2}\right).$$

Special examples

- Only $\|u\|_\infty = o(1)$.

$$\Delta + \mathcal{Z} \simeq \mathcal{N}\left(\mathcal{M}(d,u,v), V(d) + V_1(d,u,v)\right)$$

$$V_1(d,u,v) := \frac{4\theta(d)^2}{d}\left(\frac{\kappa_3}{\sqrt{n}}[\sum_i u(i)] \cdot [\sum_i v(i)^3]\right) + y\frac{\theta(d)^2}{d^2}\kappa_4[\sum_i v(i)^4].$$

- $u = e_1$ and $v = e_1$.

$$\Delta + \mathcal{Z} \simeq -2\theta(d)\sqrt{n}X_{11} + \mathcal{N}\left(0, \mathcal{V}^E(d) + \kappa_4\frac{\psi(d)^2 + y\theta(d)^2}{d^2}\right).$$

Related results

In the **supercritical** regime for deformed Wigner $(W_n + P_n)$:

- Knowles-Yin '13: limiting dist. of the outlier eigenvalues of deformed Wigner in full generality. The fluctuation depends on the "structure of deformation" and "distribution of random noise".
- Capitaine-Donati-Martin '18: fluctuation of outlier eigenvectors where P_n is diagonal and entries of Wigner have symmetric dist. and satisfy Poincaré inequality.

Related results

In the **supercritical** regime for deformed Wigner ($W_n + P_n$):

- Knowles-Yin '13: limiting dist. of the outlier eigenvalues of deformed Wigner in full generality. The fluctuation depends on the "structure of deformation" and "distribution of random noise".
- Capitaine-Donati-Martin '18: fluctuation of outlier eigenvectors where P_n is diagonal and entries of Wigner have symmetric dist. and satisfy Poincaré inequality.

New results: right singular subspace

$$V = (v_1, \cdots, v_r) \quad \text{and} \quad \widetilde{V}_r = (\widetilde{v}_1, \cdots, \widetilde{v}_r).$$

Consider

$$R := \sum_{i,j=1}^{r} |\langle \widetilde{v}_i, v_j \rangle|^2.$$

Theorem (Bao, Ding and W., 2018)

$$\sqrt{n}\Big(R - \sum_{i=1}^{r} a(d_i)\Big) \simeq \Delta + \mathcal{Z}.$$

Δ and $\mathcal{Z}$ are independent; analogous definitions to the previous result.

Sketch of proof

Knowles and Yin's approach (extreme e.v. of deformed Wigner), based on a two-step comparison. $\widetilde{W} = W + P$.

- Derive results when for W is Gaussian.
- Compare the Gaussian case with a partially Gaussian matrix $\widehat{W}$ (depending on the structure of P; large entries cause trouble).
- Compare $\widehat{W}$ with the general Wigner W.

Our approach is different and more direct.

Sketch of proof

Knowles and Yin's approach (extreme e.v. of deformed Wigner), based on a two-step comparison. $\widetilde{W} = W + P$.

- Derive results when for W is Gaussian.
- Compare the Gaussian case with a partially Gaussian matrix $\widehat{W}$ (depending on the structure of P; large entries cause trouble).
- Compare $\widehat{W}$ with the general Wigner W.

Our approach is different and more direct.

Sketch of proof for individual singular vector

$$\widetilde{S} = duv^T + X.$$

The empirical spectral distributions (ESD) of the matrices XX^T and X^TX

$$F_1(x) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{\{\lambda_i(XX^T)\le x\}}, \qquad F_2(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{\lambda_i(X^TX)\le x\}}.$$

The Marchenko-Pastur (MP) law:

$$F_1(x) \to F_{MP,1}(x), \qquad F_2(x) \to F_{MP,2}(x).$$

The Stieltjes's transforms:

$$m_1(z) := \int \frac{1}{x-z}\,\mathrm{d}F_{MP,1}(x), \qquad m_2(z) := \int \frac{1}{x-z}\,\mathrm{d}F_{MP,2}(x).$$

Sketch of proof for individual singular vector

$$\widetilde{S} = duv^T + X.$$

The empirical spectral distributions (ESD) of the matrices XX^T and X^TX

$$F_1(x) = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{\{\lambda_i(XX^T)\le x\}}, \qquad F_2(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{\lambda_i(X^TX)\le x\}}.$$

The Marchenko-Pastur (MP) law:

$$F_1(x) \to F_{MP,1}(x), \qquad F_2(x) \to F_{MP,2}(x).$$

The Stieltjes's transforms:

$$m_1(z) := \int \frac{1}{x-z}\,\mathrm{d}F_{MP,1}(x), \qquad m_2(z) := \int \frac{1}{x-z}\,\mathrm{d}F_{MP,2}(x).$$

Sketch of proof: linearization

$$\widetilde{S} = duv^T + X.$$

Linearization technique: Work with

$$\mathcal{Y} := \mathcal{U}\mathcal{D}\mathcal{U}^T + H$$

where

$$\mathcal{U} := \begin{pmatrix} u & \\ & v \end{pmatrix}, \quad \mathcal{D}(z) := \sqrt{z}\begin{pmatrix} & d \\ d & \end{pmatrix}, \quad H(z) := \sqrt{z}\begin{pmatrix} & X \\ X^T & \end{pmatrix}.$$

The Green function of H is

$$G(z) = (H - z)^{-1},$$

whose approximation is

$$\Pi(z) := m_1(z)I_m \oplus m_2(z)I_n.$$

Sketch of proof: linearization

$$\widetilde{S} = duv^T + X.$$

Linearization technique: Work with

$$\mathcal{Y} := \mathcal{U}\mathcal{D}\mathcal{U}^T + H$$

where

$$\mathcal{U} := \begin{pmatrix} u & \\ & v \end{pmatrix}, \quad \mathcal{D}(z) := \sqrt{z} \begin{pmatrix} & d \\ d & \end{pmatrix}, \quad H(z) := \sqrt{z} \begin{pmatrix} & X \\ X^T & \end{pmatrix}.$$

The Green function of H is

$$G(z) = (H - z)^{-1},$$

whose approximation is

$$\Pi(z) := m_1(z) I_m \oplus m_2(z) I_n.$$

Sketch of proof: Green function representation

Green function representation of $|\langle v, \widetilde{v}\rangle|^2$: by residue theorem,

$$|\langle v, \widetilde{v}\rangle|^2 = \frac{1}{2d^2\pi\sqrt{-1}} \oint \left(\left(\mathcal{D}^{-1} + \mathcal{U}^T G(z)\mathcal{U}\right)^{-1}\right)_{11} \frac{\mathrm{d}z}{z}.$$

By resolvent expansion,

$$\sqrt{n}(|\langle v, \widetilde{v}\rangle|^2 - a(d)) \approx \sqrt{n}\left(\mathrm{Tr}(G - \Pi)A + \mathrm{Tr}(G' - \Pi')B\right),$$

where A, B are explicit fixed-rank and bounded matrices.

Derive the law of

$$\mathcal{Q} = \sqrt{n}\left(\mathrm{Tr}(G - \Pi)A + \mathrm{Tr}(G' - \Pi')B\right).$$

Construct

$$\Delta = \underbrace{\Psi_1}_{\substack{\text{random part: linear combination of } X_{ij}\text{'s} \\ \text{w.r.t. large components of } u \text{ and } v}} + \underbrace{\Psi_2}_{\text{deterministic part: centralization}}.$$

Goals:

- The Gaussianity of $\mathcal{Q} - \Delta$.
- Independence between $\mathcal{Q} - \Delta$ and Δ.

Sketch of proof: further reduction

Derive the law of

$$\mathcal{Q} = \sqrt{n}\left(\mathrm{Tr}(G-\Pi)A + \mathrm{Tr}(G'-\Pi')B\right).$$

Construct

$$\Delta = \underbrace{\Psi_1}_{\substack{\text{random part: linear combination of } X_{ij}\text{'s} \\ \text{w.r.t. large components of } u \text{ and } v}} + \underbrace{\Psi_2}_{\text{deterministic part: centralization}}.$$

Goals:

- The Gaussianity of $\mathcal{Q} - \Delta$.
- Independence between $\mathcal{Q} - \Delta$ and Δ.

Sketch of proof: further reduction

Derive the law of

$$\mathcal{Q} = \sqrt{n}\left(\mathrm{Tr}(G-\Pi)A + \mathrm{Tr}(G'-\Pi')B\right).$$

Construct

$$\Delta = \underbrace{\Psi_1}_{\substack{\text{random part: linear combination of } X_{ij}\text{'s} \\ \text{w.r.t. large components of } u \text{ and } v}} + \underbrace{\Psi_2}_{\text{deterministic part: centralization}}.$$

Goals:

- The Gaussianity of $\mathcal{Q} - \Delta$.
- Independence between $\mathcal{Q} - \Delta$ and Δ.

Sketch of proof: recursive estimates

Our strategy is to establish the recursive estimates

$$\mathbb{E}(\mathcal{Q}-\Delta)^k \mathrm{e}^{\mathrm{i}t\Delta} = (k-1)\mathcal{V}\mathbb{E}(\mathcal{Q}-\Delta)^{k-2}\mathrm{e}^{\mathrm{i}t\Delta} + o(1).$$

The goals are achieved simultaneously.

- Gaussianity follows by taking $t = 0$.
- Independence follows from the consequence of recursive estimates:

$$\mathbb{E}\mathrm{e}^{\mathrm{i}s(\mathcal{Q}-\Delta)+\mathrm{i}t\Delta} = \mathbb{E}\mathrm{e}^{\mathrm{i}s(\mathcal{Q}-\Delta)}\mathbb{E}\mathrm{e}^{\mathrm{i}t\Delta} + o(1).$$

Sketch of proof: recursive estimates

Our strategy is to establish the recursive estimates

$$\mathbb{E}(\mathcal{Q}-\Delta)^k \mathrm{e}^{\mathrm{i}t\Delta} = (k-1)\mathcal{V}\mathbb{E}(\mathcal{Q}-\Delta)^{k-2}\mathrm{e}^{\mathrm{i}t\Delta} + o(1).$$

The goals are achieved simultaneously.

- Gaussianity follows by taking $t=0$.
- Independence follows from the consequence of recursive estimates:

$$\mathbb{E}\mathrm{e}^{\mathrm{i}s(\mathcal{Q}-\Delta)+\mathrm{i}t\Delta} = \mathbb{E}\mathrm{e}^{\mathrm{i}s(\mathcal{Q}-\Delta)}\mathbb{E}\mathrm{e}^{\mathrm{i}t\Delta} + o(1).$$

Two key ingredients in the proof of recursive estimates.

- **Cumulant expansion formula:** For $f \in C^{\ell+1}(\mathbb{R})$ and ξ a centered random variable with finite $l+2$ moments,

$$\mathbb{E}(\xi f(\xi)) = \sum_{k=1}^{\ell} \frac{\kappa_{k+1}(\xi)}{k!} \mathbb{E}(f^{(k)}(\xi)) + \mathbb{E}(\epsilon_\ell(\xi f(\xi))).$$

  Applications in RMT: Khorunzhy-Khoruzhenko-Pastur '96, Lytova-Pastur 09', Lee-Schnelli '16, He-Knowles '16.

- **Isotropic local laws:** large deviation bounds of

$$\langle u, (G^{(l)} - \Pi^{(l)})v \rangle \quad \text{for } l \in \mathbb{N}.$$

  Established in Bloemendal-Erdős-Knowles-Yau-Yin '14, Knowles-Yin '17.

Sketch of proof: key technical inputs

Two key ingredients in the proof of recursive estimates.

- **Cumulant expansion formula:** For $f \in C^{\ell+1}(\mathbb{R})$ and ξ a centered random variable with finite $l+2$ moments,

$$\mathbb{E}(\xi f(\xi)) = \sum_{k=1}^{\ell} \frac{\kappa_{k+1}(\xi)}{k!} \mathbb{E}(f^{(k)}(\xi)) + \mathbb{E}(\epsilon_\ell(\xi f(\xi))).$$

  Applications in RMT: Khorunzhy-Khoruzhenko-Pastur '96, Lytova-Pastur 09', Lee-Schnelli '16, He-Knowles '16.

- **Isotropic local laws:** large deviation bounds of

$$\langle u, (G^{(l)} - \Pi^{(l)})v \rangle \quad \text{for } l \in \mathbb{N}.$$

  Established in Bloemendal-Erdős-Knowles-Yau-Yin '14, Knowles-Yin '17.

THANK YOU!