Random Matrices with Log-Range Correlations

BIRS Workshop

Analytic vs. Combinatorial in Free Probability

Todd Kemp

UC San Diego

December 5, 2016

Empirical Spectral Distribution

Let X be a symmetric real matrix, with eigenvalues $\lambda_1 \le \cdots \le \lambda_N$. The **empirical spectral distribution** (**ESD**) of X_N is the discrete probability measure

$$\mu = \frac{1}{N}\sum_{j=1}^{N} \delta_{\lambda_j}.$$

(Note: if X is a random matrix, then μ is a random measure.)

Empirical Spectral Distribution

Let X be a symmetric real matrix, with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_N$.
The **empirical spectral distribution** (**ESD**) of X_N is the discrete probability measure

$$\mu = \frac{1}{N}\sum_{j=1}^{N} \delta_{\lambda_j}.$$

(Note: if X is a random matrix, then μ is a random measure.)
The typical quantities of interest are the **linear statistics** of the matrix:

$$\int f\, d\mu = \frac{1}{N}\sum_{j=1}^{N} f(\lambda_j) = \frac{1}{N}\mathrm{Tr}(f(X_N)).$$

A sequence of random measures μ_N converges to a deterministic measure σ weakly **in expectation** if

$$\mathbb{E}\left(\int f\, d\mu_N\right) \to \int f\, d\sigma \qquad \forall f \in C_c(\mathbb{R})$$

Empirical Spectral Distribution

Let X be a symmetric real matrix, with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_N$. The **empirical spectral distribution** (**ESD**) of X_N is the discrete probability measure

$$\mu = \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j}.$$

(Note: if X is a random matrix, then μ is a random measure.)
The typical quantities of interest are the **linear statistics** of the matrix:

$$\int f\, d\mu = \frac{1}{N} \sum_{j=1}^{N} f(\lambda_j) = \frac{1}{N} \mathrm{Tr}(f(X_N)).$$

A sequence of random measures μ_N converges to a deterministic measure σ weakly **in probability** if

$$\mathbb{P}\left(\left| \int f\, d\mu_N - \int f\, d\sigma \right| > \epsilon \right) \to 0 \qquad \forall f \in C_c(\mathbb{R}), \epsilon > 0.$$

Empirical Spectral Distribution

Let X be a symmetric real matrix, with eigenvalues $\lambda_1 \le \cdots \le \lambda_N$. The **empirical spectral distribution** (**ESD**) of X_N is the discrete probability measure

$$\mu = \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j}.$$

(Note: if X is a random matrix, then μ is a random measure.)
The typical quantities of interest are the **linear statistics** of the matrix:

$$\int f\, d\mu = \frac{1}{N} \sum_{j=1}^{N} f(\lambda_j) = \frac{1}{N} \mathrm{Tr}(f(X_N)).$$

A sequence of random measures μ_N converges to a deterministic measure σ weakly **almost surely** if

$$\mathbb{P}\left(\lim_{N\to\infty} \int f\, d\mu_N = \int f\, d\sigma \right) = 1 \qquad \forall f \in C_c(\mathbb{R}).$$

Wigner's Semicircle Law

Theorem. Let X_N be a symmetric matrix whose upper-triangular entries are i.i.d., centered with variance $\frac{1}{N}$. Then the ESD of X_N converges weakly almost surely to the semicircle law

$$\sigma(dx) = \frac{1}{2\pi}\sqrt{(4-x^2)_+}\,dx.$$

Wigner's Semicircle Law

Theorem. Let X_N be a symmetric matrix whose upper-triangular entries are i.i.d., centered with variance $\frac{1}{N}$. Then the ESD of X_N converges weakly almost surely to the semicircle law

$$\sigma(dx) = \frac{1}{2\pi}\sqrt{(4-x^2)_+}\,dx.$$

- Originally proved by Wigner in 1955, for standard normal entries (GOE_N), and with convergence in expectation only.

Wigner's Semicircle Law

Theorem. Let X_N be a symmetric matrix whose upper-triangular entries are i.i.d., centered with variance $\frac{1}{N}$. Then the ESD of X_N converges weakly almost surely to the semicircle law

$$\sigma(dx) = \frac{1}{2\pi}\sqrt{(4-x^2)_+}\,dx.$$

- Originally proved by Wigner in 1955, for standard normal entries (GOE_N), and with convergence in expectation only.
- Upgraded to a.s. converge in the 1960s — really to convergence in probability, with an explicit estimate on the rate of convergence that is summable ($O(1/N^2)$), thus yielding a.s. convergence by the Borel–Cantelli Lemma.

Wigner's Semicircle Law

Theorem. Let X_N be a symmetric matrix whose upper-triangular entries are i.i.d., centered with variance $\frac{1}{N}$. Then the ESD of X_N converges weakly almost surely to the semicircle law

$$\sigma(dx) = \frac{1}{2\pi}\sqrt{(4-x^2)_+}\,dx.$$

- Originally proved by Wigner in 1955, for standard normal entries (GOE_N), and with convergence in expectation only.
- Upgraded to a.s. converge in the 1960s — really to convergence in probability, with an explicit estimate on the rate of convergence that is summable ($O(1/N^2)$), thus yielding a.s. convergence by the Borel–Cantelli Lemma.
- Generalized to entries with any distribution having at least 2 finite moments, using similar combinatorial techniques (the method of moments).

Combinatorial Approach to Wigner's Law

Prove it for polynomial test functions f. For $f(x) = x^k$

$$\mathbb{E}\mathrm{Tr}[(X_N)^k] = \sum_{i_1,\dots,i_k} \mathbb{E}([X_N]_{i_1 i_2}[X_N]_{i_2 i_3}\cdots[X_N]_{i_k i_1}).$$

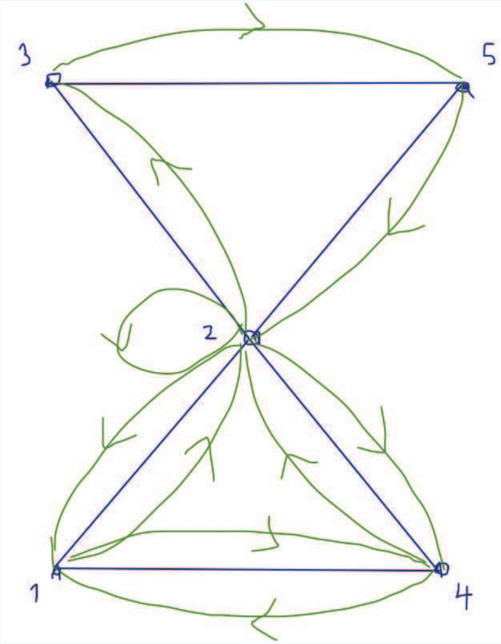
Use independence and identical distribution to collect terms; each one is associated to a walk on a graph.

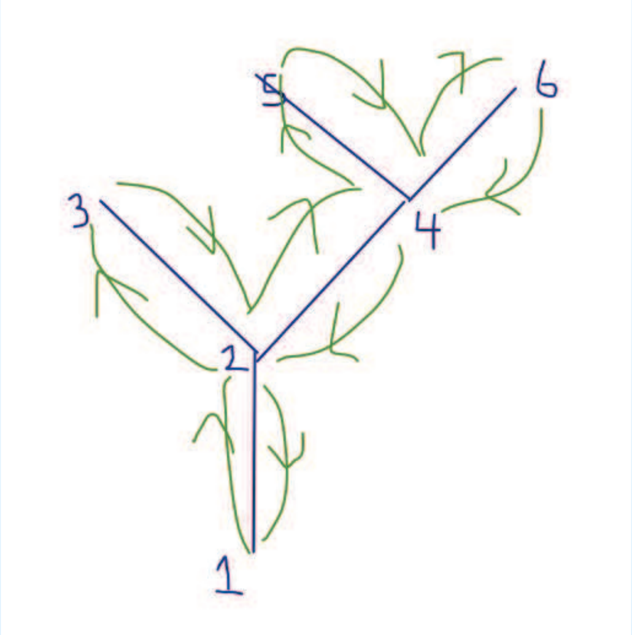
Combinatorial Approach to Wigner's Law

Prove it for polynomial test functions f. For $f(x) = x^k$

$$\mathbb{E}\mathrm{Tr}[(X_N)^k] = \sum_{i_1,\ldots,i_k} \mathbb{E}([X_N]_{i_1 i_2}[X_N]_{i_2 i_3}\cdots[X_N]_{i_k i_1}).$$

Use independence and identical distribution to collect terms; each one is associated to a walk on a graph.





Combinatorial Approach to Wigner's Law

Prove it for polynomial test functions f. For $f(x) = x^k$

$$\mathbb{E}\mathrm{Tr}[(X_N)^k] = \sum_{i_1,\ldots,i_k} \mathbb{E}([X_N]_{i_1 i_2}[X_N]_{i_2 i_3}\cdots[X_N]_{i_k i_1}).$$

Use independence and identical distribution to collect terms; each one is associated to a walk on a graph.

Using variance $= \frac{1}{N}$, find that the terms that contribute in the limit are rooted trees traversed in the unique path hitting every edge twice. Count these up, get Catalan number $=$ the moments of the semicircle law.

Combinatorial Approach to Wigner's Law

Prove it for polynomial test functions f. For $f(x) = x^k$

$$\mathbb{E}\mathrm{Tr}[(X_N)^k] = \sum_{i_1,\ldots,i_k} \mathbb{E}([X_N]_{i_1 i_2}[X_N]_{i_2 i_3} \cdots [X_N]_{i_k i_1}).$$

Use independence and identical distribution to collect terms; each one is associated to a walk on a graph.

Using variance $= \frac{1}{N}$, find that the terms that contribute in the limit are rooted trees traversed in the unique path hitting every edge twice. Count these up, get Catalan number $=$ the moments of the semicircle law.

For a.s. convergence: follow the same method to expand $\mathrm{Var}\left(\int x^k \mu_N(dx)\right)$; find that it is a sum of terms in correspondence with pairs of walks on graphs with certain constraints. Count these, find overall $O(1/N^2)$ contribution.

Band Matrices: i. but not i.d. Entries

If we relax the condition that the upper-triangular entries are i.i.d., these methods do not work.

Band Matrices: i. but not i.d. Entries

If we relax the condition that the upper-triangular entries are i.i.d., these methods do not work. But with some uniformity assumptions, the same result holds. For example, Anderson–Zeitouni showed that if the entries are still independent, and have the form

$$\sqrt{N}[X_N]_{ij} = g(i/N, j/N)\xi_{ij}$$

where ξ_{ij} are i.i.d. and "nice", and $\int_{[0,1]^2} g(x,y)\,dxdy = 1$, then the ESD of X_N converges weakly a.s. to the semicircle law.

Band Matrices: i. but not i.d. Entries

If we relax the condition that the upper-triangular entries are i.i.d., these methods do not work. But with some uniformity assumptions, the same result holds. For example, Anderson–Zeitouni showed that if the entries are still independent, and have the form

$$\sqrt{N}[X_N]_{ij} = g(i/N, j/N)\xi_{ij}$$

where ξ_{ij} are i.i.d. and "nice", and $\int_{[0,1]^2} g(x,y)\,dxdy = 1$, then the ESD of X_N converges weakly a.s. to the semicircle law.

If we **abandon independence**, however, the semicircle law is lost: it is the fixed point in the universality class of Wigner ensembles (with independent entries) only.

Band Matrices: i. but not i.d. Entries

If we relax the condition that the upper-triangular entries are i.i.d., these methods do not work. But with some uniformity assumptions, the same result holds. For example, Anderson–Zeitouni showed that if the entries are still independent, and have the form

$$\sqrt{N}[X_N]_{ij} = g(i/N, j/N)\xi_{ij}$$

where ξ_{ij} are i.i.d. and "nice", and $\int_{[0,1]^2} g(x,y)\, dxdy = 1$, then the ESD of X_N converges weakly a.s. to the semicircle law.

If we **abandon independence**, however, the semicircle law is lost: it is the fixed point in the universality class of Wigner ensembles (with independent entries) only.

The combinatorial methods of free probability can still be used to understand the limiting ESD of some matrices with correlated entries, however…

Block Matrices and Operator-Valued Free Probability

Consider a random matrix of a form like this:

$$\begin{bmatrix} X_N & Y_N \\ Y_N & Z_N \end{bmatrix}$$

where X_N, Y_N, Z_N are GOE_N matrices, but are not independent from each other; instead we specify the correlations between their entires. For simplicity we assume that the correlations between entries of X_N and Y_N are the same for each pair of entries $[X_N]_{ab}$ and $[Y_N]_{cd}$, etc.

Block Matrices and Operator-Valued Free Probability

Consider a random matrix of a form like this:

$$\begin{bmatrix} X_N & Y_N \\ Y_N & Z_N \end{bmatrix}$$

where X_N, Y_N, Z_N are GOE_N matrices, but are not independent from each other; instead we specify the correlations between their entires. For simplicity we assume that the correlations between entries of X_N and Y_N are the same for each pair of entries $[X_N]_{ab}$ and $[Y_N]_{cd}$, etc.

In 1996, Shlyakhtenko showed that this ensemble converges in $*$-distribution to an operator-valued semicircular operator:

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix}$$

where the $\{x, y, z\}$ is a semicircular family with correlations matching the ones above.

Block Matrices and a "Genus Expansion"

Say your matrix X_N is $mN \times mN$, with GOE_N blocks in an $m \times m$ array with specified covariances. By encoding the covariances between the blocks as a certain mapping $\eta\colon \mathbb{M}_m \to \mathbb{M}_m$, Speicher showed one can construct the usual non-crossing cumulant formalism for such block matrices to efficiently compute moments of their limit ESDs.

Block Matrices and a “Genus Expansion”

Say your matrix X_N is $mN \times mN$, with GOE_N blocks in an $m \times m$ array with specified covariances. By encoding the covariances between the blocks as a certain mapping $\eta \colon \mathbb{M}_m \to \mathbb{M}_m$, Speicher showed one can construct the usual non-crossing cumulant formalism for such block matrices to efficiently compute moments of their limit ESDs. Also Bryc, Oraby, Rashidi Far and Speicher used Cauchy transform methods to prove a.s. convergence in such models.

Block Matrices and a “Genus Expansion”

Say your matrix X_N is $mN \times mN$, with GOE_N blocks in an $m \times m$ array with specified covariances. By encoding the covariances between the blocks as a certain mapping $\eta\colon \mathbb{M}_m \to \mathbb{M}_m$, Speicher showed one can construct the usual non-crossing cumulant formalism for such block matrices to efficiently compute moments of their limit ESDs. Also Bryc, Oraby, Rashidi Far and Speicher used Cauchy transform methods to prove a.s. convergence in such models.

All of this analysis requires m to be fixed. It would be good (but challenging) to extend the analysis to allow m to grow with N. A starting point is the following “genus expansion” in the Gaussian case:

$$\frac{1}{mN}\mathbb{E}\mathrm{Tr}[(X_N)^{2k}] = \sum_{\pi \in \mathscr{P}_2(2k)} \frac{\alpha_\pi(N)}{N^{k+1}} \frac{1}{m^{k+1}} \sum_{i_1,\ldots,i_{2k}=1}^{m} \prod_{(\alpha,\beta)\in\pi} \mathrm{Cov}_{i_\alpha,i_{\alpha+1},i_\beta,i_{\beta+1}}$$

“Block” Matrices with Log-Range Correlations

A nice change of basis converts these matrices into the dual block form: an overall $N \times N$ block grid with independent $m \times m$ blocks in each entry. That’s the perspective we take for our main theorem.

“Block” Matrices with Log-Range Correlations

A nice change of basis converts these matrices into the dual block form: an overall $N \times N$ block grid with independent $m \times m$ blocks in each entry. That’s the perspective we take for our main theorem.

Theorem. [K, Zimmermann *late* 2016] Let X_N be a random symmetric matrix whose entries are uniformly square integrable. Let μ_N denote the ESD of X_N.

For each N, suppose there is a constant $d_N = o(\log N)$, and a partition of $\{(i,j)\colon 1 \le i \le j \le N\}$ with blocks of size $\le d_N$, such that entries of X_N in different blocks are independent.

“Block” Matrices with Log-Range Correlations

A nice change of basis converts these matrices into the dual block form: an overall $N \times N$ block grid with independent $m \times m$ blocks in each entry. That’s the perspective we take for our main theorem.

Theorem. [K, Zimmermann *late* 2016] Let X_N be a random symmetric matrix whose entries are uniformly square integrable. Let μ_N denote the ESD of X_N.

For each N, suppose there is a constant $d_N = o(\log N)$, and a partition of $\{(i,j)\colon 1 \le i \le j \le N\}$ with blocks of size $\le d_N$, such that entries of X_N in different blocks are independent. Then for any $f \in \mathrm{Lip}(\mathbb{R})$,

$$\int f\, d\mu_N - \mathbb{E}\left(\int f\, d\mu_N\right) \to_{\mathbb{P}} 0.$$

“Block” Matrices with Log-Range Correlations

A nice change of basis converts these matrices into the dual block form: an overall $N \times N$ block grid with independent $m \times m$ blocks in each entry. That’s the perspective we take for our main theorem.

Theorem. [K, Zimmermann *late* 2016] Let X_N be a random symmetric matrix whose entries are uniformly square integrable. Let μ_N denote the ESD of X_N.

For each N, suppose there is a constant $d_N = o(\log N)$, and a partition of $\{(i,j)\colon 1 \le i \le j \le N\}$ with blocks of size $\le d_N$, such that entries of X_N in different blocks are independent. Then for any $f \in \mathrm{Lip}(\mathbb{R})$,

$$\int f\, d\mu_N - \mathbb{E}\left(\int f\, d\mu_N\right) \to_{\mathbb{P}} 0.$$

(In the block matrices studied above, this handles the case $m = o(\sqrt{\log N})$. But it is much more general.)

Cutoffs and Random Gaussian Noise

The idea of the proof is as follows. First, using the uniform square integrability, a fairly standard argument shows that X_N can be replaced by a cutoff:

$$\text{replace } \; [X_N]_{ij} \; \text{ with } \; [\widehat{X}_N] := [X_N]_{ij}\mathbb{1}_{\sqrt{N}|[X_N]_{ij}|\le C}.$$

Proving the theorem for $\widehat{X}_N$ then suffices to prove it for X_N; so we may assume $\sqrt{N}X_N$ has uniformly bounded entries.

Cutoffs and Random Gaussian Noise

The idea of the proof is as follows. First, using the uniform square integrability, a fairly standard argument shows that X_N can be replaced by a cutoff:

$$\text{replace } [X_N]_{ij} \text{ with } [\widehat{X}_N] := [X_N]_{ij}\mathbb{1}_{\sqrt{N}|[X_N]_{ij}|\leq C}.$$

Proving the theorem for $\widehat{X}_N$ then suffices to prove it for X_N; so we may assume $\sqrt{N}X_N$ has uniformly bounded entries.

Next, add some Gaussian noise:

$$\widetilde{X}_N = \widehat{X}_N + tG_N$$

where G_N is a GOE_N independent from $\widehat{X}_N$. The goal is to show that the theorem holds for $\widetilde{X}_N$ for each t, and that one can let $t = t_N \to 0$ and recover the theorem for $\widehat{X}_N$.

Logarithmic Sobolev Inequalities (LSI)

The LSI is a coercive functional inequality. A measure on $\mathbb{R}^d$ satisfies the LSI with constant c if, for f with $\int f^2\,d\mu = 1$,

$$\int f^2 \log f^2\,d\mu = \mathrm{Ent}_\mu(f^2) \le c \int |\nabla f|^2\,d\mu.$$

Logarithmic Sobolev Inequalities (LSI)

The LSI is a coercive functional inequality. A measure on $\mathbb{R}^d$ satisfies the LSI with constant c if, for f with $\int f^2\, d\mu = 1$,

$$\int f^2 \log f^2\, d\mu = \mathrm{Ent}_\mu(f^2) \le c \int |\nabla f|^2\, d\mu.$$

It was first written down by Stam in 1959 (in a different form) for Gaussian measures. It was rediscovered and named by L. Gross in 1973. Since then, it has been used in literally thousands of papers, with applications to quantum field theory, geometric analysis, stochastic analysis, Markov chains, interacting particle systems, large deviations, optimal transport, random matrix theory, . . .

Logarithmic Sobolev Inequalities (LSI)

The LSI is a coercive functional inequality. A measure on $\mathbb{R}^d$ satisfies the LSI with constant c if, for f with $\int f^2\,d\mu = 1$,

$$\int f^2 \log f^2\,d\mu = \mathrm{Ent}_\mu(f^2) \le c\int |\nabla f|^2\,d\mu.$$

It was first written down by Stam in 1959 (in a different form) for Gaussian measures. It was rediscovered and named by L. Gross in 1973. Since then, it has been used in literally thousands of papers, with applications to quantum field theory, geometric analysis, stochastic analysis, Markov chains, interacting particle systems, large deviations, optimal transport, random matrix theory, . . .

Herbst concentration argument: for $F \in \mathrm{Lip}(\mathbb{R}^d)$ and $\mathbf{X} \sim \mu$,

$$\mathbb{P}(|F(\mathbf{X}) - \mathbb{E}(F(\mathbf{X}))| \ge \epsilon) \le 2\exp\left(-\frac{\epsilon^2}{c\|f\|_{\mathrm{Lip}}^2}\right).$$

Log Sobolev Inequalities in Random Matrix Theory

Proposition G. Let X_N be a symmetric random matrix. If the joint law of entries of $\sqrt{N}X_N$ satisfies the LSI with constant c, then for all $\epsilon > 0$ and $f \in \mathrm{Lip}(\mathbb{R})$,

$$\mathbb{P}\left(\left|\int f\, d\mu_N - \mathbb{E}\left(\int f\, d\mu_N\right)\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-N^2\epsilon^2}{4c||f||_{\mathrm{Lip}}^2}\right).$$

Log Sobolev Inequalities in Random Matrix Theory

Proposition G. Let X_N be a symmetric random matrix. If the joint law of entries of $\sqrt{N}X_N$ satisfies the LSI with constant c, then for all $\epsilon > 0$ and $f \in \mathrm{Lip}(\mathbb{R})$,

$$\mathbb{P}\left(\left|\int f\, d\mu_N - \mathbb{E}\left(\int f\, d\mu_N\right)\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-N^2\epsilon^2}{4c||f||_{\mathrm{Lip}}^2}\right).$$

This was essentially proved by Guionnet. She deduced the result under the stronger hypothesis that X_N has i.i.d. upper-triangular entries, *each* satisfying the LSI with constant c.

Log Sobolev Inequalities in Random Matrix Theory

Proposition G. Let X_N be a symmetric random matrix. If the joint law of entries of $\sqrt{N}X_N$ satisfies the LSI with constant c, then for all $\epsilon > 0$ and $f \in \mathrm{Lip}(\mathbb{R})$,

$$\mathbb{P}\left(\left|\int f\,d\mu_N - \mathbb{E}\left(\int f\,d\mu_N\right)\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-N^2\epsilon^2}{4c||f||^2_{\mathrm{Lip}}}\right).$$

This was essentially proved by Guionnet. She deduced the result under the stronger hypothesis that X_N has i.i.d. upper-triangular entries, *each* satisfying the LSI with constant c. This is stronger because of:

Segal's Lemma. If μ_1 satisfies the LSI with constant c_1 and μ_2 satisfies the LSI with constant c_2, then $\mu_1 \otimes \mu_2$ satisfies the LSI with constant $\max\{c_1, c_2\}$.

Log Sobolev Inequalities in Random Matrix Theory

Proposition G. Let X_N be a symmetric random matrix. If the joint law of entries of $\sqrt{N}X_N$ satisfies the LSI with constant c, then for all $\epsilon > 0$ and $f \in \mathrm{Lip}(\mathbb{R})$,

$$\mathbb{P}\left(\left|\int f\,d\mu_N - \mathbb{E}\left(\int f\,d\mu_N\right)\right| \geq \epsilon\right) \leq 2\exp\left(\frac{-N^2\epsilon^2}{4c||f||^2_{\mathrm{Lip}}}\right).$$

This was essentially proved by Guionnet. She deduced the result under the stronger hypothesis that X_N has i.i.d. upper-triangular entries, *each* satisfying the LSI with constant c. This is stronger because of:

Segal's Lemma. If μ_1 satisfies the LSI with constant c_1 and μ_2 satisfies the LSI with constant c_2, then $\mu_1 \otimes \mu_2$ satisfies the LSI with constant $\max\{c_1, c_2\}$.

Proposition G is proved noticing that, if $f \in \mathrm{Lip}(\mathbb{R})$, then $F\colon X \mapsto \mathrm{Tr}(f(X))$ is $\mathrm{Lip}(\mathbb{M}_N^{\mathrm{s.a.}})$ with $\|F\|_{\mathrm{Lip}} \leq \|f\|_{\mathrm{Lip}}$.

Mollified Log Sobolev Inequalities

Theorem KZ. Let $\mathbf{X}$ be a bounded random vector and let $\mathbf{G}$ be a standard normal random vector on $\mathbb{R}^d$. Then for each $t > 0$, $\mathrm{Law}_{\mathbf{X}+t\mathbf{G}}$ satisfies the LSI with constant

$$c(t) \le 289 |||\mathbf{X}|||_\infty^2 \exp\left(20d + \frac{5|||\mathbf{X}|||_\infty^2}{t}\right).$$

Mollified Log Sobolev Inequalities

Theorem KZ. Let $\mathbf{X}$ be a bounded random vector and let $\mathbf{G}$ be a standard normal random vector on $\mathbb{R}^d$. Then for each $t > 0$, $\mathrm{Law}_{\mathbf{X}+t\mathbf{G}}$ satisfies the LSI with constant

$$c(t) \le 289 \| |\mathbf{X}| \|_\infty^2 \exp\left(20d + \frac{5\| |\mathbf{X}| \|_\infty^2}{t}\right).$$

The 289 and the $20d$ are probably not sharp (probably should be independent of dimension). But the $\exp(C/t)$ behavior is sharp.

Mollified Log Sobolev Inequalities

Theorem KZ. Let $\mathbf{X}$ be a bounded random vector and let $\mathbf{G}$ be a standard normal random vector on $\mathbb{R}^d$. Then for each $t > 0$, $\mathrm{Law}_{\mathbf{X}+t\mathbf{G}}$ satisfies the LSI with constant

$$c(t) \le 289 \||\mathbf{X}|\|_\infty^2 \exp\left(20d + \frac{5\||\mathbf{X}|\|_\infty^2}{t}\right).$$

The 289 and the $20d$ are probably not sharp (probably should be independent of dimension). But the $\exp(C/t)$ behavior is sharp.

To apply this to $\widetilde{X}_N = \widehat{X}_N + tG_N$, break up into random vectors corresponding to the entries of the blocks of the partition. Each has dimension $\le d_N$. Apply the above theorem with

$$t_N = \frac{Cd_N}{\log N - 21 d_N}; \qquad \therefore c(t_N) = O(N).$$

By Segal's lemma, get $\mathrm{Law}_{\widetilde{X}_N}$ satisfies LSI with constant $c(t_N)$.

Comments on the Proofs

Getting from $\widetilde{X}_N$ to $\widehat{X}_N$:

Compare $\int f\, d\widehat{\mu}_N - \mathbb{E}\left(\int f\, d\widehat{\mu}_N\right)$ to $\int f\, d\widetilde{\mu}_N - \mathbb{E}\left(\int f\, d\widetilde{\mu}_N\right)$ with a standard $\epsilon/3$ argument.

Comments on the Proofs

Getting from $\widetilde{X}_N$ to $\widehat{X}_N$:

Compare $\int f\, d\widehat{\mu}_N - \mathbb{E}\left(\int f\, d\widehat{\mu}_N\right)$ to $\int f\, d\widetilde{\mu}_N - \mathbb{E}\left(\int f\, d\widetilde{\mu}_N\right)$ with a standard $\epsilon/3$ argument. The cross terms can be estimated (again using the Hoffman–Wielandt lemma) to give a term proportional to t_N, which tends to 0.

Comments on the Proofs

Getting from $\widetilde{X}_N$ to $\widehat{X}_N$:

Compare $\int f\, d\widehat{\mu}_N - \mathbb{E}\left(\int f\, d\widehat{\mu}_N\right)$ to $\int f\, d\widetilde{\mu}_N - \mathbb{E}\left(\int f\, d\widetilde{\mu}_N\right)$ with a standard $\epsilon/3$ argument. The cross terms can be estimated (again using the Hoffman–Wielandt lemma) to give a term proportional to t_N, which tends to 0. (Note: even if $c(t)$ were independent of dimension, we still need $d_N = O(\log N)$ to make $t_N \to 0$.)

Comments on the Proofs

Getting from $\widetilde{X}_N$ to $\widehat{X}_N$:

Compare $\int f\,d\widehat{\mu}_N - \mathbb{E}\left(\int f\,d\widehat{\mu}_N\right)$ to $\int f\,d\widetilde{\mu}_N - \mathbb{E}\left(\int f\,d\widetilde{\mu}_N\right)$ with a standard $\epsilon/3$ argument. The cross terms can be estimated (again using the Hoffman–Wielandt lemma) to give a term proportional to t_N, which tends to 0. (Note: even if $c(t)$ were independent of dimension, we still need $d_N = O(\log N)$ to make $t_N \to 0$.)

Proving Theorem KZ:

We use the "Lyapunov" approach, carefully tracking the dependence of the LSI constant on the Lyapunov exponents.

Comments on the Proofs

Getting from $\widetilde{X}_N$ to $\widehat{X}_N$:

Compare $\int f\,d\widehat{\mu}_N - \mathbb{E}\left(\int f\,d\widehat{\mu}_N\right)$ to $\int f\,d\widetilde{\mu}_N - \mathbb{E}\left(\int f\,d\widetilde{\mu}_N\right)$ with a standard $\epsilon/3$ argument. The cross terms can be estimated (again using the Hoffman–Wielandt lemma) to give a term proportional to t_N, which tends to 0. (Note: even if $c(t)$ were independent of dimension, we still need $d_N = O(\log N)$ to make $t_N \to 0$.)

Proving Theorem KZ:

We use the "Lyapunov" approach, carefully tracking the dependence of the LSI constant on the Lyapunov exponents. There is also an "elementary" proof (giving a worse constant) that goes like this:

If μ satisfies LSI with constant c, and $F \in \mathrm{Lip}(\mathbb{R}^d)$, then $F_*\mu$ satisfies LSI with constant $c\|F\|_{\mathrm{Lip}}^2$.

Comments on the Proofs

Getting from $\widetilde{X}_N$ to $\widehat{X}_N$:

Compare $\int f\, d\widehat{\mu}_N - \mathbb{E}\left(\int f\, d\widehat{\mu}_N\right)$ to $\int f\, d\widetilde{\mu}_N - \mathbb{E}\left(\int f\, d\widetilde{\mu}_N\right)$ with a standard $\epsilon/3$ argument. The cross terms can be estimated (again using the Hoffman–Wielandt lemma) to give a term proportional to t_N, which tends to 0. (Note: even if $c(t)$ were independent of dimension, we still need $d_N = O(\log N)$ to make $t_N \to 0$.)

Proving Theorem KZ:

We use the "Lyapunov" approach, carefully tracking the dependence of the LSI constant on the Lyapunov exponents. There is also an "elementary" proof (giving a worse constant) that goes like this:

If μ satisfies LSI with constant c, and $F \in \mathrm{Lip}(\mathbb{R}^d)$, then $F_*\mu$ satisfies LSI with constant $c\|F\|_{\mathrm{Lip}}^2$. It turns out that $\mathrm{Law}_{\mathbf{X}+t\mathbf{G}}$ *is* the push-forward of $\mathrm{Law}_{\mathbf{G}}$ under some Lipschitz map F_t.

Matrix Models to Look At

Our theorem gives a framework for upgrading *convergence in expectation* to *convergence in probability* for a wide class of matrix models. Now we need to find matrix models that converge in expectation!

Matrix Models to Look At

Our theorem gives a framework for upgrading *convergence in expectation* to *convergence in probability* for a wide class of matrix models. Now we need to find matrix models that converge in expectation!

E.g. Let X_N be a block matrix with $m \times m$ independent blocks (where m may grow with N) of the following form: given ℓ independent GOE_m matrices $G_m^{(k)}$,

$$\sum_{k=1}^{\ell} \left(A_m^{(k)} G_m^{(k)} B_m^{(k)} + B_m^{(k)*} G_m^{(k)} A_m^{(k)*} \right)$$

where $\{A_m^{(k)}, B_m^{(k)}\}_{k=1}^{\ell}$ has a limit $*$-distribution as $m \to \infty$. Can check using standard free probability techniques that you get a limit $*$-distribution if ℓ is fixed.

Matrix Models to Look At

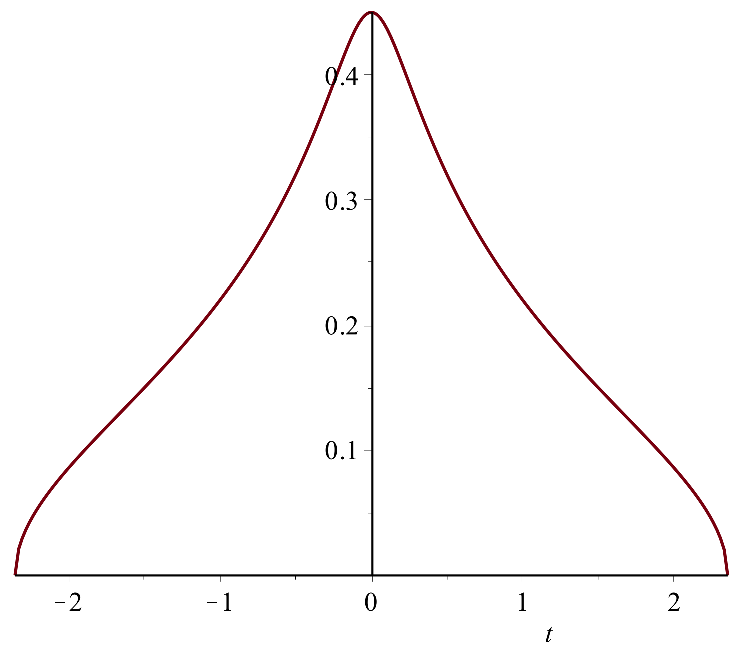
Our theorem gives a framework for upgrading *convergence in expectation* to *convergence in probability* for a wide class of matrix models. Now we need to find matrix models that converge in expectation!

E.g. Special case $\frac{1}{\sqrt{2}}(A_m G_m + G_m A_m)$, where A_m is a deterministic sequence converging in $*$-distribution to the semicircular distribution. ESD $\to$ *Tetilla Law*:

Matrix Models to Look At

Our theorem gives a framework for upgrading *convergence in expectation* to *convergence in probability* for a wide class of matrix models. Now we need to find matrix models that converge in expectation!

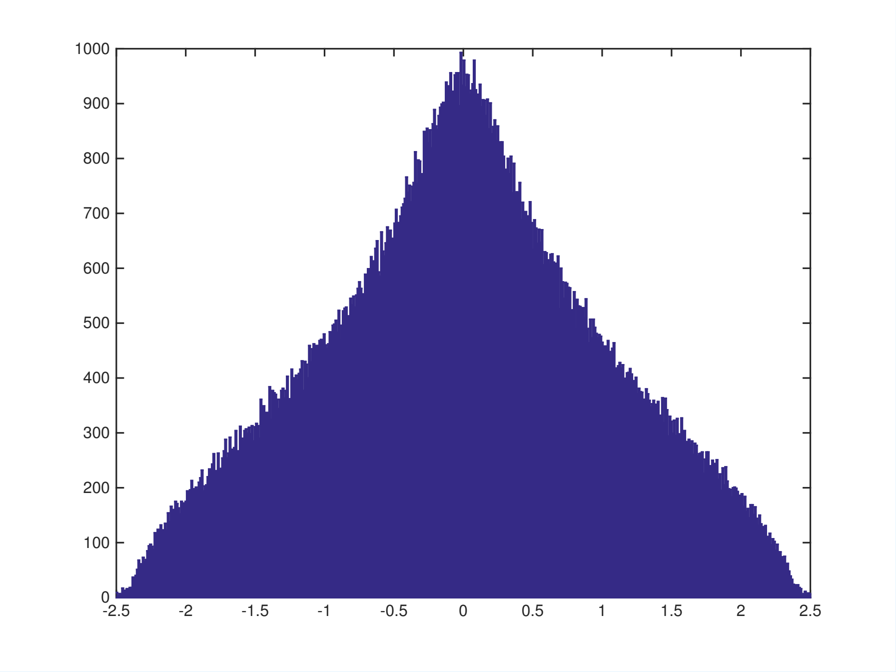
E.g. Special case $\frac{1}{\sqrt{2}}(A_m G_m + G_m A_m)$, where A_m is a deterministic sequence converging in $*$-distribution to the semicircular distribution. ESD $\to$ *Tetilla Law*:



Matrix Models to Look At

Our theorem gives a framework for upgrading *convergence in expectation* to *convergence in probability* for a wide class of matrix models. Now we need to find matrix models that converge in expectation!

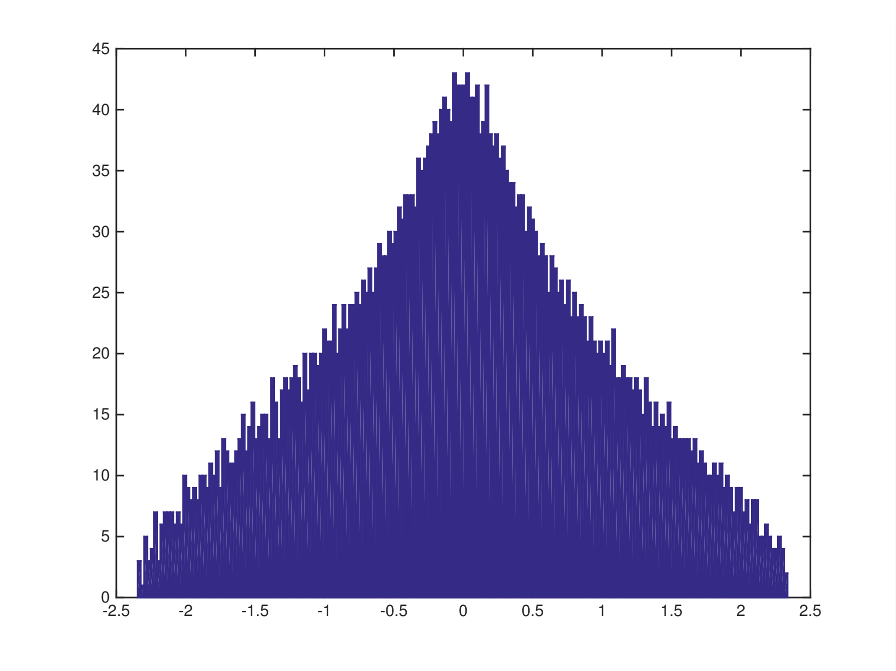
E.g. Special case $\frac{1}{\sqrt{2}}(A_m G_m + G_m A_m)$, where A_m is a deterministic sequence converging in $*$-distribution to the semicircular distribution. ESD $\rightarrow$ *Tetilla Law*:



Matrix Models to Look At

Our theorem gives a framework for upgrading *convergence in expectation* to *convergence in probability* for a wide class of matrix models. Now we need to find matrix models that converge in expectation!

E.g. Special case $\frac{1}{\sqrt{2}}(A_m G_m + G_m A_m)$, where A_m is a deterministic sequence converging in $*$-distribution to the semicircular distribution. ESD $\rightarrow$ *Tetilla Law*:



Matrix Models to Look At

Our theorem gives a framework for upgrading *convergence in expectation* to *convergence in probability* for a wide class of matrix models. Now we need to find matrix models that converge in expectation!

E.g. Let X_N be a block matrix with $m \times m$ independent blocks (where m may grow with N) of the following form: given ℓ independent GOE_m matrices $G_m^{(k)}$,

$$\sum_{k=1}^{\ell} \left(A_m^{(k)} G_m^{(k)} B_m^{(k)} + B_m^{(k)*} G_m^{(k)} A_m^{(k)*} \right)$$

where $\{A_m^{(k)}, B_m^{(k)}\}_{k=1}^{\ell}$ has a limit $*$-distribution. Can check using standard free probability techniques that you get a limit $*$-distribution if ℓ is fixed. *Can you let ℓ grow with m?*

Next Steps

- Improve to a.s. convergence. Does not follow from these estimates in general (and probably not at all with wild enough distributions). But there should be conditions on the entries. In Guionnet's approach (with independent entries), suffices to assume the laws of the entries satisfy LSIs. How about in these correlated matrices?

Next Steps

- Improve to a.s. convergence. Does not follow from these estimates in general (and probably not at all with wild enough distributions). But there should be conditions on the entries. In Guionnet's approach (with independent entries), suffices to assume the laws of the entries satisfy LSIs. How about in these correlated matrices?
- Increase the allowed size of the partition blocks (if possible). Requires a totally different approach: the sharpest theoretical estimates for our mollified LSIs will require $d_N = O(\log N)$. Combinatorial approach?

Next Steps

- Improve to a.s. convergence. Does not follow from these estimates in general (and probably not at all with wild enough distributions). But there should be conditions on the entries. In Guionnet's approach (with independent entries), suffices to assume the laws of the entries satisfy LSIs. How about in these correlated matrices?
- Increase the allowed size of the partition blocks (if possible). Requires a totally different approach: the sharpest theoretical estimates for our mollified LSIs will require $d_N = O(\log N)$. Combinatorial approach?
- Alternative framework: use LSI to show convergence in probability even if *all* entries are correlated, but correlation strength decays quickly with distance between entries. (Work in progress.)

Next Steps

- Improve to a.s. convergence. Does not follow from these estimates in general (and probably not at all with wild enough distributions). But there should be conditions on the entries. In Guionnet's approach (with independent entries), suffices to assume the laws of the entries satisfy LSIs. How about in these correlated matrices?
- Increase the allowed size of the partition blocks (if possible). Requires a totally different approach: the sharpest theoretical estimates for our mollified LSIs will require $d_N = O(\log N)$. Combinatorial approach?
- Alternative framework: use LSI to show convergence in probability even if *all* entries are correlated, but correlation strength decays quickly with distance between entries. (Work in progress.)