Large deviations of the spectrum of supercritical sparse Erdős-Rényi graphs

Fanny Augeri

Université Paris Cité

Bernoulli matrices

Let $(A_{i,j})_{1\leq i\leq j\leq n}$ be i.i.d. $\mathrm{Ber}(1/2)$.

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Bernoulli matrices

Let $(A_{i,j})_{1\leq i\leq j\leq n}$ be i.i.d. $\mathrm{Ber}(1/2)$.

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

If $\mu_{A/\sqrt{n}}$ is the empirical spectral measure, then

(Wigner's law) $\qquad \mu_{A/\sqrt{n}} \underset{n\to+\infty}{\Longrightarrow} \mu_{\mathsf{sc}} \qquad$ in probability,

where $\mu_{\mathsf{sc}} = (2/\pi)\sqrt{(1-x^2)_+}dx$.

Bernoulli matrices

Let $(A_{i,j})_{1\leq i\leq j\leq n}$ be i.i.d. $\mathrm{Ber}(1/2)$.

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

If $\mu_{A/\sqrt{n}}$ is the empirical spectral measure, then

(Wigner's law) $$\mu_{A/\sqrt{n}} \underset{n\to+\infty}{\Longrightarrow} \mu_{\mathsf{sc}}$$ in probability,

where $\mu_{\mathsf{sc}} = (2/\pi)\sqrt{(1-x^2)_+}dx$.

What is the large deviation behaviour of $\mu_{A/\sqrt{n}}$?

Bernoulli matrices

- Large deviations of $\mu_{A/\sqrt{n}}$ turn out to be very difficult.

$$\text{(LDP)} \qquad \log \mathbb{P}(\mu_{A/\sqrt{n}} \simeq \nu) \overset{?}{\asymp} e^{-v_n I(\nu)}, \ \nu \neq \mu_{\mathsf{sc}},$$

where v_n is the speed and I the rate function.

Bernoulli matrices

- Large deviations of $\mu_{A/\sqrt{n}}$ turn out to be very difficult.

$$\text{(LDP)} \qquad \log \mathbb{P}(\mu_{A/\sqrt{n}} \simeq \nu) \overset{?}{\asymp} e^{-v_n I(\nu)}, \ \nu \neq \mu_{\mathsf{sc}},$$

  where v_n is the speed and I the rate function.
- What do we know ? Not much...

Bernoulli matrices

- Large deviations of $\mu_{A/\sqrt{n}}$ turn out to be very difficult.

$$\text{(LDP)} \qquad \log \mathbb{P}(\mu_{A/\sqrt{n}} \simeq \nu) \overset{?}{\asymp} e^{-v_n I(\nu)}, \ \nu \neq \mu_{\mathsf{sc}},$$

  where v_n is the speed and I the rate function.
- What do we know ? Not much...
  - (Guionnet-Zeitouni): Speed of large deviation is n^2

$$\log \mathbb{P}(d(\mu_{A/\sqrt{n}}, \mu_{\mathsf{sc}}) > \varepsilon) \leq -C_\varepsilon n^2.$$

Bernoulli matrices

- Large deviations of $\mu_{A/\sqrt{n}}$ turn out to be very difficult.

$$\text{(LDP)} \qquad \log \mathbb{P}(\mu_{A/\sqrt{n}} \simeq \nu) \overset{?}{\asymp} e^{-v_n I(\nu)}, \ \nu \neq \mu_{\mathsf{sc}},$$

where v_n is the speed and I the rate function.

- What do we know ? Not much...
    - (Guionnet-Zeitouni): Speed of large deviation is n^2

$$\log \mathbb{P}(d(\mu_{A/\sqrt{n}}, \mu_{\mathsf{sc}}) > \varepsilon) \leq -C_\varepsilon n^2.$$

    - If the rate function I exists, then

$$I(\mu) = +\infty, \quad \text{if } \mu(x^2) > 1/4,$$

since $n^{-2}\mathrm{tr}(A - \mathbb{E}A)^2 = 1/4$ and $\mathbb{E}A$ rank 1.

Bernoulli matrices

- Large deviations of $\mu_{A/\sqrt{n}}$ turn out to be very difficult.

$$\text{(LDP)} \qquad \log \mathbb{P}(\mu_{A/\sqrt{n}} \simeq \nu) \overset{?}{\asymp} e^{-v_n I(\nu)}, \ \nu \neq \mu_{\mathsf{sc}},$$

  where v_n is the speed and I the rate function.
- What do we know ? Not much...
  - (Guionnet-Zeitouni): Speed of large deviation is n^2

$$\log \mathbb{P}(d(\mu_{A/\sqrt{n}}, \mu_{\mathsf{sc}}) > \varepsilon) \leq -C_\varepsilon n^2.$$

  - If the rate function I exists, then

$$I(\mu) = +\infty, \quad \text{if } \mu(x^2) > 1/4,$$

    since $n^{-2}\mathrm{tr}(A - \mathbb{E}A)^2 = 1/4$ and $\mathbb{E}A$ rank 1.
  - I cannot be universal: $I \neq I_{\mathrm{GOE}}$.

$$\delta_0 \in \mathcal{D}_I \text{ but } \delta_0 \notin \mathcal{D}_{I_{\mathsf{GOE}}}.$$

Some universality in large deviation behaviour

Let $\underline{A} = A - \mathbb{E}A$. By Cauchy interlacing inequalities

$$d_{\mathrm{KS}}(\mu_{\underline{A}/\sqrt{n}}, \mu_{A/\sqrt{n}}) \leq \frac{1}{n}\mathrm{rk}(A - \underline{A}) = \frac{1}{n}.$$

- Large deviations of $\mu_{\underline{A}/\sqrt{n}}$ and $\mu_{A/\sqrt{n}}$ are the same.

Some universality in large deviation behaviour

Let $\underline{A} = A - \mathbb{E}A$. By Cauchy interlacing inequalities

$$d_{\mathrm{KS}}(\mu_{\underline{A}/\sqrt{n}}, \mu_{A/\sqrt{n}}) \leq \frac{1}{n}\mathrm{rk}(A - \underline{A}) = \frac{1}{n}.$$

- Large deviations of $\mu_{\underline{A}/\sqrt{n}}$ and $\mu_{A/\sqrt{n}}$ are the same.

What about other spectral observables?

Some universality in large deviation behaviour

Let $\underline{A} = A - \mathbb{E}A$. By Cauchy interlacing inequalities

$$d_{\mathrm{KS}}(\mu_{\underline{A}/\sqrt{n}}, \mu_{A/\sqrt{n}}) \leq \frac{1}{n}\mathrm{rk}(A - \underline{A}) = \frac{1}{n}.$$

➤ Large deviations of $\mu_{\underline{A}/\sqrt{n}}$ and $\mu_{A/\sqrt{n}}$ are the same.

What about other spectral observables?

- (Guionnet-Husson) $\lambda_{\underline{A}/\sqrt{n}}$ satisfies a LDP

  with speed n and rate function $J = J_{\mathrm{GOE}}$.

Some universality in large deviation behaviour

Let $\underline{A} = A - \mathbb{E}A$. By Cauchy interlacing inequalities

$$d_{\mathrm{KS}}(\mu_{\underline{A}/\sqrt{n}}, \mu_{A/\sqrt{n}}) \leq \frac{1}{n}\mathrm{rk}(A - \underline{A}) = \frac{1}{n}.$$

➤ Large deviations of $\mu_{\underline{A}/\sqrt{n}}$ and $\mu_{A/\sqrt{n}}$ are the same.

What about other spectral observables?

- (Guionnet-Husson) $\lambda_{\underline{A}/\sqrt{n}}$ satisfies a LDP

  with speed n and rate function $J = J_{\mathrm{GOE}}$.

- (A.) $\mu_{\underline{A}/\sqrt{n}}(x^k)$ satisfies a LDP

  with speed $n^{1+\frac{2}{k}}$ and rate function $L^k = L^k_{\mathrm{GOE}}$.

Some universality in large deviation behaviour

Let $\underline{A} = A - \mathbb{E}A$. By Cauchy interlacing inequalities

$$d_{\mathrm{KS}}(\mu_{\underline{A}/\sqrt{n}}, \mu_{A/\sqrt{n}}) \leq \frac{1}{n}\mathrm{rk}(A - \underline{A}) = \frac{1}{n}.$$

➤ Large deviations of $\mu_{\underline{A}/\sqrt{n}}$ and $\mu_{A/\sqrt{n}}$ are the same.

What about other spectral observables?

- (Guionnet-Husson) $\lambda_{\underline{A}/\sqrt{n}}$ satisfies a LDP

  with speed n and rate function $J = J_{\mathrm{GOE}}$.

- (A.) $\mu_{\underline{A}/\sqrt{n}}(x^k)$ satisfies a LDP

  with speed $n^{1+\frac{2}{k}}$ and rate function $L^k = L^k_{\mathrm{GOE}}$.

➤ Universality is a manifestation of moderate deviation behaviour.

Sparsification as a relaxation of the problem

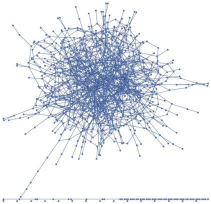
Take $p \ll 1$ and consider A the adjacency matrix of the Erdős-Rényi graph $G(n, p)$.

Sparsification as a relaxation of the problem

Take $p \ll 1$ and consider A the adjacency matrix of the Erdős-Rényi graph $G(n, p)$.

0 — $\log n$ — n

$np \lesssim \log n$	$\log n \ll np$
Inhomogeneous graph:	Homogeneous graph:
• isolated vertices • hubs • leaves	• vertices have all almost the same degrees • connected w.h.p



The subcritical regime $np = \lambda$

- Asymptotically, $G(n, \frac{\lambda}{n})$ looks like a Galton-Watson tree with degree distribution $\mathrm{Poi}(\lambda)$.

The subcritical regime $np = \lambda$

➤ Asymptotically, $G(n, \frac{\lambda}{n})$ looks like a Galton-Watson tree with degree distribution $\mathrm{Poi}(\lambda)$.

(Abert-Thom-Virág) $\qquad \mu_A \underset{n\to+\infty}{\Longrightarrow} \mu_\lambda, \qquad$ in probability

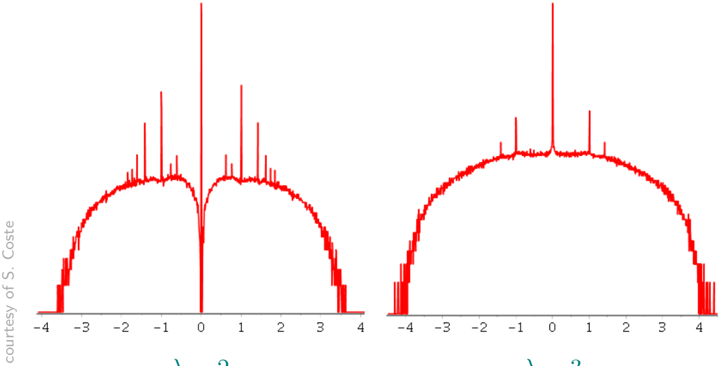
where μ_λ is the expected spectral measure of $\mathrm{GWT}(\mathrm{Poi}(\lambda))$.

The subcritical regime $np = \lambda$

➤ Asymptotically, $G(n, \frac{\lambda}{n})$ looks like a Galton-Watson tree with degree distribution $\mathrm{Poi}(\lambda)$.

(Abert-Thom-Virág) $\mu_A \underset{n\to+\infty}{\Longrightarrow} \mu_\lambda,$ in probability

where μ_λ is the expected spectral measure of $\mathrm{GWT}(\mathrm{Poi}(\lambda))$.



$\lambda = 2$ $\lambda = 3$

The subcritical regime $np = \lambda$

➤ Asymptotically, $G(n, \frac{\lambda}{n})$ looks like a Galton-Watson tree with degree distribution $\mathrm{Poi}(\lambda)$.

(Abért-Thom-Virág) $\mu_A \underset{n \to +\infty}{\Longrightarrow} \mu_\lambda,$ in probability

where μ_λ is the expected spectral measure of $\mathrm{GWT}(\mathrm{Poi}(\lambda))$.

- (Bordenave-Lelarge-Salez): μ_λ is symmetric

The subcritical regime $np = \lambda$

➤ Asymptotically, $G(n, \frac{\lambda}{n})$ looks like a Galton-Watson tree with degree distribution $\mathrm{Poi}(\lambda)$.

(Abért-Thom-Virág) $$\mu_A \underset{n \to +\infty}{\Longrightarrow} \mu_\lambda,$$ in probability

where μ_λ is the expected spectral measure of $\mathrm{GWT}(\mathrm{Poi}(\lambda))$.

- (Bordenave-Lelarge-Salez): μ_λ is symmetric
- (Salez): Dense set of atoms

The subcritical regime $np = \lambda$

➤ Asymptotically, $G(n, \frac{\lambda}{n})$ looks like a Galton-Watson tree with degree distribution $\mathrm{Poi}(\lambda)$.

(Abért-Thom-Virág) $$\mu_A \underset{n \to +\infty}{\Longrightarrow} \mu_\lambda,$$ in probability

where μ_λ is the expected spectral measure of $\mathrm{GWT}(\mathrm{Poi}(\lambda))$.

- (Bordenave-Lelarge-Salez): μ_λ is symmetric
- (Salez): Dense set of atoms
- (Bordenave-Sen-Virág): μ_λ is purely atomic iff $\lambda \leq 1$.

The subcritical regime $np = \lambda$

➤ Asymptotically, $G(n, \frac{\lambda}{n})$ looks like a Galton-Watson tree with degree distribution $\mathrm{Poi}(\lambda)$.

(Abért-Thom-Virág) $$\mu_A \underset{n\to+\infty}{\Longrightarrow} \mu_\lambda,$$ in probability

where μ_λ is the expected spectral measure of $\mathrm{GWT}(\mathrm{Poi}(\lambda))$.

- (Bordenave-Lelarge-Salez): μ_λ is symmetric
- (Salez): Dense set of atoms
- (Bordenave-Sen-Virág): μ_λ is purely atomic iff $\lambda \leq 1$.
- (Arras-Bordenave): $\mu_{\lambda,\mathrm{ac}} \neq 0$ if λ large enough.

The subcritical regime $np = \lambda$

➤ Asymptotically, $G(n, \frac{\lambda}{n})$ looks like a Galton-Watson tree with degree distribution $\mathrm{Poi}(\lambda)$.

(Abért-Thom-Virág) $\mu_A \underset{n \to +\infty}{\Longrightarrow} \mu_\lambda,$ in probability

where μ_λ is the expected spectral measure of $\mathrm{GWT}(\mathrm{Poi}(\lambda))$.

- (Bordenave-Lelarge-Salez): μ_λ is symmetric
- (Salez): Dense set of atoms
- (Bordenave-Sen-Virág): μ_λ is purely atomic iff $\lambda \leq 1$.
- (Arras-Bordenave): $\mu_{\lambda,\text{ac}} \neq 0$ if λ large enough.

(Bordenave-Caputo) μ_A satisfies a LDP with speed n.

➤ Obtained by contracting the LDP for $G(n, \lambda/n)$ with respect to the local weak topology.

The supercritical regime $np \gg \log n$

The supercritical regime $np \gg \log n$

As soon as $np \gg 1$, Wigner's law holds

$$\mu_{A/\sqrt{np}} \underset{n\to+\infty}{\Longrightarrow} \mu_{\mathsf{sc}},\ \text{in probability.}$$

The supercritical regime $np \gg \log n$

As soon as $np \gg 1$, Wigner's law holds

$$\mu_{A/\sqrt{np}} \underset{n \to +\infty}{\Longrightarrow} \mu_{\mathsf{sc}}, \text{ in probability.}$$

(A.) If $\log n \ll np \ll n$, $\mu_{A/\sqrt{np}}$ satisfies a LDP

with speed $n^2 p$ and good rate function I.

The supercritical regime $np \gg \log n$

As soon as $np \gg 1$, Wigner's law holds

$$\mu_{A/\sqrt{np}} \underset{n\to+\infty}{\Longrightarrow} \mu_{\mathsf{sc}}, \text{ in probability.}$$

(A.) If $\log n \ll np \ll n$, $\mu_{A/\sqrt{np}}$ satisfies a LDP

with speed n^2p and good rate function I.

➤ The rate function is the solution of a certain variational problem

The supercritical regime $np \gg \log n$

As soon as $np \gg 1$, Wigner's law holds

$$\mu_{A/\sqrt{np}} \underset{n\to+\infty}{\Longrightarrow} \mu_{\mathsf{sc}}, \text{ in probability.}$$

(A.) If $\log n \ll np \ll n$, $\mu_{A/\sqrt{np}}$ satisfies a LDP

with speed n^2p and good rate function I.

➤ The rate function is the solution of a certain variational problem

- $I(\mu) = 0 \Longleftrightarrow \mu = \mu_{\mathsf{sc}}$.

The supercritical regime $np \gg \log n$

As soon as $np \gg 1$, Wigner's law holds

$$\mu_{A/\sqrt{np}} \underset{n\to+\infty}{\Longrightarrow} \mu_{\mathsf{sc}}, \text{ in probability.}$$

(A.) If $\log n \ll np \ll n$, $\mu_{A/\sqrt{np}}$ satisfies a LDP

with speed n^2p and good rate function I.

➤ The rate function is the solution of a certain variational problem

- $I(\mu) = 0 \Longleftrightarrow \mu = \mu_{\mathsf{sc}}$.
- $I(\mu) = +\infty$ if μ is not symmetric.

The supercritical regime $np \gg \log n$

As soon as $np \gg 1$, Wigner's law holds

$$\mu_{A/\sqrt{np}} \underset{n\to+\infty}{\Longrightarrow} \mu_{\mathsf{sc}}, \text{ in probability.}$$

(A.) If $\log n \ll np \ll n$, $\mu_{A/\sqrt{np}}$ satisfies a LDP

with speed $n^2 p$ and good rate function I.

➤ The rate function is the solution of a certain variational problem

- $I(\mu) = 0 \Longleftrightarrow \mu = \mu_{\mathsf{sc}}$.
- $I(\mu) = +\infty$ if μ is not symmetric.

➤ The only possible deviations are around measures coming from Quadratic Vector Equations.

Quadratic Vector Equations

➤ (Girko): Equations for the Stieltjes transform of limits of empirical spectral measures of Wigner-type matrices.

Quadratic Vector Equations

➤ (Girko): Equations for the Stieltjes transform of limits of empirical spectral measures of Wigner-type matrices.

Let $W : [0,1]^2 \to \mathbb{R}_+$ be a Borel measurable symmetric kernel s.t.

$$\sup_{x\in[0,1]} \int_0^1 W(x,y)dy < +\infty,$$

i.e. W has a bounded degree function $d_W = \int_0^1 W(.,y)dy$.

Quadratic Vector Equations

➤ (Girko): Equations for the Stieltjes transform of limits of empirical spectral measures of Wigner-type matrices.

Let $W : [0,1]^2 \to \mathbb{R}_+$ be a Borel measurable symmetric kernel s.t.

$$\sup_{x \in [0,1]} \int_0^1 W(x,y)dy < +\infty,$$

i.e. W has a bounded degree function $d_W = \int_0^1 W(.,y)dy$.

The QVE associated with W is

$$-\frac{1}{m(z,x)} = z + \int_0^1 W(x,y)m(z,y)dy, \ z \in \mathbb{H}, x \in [0,1],$$

where $\mathbb{H} = \{z \in \mathbb{C} : \Im z > 0\}$.

Quadratic Vector Equations

➤ (Girko): Equations for the Stieltjes transform of limits of empirical spectral measures of Wigner-type matrices.

Let $W : [0,1]^2 \to \mathbb{R}_+$ be a Borel measurable symmetric kernel s.t.

$$\sup_{x\in[0,1]} \int_0^1 W(x,y)dy < +\infty,$$

i.e. W has a bounded degree function $d_W = \int_0^1 W(.,y)dy$.

The QVE associated with W is

$$-\frac{1}{m(z,x)} = z + \int_0^1 W(x,y)m(z,y)dy,\ z \in \mathbb{H}, x \in [0,1],$$

where $\mathbb{H} = \{z \in \mathbb{C} : \Im z > 0\}$.

(Ajanki, Erdős, Krüger): Existence and uniqueness in $\mathcal{B}^+$, the space of $m : \mathbb{H} \times [0,1] \to \mathbb{H}$ s.t. $m(z,.)$ is bounded for any $z \in \mathbb{H}$.

QVE measure of a kernel

If $m = (m(z,x))_{z\in\mathbb{H},x\in[0,1]}$ is the unique solution in $\mathcal{B}^+$ of the QVE

$$-\frac{1}{m(z,x)} = z + \int_0^1 W(x,y)m(z,y)dy,\ \ z \in \mathbb{H},$$

then for any $x \in [0,1]$,

$m(.,x)$ is the Stieltjes transform of some $v_x \in \mathcal{P}(\mathbb{R})$.

QVE measure of a kernel

If $m = (m(z,x))_{z\in\mathbb{H}, x\in[0,1]}$ is the unique solution in $\mathcal{B}^+$ of the QVE

$$-\frac{1}{m(z,x)} = z + \int_0^1 W(x,y)m(z,y)dy, \ z \in \mathbb{H},$$

then for any $x \in [0,1]$,

$m(.,x)$ is the Stieltjes transform of some $\upsilon_x \in \mathcal{P}(\mathbb{R})$.

Define the QVE measure of the kernel W as $\upsilon_W = \int_0^1 \upsilon_x dx$.

QVE measure of a kernel

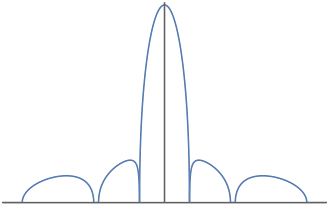
If $m = (m(z,x))_{z\in\mathbb{H}, x\in[0,1]}$ is the unique solution in $\mathcal{B}^+$ of the QVE

$$-\frac{1}{m(z,x)} = z + \int_0^1 W(x,y)m(z,y)dy, \ z \in \mathbb{H},$$

then for any $x \in [0,1]$,

$m(.,x)$ is the Stieltjes transform of some $\upsilon_x \in \mathcal{P}(\mathbb{R})$.

Define the QVE measure of the kernel W as $\upsilon_W = \int_0^1 \upsilon_x dx$.



(Ajanki, Erdős, Krüger):

υ_W is symmetric

$\mathrm{supp}(\upsilon_W) \subset [-2\|d_W\|_\infty^{1/2}, 2\|d_W\|_\infty^{1/2}]$.

QVE measure of a kernel

If $m = (m(z,x))_{z\in\mathbb{H},x\in[0,1]}$ is the unique solution in $\mathcal{B}^+$ of the QVE

$$-\frac{1}{m(z,x)} = z + \int_0^1 W(x,y)m(z,y)dy,\ \ z\in\mathbb{H},$$

then for any $x\in[0,1]$,

$m(.,x)$ is the Stieltjes transform of some $\upsilon_x \in \mathcal{P}(\mathbb{R})$.

Define the QVE measure of the kernel W as

$$\upsilon_W = \int_0^1 \upsilon_x dx.$$

(Erdős-Mülhbacher), (Zhu): Moments are tree homomorphisms densities:

QVE measure of a kernel

If $m = (m(z,x))_{z \in \mathbb{H}, x \in [0,1]}$ is the unique solution in $\mathcal{B}^+$ of the QVE

$$-\frac{1}{m(z,x)} = z + \int_0^1 W(x,y) m(z,y) dy, \;\; z \in \mathbb{H},$$

then for any $x \in [0,1]$,

$m(.,x)$ is the Stieltjes transform of some $\upsilon_x \in \mathcal{P}(\mathbb{R})$.

Define the QVE measure of the kernel W as

$$\upsilon_W = \int_0^1 \upsilon_x dx.$$

(Erdős-Mülhbacher), (Zhu): Moments are tree homomorphisms densities:

$$\upsilon_W(\tau^{2k}) = \sum_{(F,o) \in \mathcal{T}_k} t(F,W), \;\; k \in \mathbb{N}$$

where $\mathcal{T}_k$ is the set of unlabelled rooted planar trees with k edges

$$t(F,W) = \int_{[0,1]^k} \prod_{ij \in E(F)} W(x_i, x_j) \prod_{\ell=1}^{k} dx_\ell.$$

Inhomogeous sparse Erdős-Rényi graphs

Inhomogeous sparse Erdős-Rényi graphs

Take W a kernel with bounded degree function and set

$$W_{ij} = n^2 \int_{(\frac{i-1}{n},\frac{i}{n}]\times(\frac{j-1}{n},\frac{j}{n}]} W(x,y)dxdy, \quad i,j = 1,\ldots,n.$$

Inhomogeous sparse Erdős-Rényi graphs

Take W a kernel with bounded degree function and set

$$W_{ij} = n^2 \int_{(\frac{i-1}{n},\frac{i}{n}]\times(\frac{j-1}{n},\frac{j}{n}]} W(x,y)dxdy, \quad i,j = 1,\ldots,n.$$

Let $G(n,pW)$ be a graph on n vertices where the edge $\{i,j\}$ is present with probability pW_{ij} independently of the others.

Inhomogeous sparse Erdős-Rényi graphs

Take W a kernel with bounded degree function and set

$$W_{ij} = n^2 \int_{(\frac{i-1}{n},\frac{i}{n}]\times(\frac{j-1}{n},\frac{j}{n}]} W(x,y)dxdy, \quad i,j = 1,\ldots,n.$$

Let $G(n, pW)$ be a graph on n vertices where the edge $\{i,j\}$ is present with probability pW_{ij} independently of the others.

(Girko): If $np \gg 1$,

$$\mu_{A/\sqrt{np}} \underset{n\to+\infty}{\Longrightarrow} \upsilon_W, \quad \text{in probability}$$

where A is the adjacency matrix of $G(n, pW)$ and υ_W the QVE measure of W.

Inhomogeous sparse Erdős-Rényi graphs

Take W a kernel with bounded degree function and set

$$W_{ij} = n^2 \int_{(\frac{i-1}{n},\frac{i}{n}]\times(\frac{j-1}{n},\frac{j}{n}]} W(x,y)dxdy, \quad i,j = 1,\ldots,n.$$

Let $G(n,pW)$ be a graph on n vertices where the edge $\{i,j\}$ is present with probability pW_{ij} independently of the others.

(Girko): If $np \gg 1$,

$$\mu_{A/\sqrt{np}} \underset{n\to+\infty}{\Longrightarrow} \upsilon_W, \quad \text{in probability}$$

where A is the adjacency matrix of $G(n,pW)$ and υ_W the QVE measure of W.

➤ The optimal large deviation strategies of $\mu_{A/\sqrt{np}}$ correspond to changing $G(n,p)$ into one inhomogeous Erdős-Rényi graph $G(n,pW)$.

QVE measure of integrable kernels

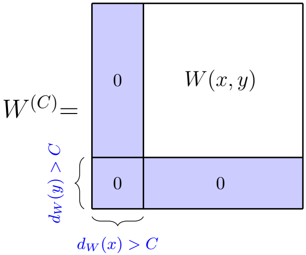
QVE measure of integrable kernels

Extend the definition of the QVE measure to integrable kernels W:

$$\upsilon_W = \lim_{C \to +\infty} \upsilon_{W^{(C)}},$$

where $W^{(C)}$ is the degree truncated kernel

$$W^{(C)}(x,y) = W(x,y) \mathbb{1}_{d_W(x) \le C} \mathbb{1}_{d_W(y) \le C}, \; (x,y) \in [0,1]^2.$$



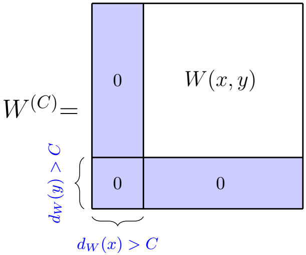
QVE measure of integrable kernels

Extend the definition of the QVE measure to integrable kernels W:

$$\upsilon_W = \lim_{C \to +\infty} \upsilon_{W^{(C)}},$$

where $W^{(C)}$ is the degree truncated kernel

$$W^{(C)}(x,y) = W(x,y) \mathbb{1}_{d_W(x) \le C} \mathbb{1}_{d_W(y) \le C}, \; (x,y) \in [0,1]^2.$$



Since $d_W \in L^1$,

$$\|W^{(C)} - W\|_1 \underset{C \to +\infty}{\longrightarrow} 0.$$

If U, V are kernels with bounded degree functions

$$\mathscr{W}_2(\upsilon_U, \upsilon_V) \le \|U - V\|_1^{1/2}$$

The rate function in the supercritical sparse case

The rate function in the supercritical sparse case

Recall that A is the adjacency matrix of $G(n,p)$.

(A.) If $\log \ll np \ll n$, then $\mu_{A/\sqrt{np}}$ satisfies a LDP with speed $n^2 p$ and rate function I:

The rate function in the supercritical sparse case

Recall that A is the adjacency matrix of $G(n,p)$.

(A.) If $\log \ll np \ll n$, then $\mu_{A/\sqrt{np}}$ satisfies a LDP with speed $n^2 p$ and rate function I:

$$I(\nu) = \inf\left\{H(W) : \upsilon_W = \nu,\ W \in \mathcal{W}\right\},\ \nu \in \mathcal{P}(\mathbb{R}),$$

where $\mathcal{W}$ is the set of integrable kernels on $[0,1]^2$ and

$$H(W) = \int_{[0,1]^2} h(W(x,y))dxdy,\ W \in \mathcal{W},$$

with $h(u) = u\log u - u + 1$.

The rate function in the supercritical sparse case

Recall that A is the adjacency matrix of $G(n,p)$.

(A.) If $\log \ll np \ll n$, then $\mu_{A/\sqrt{np}}$ satisfies a LDP with speed $n^2 p$ and rate function I:

$$I(\nu) = \inf\left\{H(W) : \upsilon_W = \nu,\ W \in \mathcal{W}\right\},\ \nu \in \mathcal{P}(\mathbb{R}),$$

where $\mathcal{W}$ is the set of integrable kernels on $[0,1]^2$ and

$$H(W) = \int_{[0,1]^2} h(W(x,y))dxdy,\ W \in \mathcal{W},$$

with $h(u) = u \log u - u + 1$.

- $H\big(G(n,pW) \mid G(n,p)\big) \simeq n^2 p H(W)$.

The rate function in the supercritical sparse case

Recall that A is the adjacency matrix of $G(n,p)$.

(A.) If $\log \ll np \ll n$, then $\mu_{A/\sqrt{np}}$ satisfies a LDP with speed n^2p and rate function I:

$$I(\nu) = \inf\left\{H(W) : \upsilon_W = \nu,\ W \in \mathcal{W}\right\},\ \nu \in \mathcal{P}(\mathbb{R}),$$

where $\mathcal{W}$ is the set of integrable kernels on $[0,1]^2$ and

$$H(W) = \int_{[0,1]^2} h(W(x,y))dxdy,\ W \in \mathcal{W},$$

with $h(u) = u\log u - u + 1$.

- $H\big(G(n,pW) \mid G(n,p)\big) \simeq n^2pH(W)$.
- If $I(\nu) < +\infty$ then the infimum is achieved.

The rate function in the supercritical sparse case

Recall that A is the adjacency matrix of $G(n,p)$.

(A.) If $\log \ll np \ll n$, then $\mu_{A/\sqrt{np}}$ satisfies a LDP with speed n^2p and rate function I:

$$I(\nu) = \inf\left\{H(W) : \upsilon_W = \nu,\ W \in \mathcal{W}\right\},\ \nu \in \mathcal{P}(\mathbb{R}),$$

where $\mathcal{W}$ is the set of integrable kernels on $[0,1]^2$ and

$$H(W) = \int_{[0,1]^2} h(W(x,y))dxdy,\ W \in \mathcal{W},$$

with $h(u) = u\log u - u + 1$.

- $H\big(G(n,pW) \mid G(n,p)\big) \simeq n^2pH(W)$.
- If $I(\nu) < +\infty$ then the infimum is achieved.
- Conditioned on a large deviation of its spectrum, $G(n,p)$ is expected to look like $G(n,pW)$ for some $W \in \mathcal{W}$.

Supercritical sparse Wigner matrices

Supercritical sparse Wigner matrices

Consider M a Wigner matrix with bounded entries,
A the adjacency matrix of $G(n, p)$ independent of M.

Set:

$$X = M \circ A,$$

where $\circ$ denotes the Hadamard product.

Supercritical sparse Wigner matrices

Consider M a Wigner matrix with bounded entries,

A the adjacency matrix of $G(n, p)$ independent of M.

Set:

$$X = M \circ A,$$

where $\circ$ denotes the Hadamard product.

➤ neither A or $A - \mathbb{E}A$ are sparse Wigner matrices but:

Supercritical sparse Wigner matrices

Consider M a Wigner matrix with bounded entries,

A the adjacency matrix of $G(n,p)$ independent of M.

Set:

$$X = M \circ A,$$

where $\circ$ denotes the Hadamard product.

➤ neither A or $A - \mathbb{E}A$ are sparse Wigner matrices but:

(Tikhomirov-Youssef): $A - \varepsilon A'$ is a sparse Wigner matrix

$A'_{ij} \sim \mathrm{Ber}(p/\varepsilon)$ and $A' \perp\!\!\!\perp A$.

Supercritical sparse Wigner matrices

Consider M a Wigner matrix with bounded entries,
A the adjacency matrix of $G(n,p)$ independent of M.

Set:

$$X = M \circ A,$$

where $\circ$ denotes the Hadamard product.

➤ neither A or $A - \mathbb{E}A$ are sparse Wigner matrices but:

(Tikhomirov-Youssef): $A - \varepsilon A'$ is a sparse Wigner matrix
$A'_{ij} \sim \mathrm{Ber}(p/\varepsilon)$ and $A' \perp\!\!\!\perp A$.

(A.) If $\log n \ll np \ll n$, $\mu_{X/\sqrt{np}}$ satisfies a LDP with speed $n^2 p$ and rate function I_L:

$$I_L(\nu) = \inf\left\{H_L(W) : \upsilon_W = \nu, W \in \mathcal{W}\right\},\ \nu \in \mathcal{P}(\mathbb{R}),$$

where $H_L(W) = \int_{[0,1]^2} h_L(W(x,y))dxdy$,

$$h_L(u) = \sup_{\theta \in \mathbb{R}}\{\theta u - L(\theta)\}, \quad L(\theta) = \mathbb{E}(e^{\theta M_{1,2}^2}) - 1,\ u, \theta \in \mathbb{R}.$$

Why does sparsity help?

Why does sparsity help?

Starting point: Schur's complement formula

$$-\frac{1}{G_{ii}(z)} = z + \sum_{j,k}^{(i)} G^{(i)}_{jk}(z) X_{ij} X_{ik}, \quad i \in \{1, \dots, n\}$$

where G, resp. $G^{(i)}$, is the resolvent of X, resp. $X^{(i)}$.

Why does sparsity help?

Starting point: Schur's complement formula

$$-\frac{1}{G_{ii}(z)} = z + \sum_{j,k}^{(i)} G_{jk}^{(i)}(z) X_{ij} X_{ik}, \quad i \in \{1, \dots, n\}$$

where G, resp. $G^{(i)}$, is the resolvent of X, resp. $X^{(i)}$.
Thus, $(G_{ii})_{1 \le i \le n}$ satisfies a perturbed QVE:

$$-\frac{1}{G_{ii}(z)} = z + \sum_{j} X_{ij}^2 G_{jj}(z) + d_i(z), \quad i \in \{1, \dots, n\}$$

where $d_i(z) = \sum_j^{(i)} (G_{jj} - G_{jj}^{(i)}) X_{ij}^2 + \sum_{j \neq k} G_{jk}^{(i)} X_{ij} X_{ik}$.

Why does sparsity help?

Starting point: Schur's complement formula

$$-\frac{1}{G_{ii}(z)} = z + \sum_{j,k}^{(i)} G^{(i)}_{jk}(z) X_{ij} X_{ik}, \quad i \in \{1, \dots, n\}$$

where G, resp. $G^{(i)}$, is the resolvent of X, resp. $X^{(i)}$.
Thus, $(G_{ii})_{1 \le i \le n}$ satisfies a perturbed QVE:

$$-\frac{1}{G_{ii}(z)} = z + \sum_j X_{ij}^2 G_{jj}(z) + d_i(z), \quad i \in \{1, \dots, n\}$$

where $d_i(z) = \sum_j^{(i)} (G_{jj} - G^{(i)}_{jj}) X_{ij}^2 + \sum_{j \neq k} G^{(i)}_{jk} X_{ij} X_{ik}$.

➤ Claim: at the large deviation scale $n^2 p$,

$$\#\{i : |d_i(z)| > \varepsilon\} = o(n),$$

for any $\varepsilon > 0, z \in \mathbb{H}$.

Why does sparsity help?

- Claim: at the large deviation scale $n^2 p$,

$$\Lambda_n := \#\{i : |d_i(z)| > \varepsilon\} = o(n),$$

for any $\varepsilon > 0, z \in \mathbb{H}$.

Why does sparsity help?

- Claim: at the large deviation scale $n^2 p$,

$$\Lambda_n := \#\{i : |d_i(z)| > \varepsilon\} = o(n),$$

for any $\varepsilon > 0, z \in \mathbb{H}$.

- Resolvent identities: $\sum_j^{(i)} (G_{jj} - G_{jj}^{(i)}) X_{ij}^2 = O(1/np)$ a.s.

Why does sparsity help?

- Claim: at the large deviation scale n^2p,

$$\Lambda_n := \#\{i : |d_i(z)| > \varepsilon\} = o(n),$$

for any $\varepsilon > 0, z \in \mathbb{H}$.

- Resolvent identities: $\sum_j^{(i)}(G_{jj} - G_{jj}^{(i)})X_{ij}^2 = O(1/np)$ a.s.
- Conditioned on $X^{(i)}$, $\sum_{j \neq k} G_{jk}^{(i)} X_{ij} X_{ik}$ is a "sparse" chaos of order 2:

Why does sparsity help?

- Claim: at the large deviation scale n^2p,

$$\Lambda_n := \#\{i : |d_i(z)| > \varepsilon\} = o(n),$$

for any $\varepsilon > 0, z \in \mathbb{H}$.

  - Resolvent identities: $\sum_j^{(i)}(G_{jj} - G_{jj}^{(i)})X_{ij}^2 = O(1/np)$ a.s.
  - Conditioned on $X^{(i)}$, $\sum_{j\neq k} G_{jk}^{(i)} X_{ij} X_{ik}$ is a "sparse" chaos of order 2:

$$\mathbb{P}^{(i)}\Big(|\sum_{j\neq k} G_{jk}^{(i)} X_{ij} X_{ik}| > \varepsilon\Big) \leq 2e^{-c_\varepsilon np\sqrt{\log(1/p)}},$$

Why does sparsity help?

- Claim: at the large deviation scale n^2p,

$$\Lambda_n := \#\{i : |d_i(z)| > \varepsilon\} = o(n),$$

for any $\varepsilon > 0, z \in \mathbb{H}$.

- Resolvent identities: $\sum_j^{(i)} (G_{jj} - G_{jj}^{(i)}) X_{ij}^2 = O(1/np)$ a.s.
- Conditioned on $X^{(i)}$, $\sum_{j \neq k} G_{jk}^{(i)} X_{ij} X_{ik}$ is a "sparse" chaos of order 2:

$$\mathbb{P}^{(i)}\Big(|\sum_{j \neq k} G_{jk}^{(i)} X_{ij} X_{ik}| > \varepsilon \Big) \leq 2e^{-c_\varepsilon np \sqrt{\log(1/p)}},$$

Thus, $\mathbb{E}(\Lambda_n) \lesssim e^{-c_\varepsilon np \sqrt{\log(1/p)}}$.

Why does sparsity help?

➤ Claim: at the large deviation scale n^2p,

$$\Lambda_n := \#\{i : |d_i(z)| > \varepsilon\} = o(n),$$

for any $\varepsilon > 0, z \in \mathbb{H}$.

- Resolvent identities: $\sum_j^{(i)}(G_{jj} - G_{jj}^{(i)})X_{ij}^2 = O(1/np)$ a.s.
- Conditioned on $X^{(i)}$, $\sum_{j\neq k} G_{jk}^{(i)} X_{ij}X_{ik}$ is a "sparse" chaos of order 2:

$$\mathbb{P}^{(i)}\Big(|\sum_{j\neq k} G_{jk}^{(i)} X_{ij}X_{ik}| > \varepsilon\Big) \leq 2e^{-c_\varepsilon np\sqrt{\log(1/p)}},$$

Thus, $\mathbb{E}(\Lambda_n) \lesssim e^{-c_\varepsilon np\sqrt{\log(1/p)}}$.
If the d_i's were independent, then by Bennett's inequality

$$\mathbb{P}(\Lambda_n > \delta n) \leq \exp\Big(-\delta n \log \frac{\delta n}{3\mathbb{E}(\Lambda_n)}\Big) \leq e^{-C_{\delta,\varepsilon} n^2 p\sqrt{\log(1/p)}}.$$

Why does sparsity help?

- Claim: at the large deviation scale n^2p,

$$\Lambda_n := \#\{i : |d_i(z)| > \varepsilon\} = o(n),$$

for any $\varepsilon > 0, z \in \mathbb{H}$.

  - Resolvent identities: $\sum_j^{(i)}(G_{jj} - G_{jj}^{(i)})X_{ij}^2 = O(1/np)$ a.s.
  - Conditioned on $X^{(i)}$, $\sum_{j\neq k} G_{jk}^{(i)} X_{ij} X_{ik}$ is a "sparse" chaos of order 2:

$$\mathbb{P}^{(i)}\Big(|\sum_{j\neq k} G_{jk}^{(i)} X_{ij} X_{ik}| > \varepsilon\Big) \leq 2e^{-c_\varepsilon np\sqrt{\log(1/p)}},$$

Thus, $\mathbb{E}(\Lambda_n) \lesssim e^{-c_\varepsilon np\sqrt{\log(1/p)}}$.
If the d_i's were independent, then by Bennett's inequality

$$\mathbb{P}(\Lambda_n > \delta n) \leq \exp\Big(- \delta n \log \frac{\delta n}{3\mathbb{E}(\Lambda_n)}\Big) \leq e^{-C_{\delta,\varepsilon} n^2 p\sqrt{\log(1/p)}}.$$

- Develop a Bennett-type inequality for dependent variables.

Thank you for your attention!