Lectures on Random Matrices (Spring 2025)
Lecture 7: Double contour integral kernel. Steepest descent and local statistics

Leonid Petrov

Monday, February 24, 2025[*]

Contents

1

1   Steepest descent for the GUE kernel

1.1   Recap

We continue the asymptotic analysis of the GUE kernel.

The GUE correlation kernel is defined by

$$K_n(x,y) = \sum_{j=0}^{n-1} \psi_j(x)\psi_j(y),$$

where the functions

$$\psi_j(x) = \frac{1}{\sqrt{h_j}}\, p_j(x)\, e^{-x^2/4}$$

are built from the monic Hermite polynomials $p_j(x)$ with normalization constants h_j ensuring that the ψ_j's form an orthonormal system in $L^2(\mathbb{R})$.

Using the generating function

$$\exp\!\left(xt - \frac{t^2}{2}\right) = \sum_{n\geq 0} p_n(x)\frac{t^n}{n!},$$

one obtains by Cauchy's integral formula

$$p_n(x) = \frac{n!}{2\pi i}\oint_C \frac{\exp\!\left(xt - \frac{t^2}{2}\right)}{t^{n+1}}\, dt,$$

which leads to

$$\psi_n(x) = \frac{e^{-x^2/4}}{\sqrt{h_n}}\,\frac{n!}{2\pi i}\oint_C \frac{\exp\!\left(xt - \frac{t^2}{2}\right)}{t^{n+1}}\, dt.$$

Starting from the Fourier transform identity

$$\int_{-\infty}^{\infty} \exp\!\left(-\frac{t^2}{2} + i\,t\,x\right) dt = \sqrt{2\pi}\, e^{-x^2/2},$$

and differentiating with respect to x, then changing variables, one obtains

$$\psi_n(x) = \frac{i\,e^{x^2/4}}{\sqrt{2\pi\,h_n}} \int_{-i\infty}^{i\infty} s^n\, e^{s^2/2 - s\,x}\, ds.$$

By inserting the above representations for $\psi_n(x)$ into the kernel sum, one arrives at the double contour integral formula (after conjugation and the trick with removing $1/(s-t)$):

$$K_n(x,y) = \frac{1}{(2\pi)^2} \oint_C dt \int_{-i\infty}^{i\infty} ds\, \frac{\exp\left\{ \frac{s^2}{2} - sy - \frac{t^2}{2} + tx \right\}}{s-t} \left(\frac{s}{t}\right)^n.$$

The integration contour C is a small contour around 0, and s is passing to the right of C.

This representation is especially useful for performing asymptotic analysis (for example, via the steepest descent method) and for deriving results such as the semicircle law.

1.2   Scaling

Let us now consider the GUE kernel (**??**),

$$K_n(x,y) = \frac{1}{(2\pi)^2} \oint_C dt \int_{-i\infty}^{i\infty} ds\, \frac{\exp\left\{ \frac{s^2}{2} - sy - \frac{t^2}{2} + tx \right\}}{s-t} \left(\frac{s}{t}\right)^n.$$

We know from the Wigner semicircle law (established for real symmetric matrices with general iid entries in in Lecture 2, and for the GUE in Lecture 4) that the eigenvalues live on the scare $\sqrt{n}$. This means that to capture the local asymptotics, we need to scale

$$x = X\sqrt{n} + \frac{\Delta x}{\sqrt{n}}, \qquad y = Y\sqrt{n} + \frac{\Delta y}{\sqrt{n}}, \qquad \Delta x, \Delta y \in \mathbb{R}. \tag{1.1}$$

Moreover, if $X \neq Y$ (i.e., different global positions), one can check that the kernel vanishes. In other words, the local behaviors at different global positions are independent. See Problem **??**. In what follows, we take $Y = X$.

Let us also make a change of the integration variables:

$$t = z\sqrt{n}, \qquad s = w\sqrt{n}.$$

The integration contours for z and w look the same as in **??**, up to a rescaling. However, as 0 and $t = s$ are the only singularities in the integrand, we can deform the z, w contours as we wish, while keeping $|z| < |w|$ and the general shape as in **??**.

We thus have:

$$K_n(X\sqrt{n} + \Delta x/\sqrt{n}, X\sqrt{n} + \Delta y/\sqrt{n})$$
$$= \frac{\sqrt{n}}{(2\pi)^2} \oint_C dz \int_{-i\infty}^{i\infty} dw\, \frac{\exp\left\{ n\left(\log w - \log z + \frac{w^2}{2} - \frac{z^2}{2} + X(z-w) + \frac{z\Delta x - w\Delta y}{n} \right) \right\}}{w-z}. \tag{1.2}$$

Remark 1.1. The logarithms in the exponent are harmless, since for the estimates we only need the real parts of the logarithms, and for the main contributions, we will have $z \approx w$, so any phases of the logarithms would cancel.

The asymptotic analysis of double contour integrals like (1.2) in the context of determinantal point processes was pioneered in [Oko02, Section 3].

3

1.3 Critical points

Let us define
$$S(z) := \frac{z^2}{2} + \log z - Xz.$$

Then the exponent contains $n\,(S(w) - S(z))$. According to the steepest descent ideology, we should deform the integration contours to pass through the critical point(s) z_{cr} of $S(z)$. Moreover, the new w contour should maximize the real part of $S(z)$ at z_{cr}, and the new z contour should minimize it. If $S''(z_{cr}) \neq 0$, it is possible to locally choose such contours, they will be perpendicular to each other at z_{cr}.

Thus, we need to find the critical points of $S(z)$. They are found from the quadratic equation:

$$S'(z) = z + \frac{1}{z} - X = 0, \qquad z_{cr} = \frac{X \pm \sqrt{X^2 - 4}}{2}. \tag{1.3}$$

Depending on whether $|X| < 2$, there are three cases. Unless $|X| = 2$, equation (1.3) has a single root, and thus $S''(z_{cr}) \neq 0$. We will consider the three cases in Sections 1.4 to 1.6 below.

1.4 Imaginary critical points: $|X| < 2$, "bulk"

When $|X| < 2$, the critical points are complex conjugate. Denote them by z_{cr} and $\overline{z_{cr}}$. Since $S(z)$ has real coefficients, we have

$$\operatorname{Re} S(z_{cr}) = \operatorname{Re} S(\overline{z_{cr}}).$$

Thus, we need to consider the contribution from both points. For simplicity of the computations, let us consider only the case $X = 0$. See Problem **??**. We have

$$z_{cr} = i, \qquad S''(z_{cr}) = 2.$$

The behavior of $\operatorname{Re} S(z)$ on the complex plane can be illustrated by a 3D plot or by a region plot of the regions where $\operatorname{Re} S(z) - \operatorname{Re} S(z_{cr})$ has constant sign. See Figure 1 for an illustration in the case $X = \frac{1}{2}$. (We take $X \neq 0$ to break symmetry, for a better intuition.)

From the region plot, we see that the new z contour should pass through the shaded region $\operatorname{Re} S(z) - \operatorname{Re} S(z_{cr}) > 0$, and the new w contour should pass through the unshaded region $\operatorname{Re} S(z) - \operatorname{Re} S(z_{cr}) < 0$.

Deforming the contours from **??** to the new contours is impossible without passing through the residue at $w = z$. Moreover, this residue appears only for certain values of z. Namely, for $X = 0$, let us first make the z contour to be the positively (counterclockwise) oriented unit circle. It passes through the critical points $z_{cr} = i$ and $\overline{z_{cr}} = -i$. Since the original w contour is to the right of the z contour, we only encounter the residue when z is in the right half of the circle.

Thus, we can write

$$\oiint_{\text{old contours}} = \oiint_{\text{new contours}} + \int_{-i}^{i} 2\pi i \, \operatorname{Res}_{w=z} \, dz, \tag{1.4}$$

where in the single integral, the z contour passes to the right of the origin, along the right half of the unit circle.
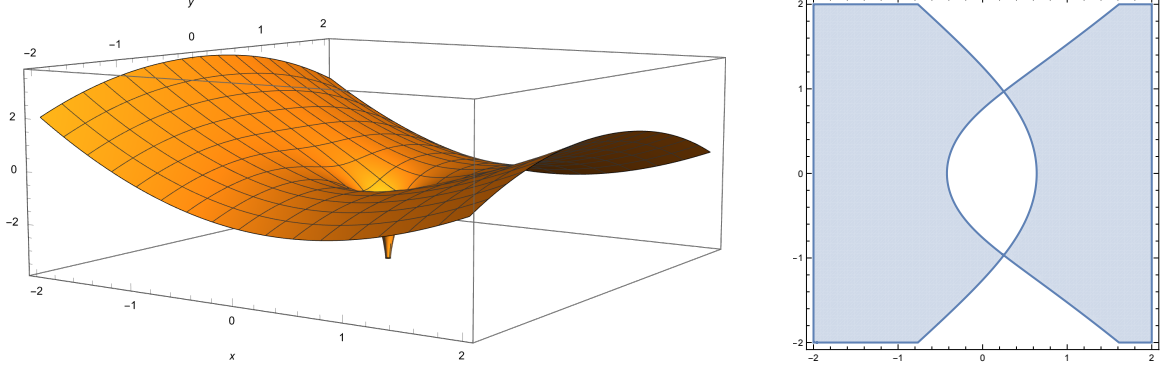
Figure 1: A 3D plot and a region plot of the regions where $\operatorname{Re} S(z) - \operatorname{Re} S(z_{cr})$ is positive (highlighted) or negative, in the case $X = \frac{1}{2}$. In this case, $z_{cr} \approx 0.25 + 0.96i$.

It remains to consider the two integrals in the right-hand side of (1.4). Recall that the correlation functions are defined relative to a reference measure, and the right object to scale is

$$K_n(x,y)dy = \frac{1}{\sqrt{n}} d(\Delta y).$$

The extra factor $n^{-1/2}$ compensates the prefactor $\sqrt{n}$ in (1.2).

The single integral takes the form

$$\frac{-i}{2\pi} \int_{-i}^{i} e^{z(\Delta x - \Delta y)} dz = \frac{\sin(\Delta x - \Delta y)}{\pi(\Delta x - \Delta y)}, \qquad \Delta x, \Delta y \in \mathbb{R}. \tag{1.5}$$

Definition 1.2. The *sine kernel* is defined as

$$K_{\text{sine}}(x,y) := \begin{cases} \dfrac{\sin(x-y)}{\pi(x-y)}, & x \neq 0, \\ \dfrac{1}{\pi}, & x = 0. \end{cases}$$

(The value at $x = y$ is defined by continuity.)

This kernel is translation invariant, and is often defined with a single argument, as $K_{\text{sine}}(x-y)$.

The double integral has both contours in the "steepest descent" regime, which means that the main contribution is

$$\text{const} \cdot \frac{e^{n(\operatorname{Re} S(z_{cr}) - \operatorname{Re} S(z_{cr}))}}{\sqrt{n}} \sim \frac{\text{const}}{\sqrt{n}}.$$

At this rate, the double integral over the new contours *does not* contribute to the asymptotics of the correlation functions. Recall that the correlation functions are expressed as finite-dimensional determinants of the kernel $K_n(x,y)$, and the error $O(n^{-1/2})$ is negligible in the limit $n \to +\infty$. This is because the main term comes from the single integral, which does not vanish.

We have established the following result:

5

Proposition 1.3 (Bulk asymptotics at $X = 0$). *The correlation kernel K_n of the GUE has the following asymptotics close to zero as $n \to +\infty$:*

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} K_n \left(\frac{\Delta x}{\sqrt{n}}, \frac{\Delta y}{\sqrt{n}} \right) = K_{\text{sine}}(\Delta x, \Delta y), \qquad \Delta x, \Delta y \in \mathbb{R}.$$

Consequently, the eigenvalues of the GUE converge to the sine process determined by the sine kernel (Definition 1.2), in the sense of finite-dimensional distributions.

Remark 1.4. Beyond $X = 0$, the local correlations are essentially the same, up to rescaling of the real line by a constant factor (depending on the semicircle density). See Problem **??**.

1.5 Real critical points: $|X| > 2$, "large deviations"

For $X^2 > 4$, both solutions (1.3) are real. Let us assume $X > 2$, the case $X < 2$ is similar. For $X > 2$, both solutions are positive. Label these solutions as

$$z_+ = \frac{X + \sqrt{X^2 - 4}}{2}, \qquad z_- = \frac{X - \sqrt{X^2 - 4}}{2}, \quad \text{so that} \quad z_+ z_- = 1.$$

A straightforward check reveals that $z_+ > 1$ and $z_- < 1$ (for $X > 2$). Note that $S''(z) = 1 - z^{-2}$, which is positive for $z_+ > 1$ and negative for $z_- < 1$. Thus, the critical points z_+ and z_- are a local minimum and a local maximum. A crucial observation is that

$$S(z_+) < S(z_-).$$

One can deform the z integration contour to pass through z_- and the w contour to pass through z_+. Then, on these contours, one can show that

$$\text{Re}\, S(w) - \text{Re}\, S(z) < 0.$$

According to the steepest descent ideology, we see that the main exponential behavior of the double contour integral is

$$\exp\left\{ n \left(\text{Re}\, S(z_+) - \text{Re}\, S(z_-) \right) \right\} = O(e^{-\delta(X)n}), \qquad |X| > 2. \tag{1.6}$$

Here $\delta(X) > 0$ for $|X| > 2$, and $\delta(X) \to 0$ when $|X| \to 2$.

The outcome (1.6) reflects the fact that the Wigner semicircle law places all eigenvalues inside the interval $|X| \leq 2$. The probability to see even a single eigenvalue outside $[-2, 2]$ is exponentially small.

This exponential decay corresponds to a large deviation regime. Indeed, if at least one of the diagonal entries of the matrix is unusually large, this corresponds to the maximal eigenvalue to get outside the interval $[-2, 2]$. See also Problem **??**.

1.6 Double critical point: $|X| = 2$, "edge"

Throughout the subsection, we assume that $X = 2$. The case $X = -2$ is symmetric.

When $X = 2$, the two solutions in (1.3) merge into a double critical point $z_{cr} = 1$. We have

$$S'(1) = 0, \qquad S''(1) = 0, \qquad S'''(1) = 2.$$

6

Thus, the usual quadratic approximation fails and one must expand to third order. Writing

$$z = 1 + u, \qquad w = 1 + v,$$

with u, v small, we have

$$S(1+u) = S(1) + \frac{S'''(1)}{6} u^3 + O(u^4) = S(1) + \frac{u^3}{3} + O(u^4),$$

and similarly for $S(1+v)$. Hence, the difference in the exponents becomes

$$S(1+v) - S(1+u) = \frac{v^3 - u^3}{3} + O(u^4 + v^4).$$

To capture the correct asymptotics, we rescale the local variables by setting

$$u = \frac{U}{n^{1/3}}, \qquad v = \frac{V}{n^{1/3}},$$

so that

$$n\Big[S(1+v) - S(1+u)\Big] = \frac{V^3 - U^3}{3} + O\Big(n^{-1/3}\Big).$$

Moreover, the correct edge scaling for the spatial variables is obtained by writing

$$x = 2\sqrt{n} + \frac{\xi}{n^{1/6}}, \qquad y = 2\sqrt{n} + \frac{\eta}{n^{1/6}}, \qquad \xi, \eta \in \mathbb{R}.$$

We have

$$n\left(S(w) - S(z)\right) = n^{1/3}(\xi - \eta) + \frac{V^3 - U^3}{3} + \xi U - \eta V + O\Big(n^{-1/3}\Big).$$

The terms $n^{1/3}(\xi - \eta)$ are harmless as they can be removed by conjugation.

The region plot of $\operatorname{Re} S(z) - \operatorname{Re} S(1)$ (shown in Figure 2) makes sure that we can deform the z contour so that it passes through $z_{cr} = 1$ as the new U contour at the angles $\pm\frac{2\pi}{3}$ (where $\operatorname{Re} U^3 > 0$), we can deform the w contour so that it passes through $z_{cr} = 1$ as the new V contour at the angles $\pm\frac{\pi}{3}$ (where $\operatorname{Re} V^3 < 0$). This will ensure the convergence of the new double integral.
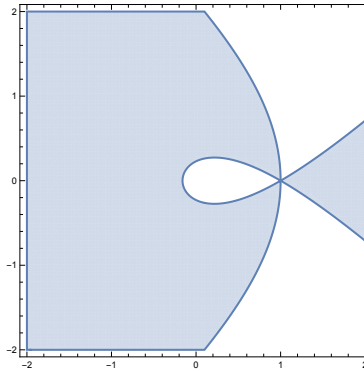

Figure 2: The plot of the region $\operatorname{Re} S(z) - \operatorname{Re} S(1) > 0$ for $X = 2$.

Thus, we have shown that under the rescaling, the GUE correlation kernel $K_n(x, y)\,dy$ converges to a new kernel.

Definition 1.5. Define the *Airy kernel* on $\mathbb{R}$ by

$$K_{\mathrm{Ai}}(\xi, \eta) = \frac{1}{(2\pi i)^2} \int_{e^{-\frac{\pi i}{3}}\infty}^{e^{\frac{\pi i}{3}}\infty} dV \int_{e^{-\frac{2\pi i}{3}}\infty}^{e^{2\frac{\pi i}{3}}\infty} dU \, \frac{\exp\left\{\frac{V^3 - U^3}{3} + U\,\xi - V\,\eta\right\}}{V - U}.$$

For another formula for the Airy kernel which does not involve integrals, see Problem G.1.

Proposition 1.6. *We have*

$$\lim_{n\to\infty} \frac{1}{n^{1/6}} K_n\left(2\sqrt{n} + \frac{\xi}{n^{1/6}}, \, 2\sqrt{n} + \frac{\eta}{n^{1/6}}\right) \to K_{\mathrm{Ai}}(\xi, \eta).$$

Consequently, the eigenvalue statistics at the edge of the spectrum converge to the Airy point process, in the sense of fine-dimensional distributions.

1.7 Airy kernel, Tracy–Widom distribution, and convergence of the maximal eigenvalue

Let us make a few remarks on the asymptotic results of Propositions 1.3 and 1.6. First, a rigorous justification of convergence of contour integrals requires some estimates on the error terms in the steepest descent analysis, but these estimates are mild and not hard to obtain.

Second, the GUE has the maximal eigenvalue λ_{max}. It is reasonable to assume that the Airy process also (almost surely) admits a maximal point (usually denoted by $\mathfrak{a}_1$), and that λ_{max} converges to $\mathfrak{a}_1$ under appropriate rescaling:

$$\lim_{n\to\infty} n^{\frac{1}{6}}\left(\lambda_{max} - 2\sqrt{n}\right) = \mathfrak{a}_1. \tag{1.7}$$

This is indeed the case, but to show (1.7), one needs to show the convergence in distribution:

$$\lim_{n\to\infty} \mathbb{P}\left(n^{1/6}(\lambda_{max} - 2\sqrt{n}) \leq x\right) \to \mathbb{P}(\mathfrak{a}_1 \leq x). \tag{1.8}$$

Both events (1.8) are so-called *gap probabilities*, for example,

$$\mathrm{P}(\mathfrak{a}_1 \leq x) = \mathrm{P}(\text{there are no eigenvalues in the interval } (x, \infty)),$$

which is expressed as the Fredholm determinant

$$\det(1 - K_{\mathrm{Ai}})_{(x,\infty)} = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \int_x^{\infty} dy_1 \int_x^{\infty} dy_2 \cdots \int_x^{\infty} dy_m \det_{i,j=1}^{m} K_{\mathrm{Ai}}(y_i, y_j). \tag{1.9}$$

Thus, to get (1.8)), one needs to show the convergence of sums like this for the GUE kernel to the corresponding sums for the Airy kernel. This is doable, but tedious.

Moreover, to get convergence in distribution of random variables, one would also have to argue either *tightness*, or independently show that (1.9) defines a cumulative probability distribution function in x:

$$F_2(x) = \det(1 - K_{\mathrm{Ai}})_{(x,\infty)}. \tag{1.10}$$

The distribution (1.10) is known as the *GUE Tracy–Widom distribution*. The subscript 2 indicates that $\beta = 2$. There are distributions F_β for all beta, most notably, the GOE and GSE distributions. The classical distributions F_1, F_2, F_4 also appear as fluctuation distributions in interacting particle systems, while other beta values do not quite appear in the particle systems domain.

More details may be found in the original papers [TW93], [For93], [TW94].

2   Introduction and Motivation

In random matrix theory, one often studies the entire spectrum of an $n \times n$ matrix ensemble such as the Gaussian Unitary Ensemble (GUE), the Gaussian Orthogonal Ensemble (GOE), or, more generally, β-ensembles. However, it is also natural to examine the spectra of *principal minors* of such matrices.

When we say "cutting corners," we typically refer to extracting a top-left $k \times k$ submatrix (or *corner*) out of an $n \times n$ random matrix H and then looking at the interplay among the eigenvalues of all corners $k = 1, \ldots, n$. This forms a *nested* family of spectra, often described by interlacing (or Gelfand–Tsetlin) patterns.

The *GUE corners process* is a classical example of this phenomenon. Concretely, if H is an $n \times n$ GUE matrix, then the top-left $k \times k$ corners (for $1 \leq k \leq n$) have jointly distributed eigenvalues that exhibit remarkable determinantal structures, interlacing inequalities, and limit theorems. Similar statements hold for the GOE, the Gaussian Symplectic Ensemble (GSE), and more general β-ensembles (algebraic generalizations of GUE/GOE/GSE that we also discuss).

2.1   Outline

These notes proceed as follows:

- §3 **Preliminaries.** We recall the GUE definition, its diagonalization, and the general β-ensembles.

- §4 **Corners of Random Matrices.** We define the corner (minor) processes and recall the fundamental interlacing property.

- §5 **GUE Corners: Joint Distribution and Determinantal Structure.** We outline how to compute the joint distribution of the spectra of all corners, show the interlacing, and discuss the determinantal kernel.

- §6 **General β Corners.** We show how the GUE corners result has a natural extension to the tridiagonal β-ensembles (Dumitriu–Edelman) and mention connections to Wishart/Laguerre and Jacobi corners.

- §7 **Local Limits.** We review the bulk (sine) and edge (Airy) universality in each corner and highlight how the entire triangular array has consistent local limits.

- §8 **Connections and Applications.** We discuss ties to Gelfand–Tsetlin patterns, representation theory, partial Haar unitaries, and beyond.

- §9 **Exercises.** We present problem sets illustrating these concepts.

3 Preliminaries on Gaussian and β-Ensembles

3.1 GUE Definition and Basic Facts

The Gaussian Unitary Ensemble (GUE_n) is the probability distribution on $n \times n$ Hermitian matrices whose density is proportional to

$$\exp\left(-\tfrac{1}{2}\operatorname{Tr}(H^2)\right) dH,$$

where dH denotes the Lebesgue measure on the space of Hermitian $n \times n$ matrices. Equivalently, one can specify that the entries H_{ij} for $i < j$ are i.i.d. complex Gaussians with mean zero and variance $1/2$, and the diagonal entries H_{ii} are i.i.d. real Gaussians with mean zero and variance 1.

A fundamental property is that the joint distribution of eigenvalues $(\lambda_1, \dots, \lambda_n)$ (ordered in any way, typically $\lambda_1 \geq \cdots \geq \lambda_n$) is given by the well-known *Hermite (or GUE) $\beta = 2$-ensemble* formula:

$$p(\lambda_1, \dots, \lambda_n) \;=\; \frac{1}{Z_n} \prod_{i<j} (\lambda_i - \lambda_j)^2 \exp\left(-\tfrac{1}{2}\sum_{k=1}^{n} \lambda_k^2\right). \tag{3.1}$$

Here Z_n is the normalizing constant. The $\beta = 2$ in the exponent of the Vandermonde product $\prod_{i<j}(\lambda_i - \lambda_j)^\beta$ reflects the unitary symmetry class.

3.2 General β-Ensembles

More generally, one can define a one-parameter family of ensembles indexed by $\beta > 0$, called *β-ensembles*:

$$p_\beta(\lambda_1, \dots, \lambda_n) \;=\; \frac{1}{Z_{n,\beta}} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \prod_{k=1}^{n} e^{-V(\lambda_k)}, \tag{3.2}$$

where $V(x)$ is a confining potential, often taken as $V(x) = \frac{x^2}{2}$ (Gaussian case) or $V(x)$ suitable for other classical ensembles (e.g., Laguerre/Wishart, Jacobi, etc.). For $\beta = 1, 2, 4$ these correspond to the classical GOE, GUE, GSE, respectively, but β need not be an integer or even rational.

An important way to realize the β-ensembles (with Gaussian potential) is via the *Dumitriu–Edelman* tridiagonal representation: one constructs an $n \times n$ tridiagonal matrix T_β whose diagonal entries are i.i.d. Gaussians (with certain means and variances) and whose sub- and super-diagonal entries are independent χ-distributed random variables. For $\beta = 2$, this recovers the GUE tridiagonal matrix. All of these β-ensembles share the fundamental property that their eigenvalues form a *repulsive point process* governed by (3.2).

4 Corners of Hermitian Matrices: Definition and Interlacing

4.1 Principal Corners (Minors)

Let H be an $n \times n$ Hermitian matrix. For each $1 \leq k \leq n$, define the *top-left $k \times k$ corner* $H^{(k)}$ by

$$H^{(k)} \;=\; \big[H_{ij}\big]_{1 \leq i,j \leq k}.$$

Since H is Hermitian, each $H^{(k)}$ is also Hermitian. Let

$$\lambda_1^{(k)} \;\geq\; \lambda_2^{(k)} \;\geq\; \cdots \;\geq\; \lambda_k^{(k)}$$

denote the eigenvalues of $H^{(k)}$. Then the collection

$$\left\{ \lambda_j^{(k)} : 1 \leq j \leq k \leq n \right\}$$

is called the *corners spectrum* (or *minor spectrum*) of H. When H is random, this entire triangular array of eigenvalues becomes a random point configuration in the two-dimensional set $\{1, \ldots, n\} \times \mathbb{R}$.

4.2   Interlacing Property

A fundamental feature of Hermitian matrices is that the eigenvalues of corners interlace with the eigenvalues of the full matrix. More precisely, if $\nu_1 \geq \cdots \geq \nu_n$ are the eigenvalues of H itself (i.e., the full $n \times n$ matrix), and $\mu_1 \geq \cdots \geq \mu_k$ are the eigenvalues of $H^{(k)}$, then we have:

$$\nu_1 \;\geq\; \mu_1 \;\geq\; \nu_2 \;\geq\; \mu_2 \;\geq\; \ldots \;\geq\; \nu_k \;\geq\; \mu_k \;\geq\; \nu_{k+1}.$$

In particular,
$$\lambda_1^{(k+1)} \;\leq\; \lambda_1^{(k)} \;\leq\; \lambda_2^{(k+1)} \;\leq\; \cdots \;\leq\; \lambda_k^{(k)} \;\leq\; \lambda_{k+1}^{(k+1)}.$$

Graphically, one can depict $\{\lambda_j^{(k)}\}$ in a triangular Gelfand–Tsetlin pattern form, reflecting these interlacing inequalities.

Remark 4.1 (Schur Complement Interpretation)**.** The interlacing property can be seen via Schur complements: when passing from H to its $(n-1) \times (n-1)$ corner, one effectively removes the last row and column, so the rank-one update in the Schur complement triggers the Weilandt–Hoffman/Cauchy interlacing inequalities.

5   GUE Corners: Joint Distribution and Determinantal Structure

Consider now the *joint* distribution of all corners of a GUE_n matrix H. That is, we have the random matrices
$$H^{(1)}, \; H^{(2)}, \; \ldots, \; H^{(n)} = H,$$

and want to understand the collection $\{\lambda_j^{(k)}\}$ for $1 \leq j \leq k \leq n$ as a single random point process.

5.1   Spectral Decomposition and Haar Unitary

Recall that H can be diagonalized:

$$H = U \Lambda U^\dagger, \quad \Lambda = diag(\lambda_1, \ldots, \lambda_n),$$

where Λ is the real diagonal matrix of H's eigenvalues (in descending order) and U is Haar-distributed on the unitary group $\mathrm{U}(n)$. The top-left $k \times k$ corner $H^{(k)}$ can be written in terms

of sub-blocks of U and Λ. In principle, one then integrates over the Haar measure to derive the joint law of $(H^{(1)}, \ldots, H^{(n)})$.

While the resulting distribution is complicated, it is nevertheless highly structured and, in fact, forms a *determinantal point process* (DPP) in the two-dimensional space of "row index k" and "spectral variable x."

5.2 Determinantal Form: GUE Corners Process

The formal statement (see, e.g., [?Johansson-2005, ?Johansson-2006, ?baryshnikov2001gues, ?forrester2010log] for references) is:

Theorem 5.1 (GUE Corners as a 2D Determinantal Process). *Let H be an $n \times n$ GUE matrix and let $\{\lambda_j^{(k)}\}_{1 \le j \le k \le n}$ be the eigenvalues of its top-left corners of sizes $k = 1, \ldots, n$. Then, viewed as a random point set in $\{1, \ldots, n\} \times \mathbb{R}$, this collection is a* determinantal point process*:*

$$\mathbb{P}\big[(k_1, x_1), \ldots, (k_m, x_m) \in \text{the process}\big] = \det\Big[K\big((k_i, x_i), (k_j, x_j)\big)\Big]_{i,j=1}^{m},$$

where K is the extended correlation kernel. *In particular, correlation functions for the entire triangular array are given by minors of K.*

Explicit formulas for $K\big((k, x), (k', y)\big)$ exist, but are somewhat more involved than the single-size GUE kernel. Nevertheless, one can still identify them in terms of *orthogonal polynomials* (Hermite polynomials) and certain additional matrix integrals.

Remark 5.2. For $k = n$ (the largest corner), we recover the usual 1D GUE correlation kernel restricted to the $\lambda_i^{(n)}$ alone. The extended 2D kernel encapsulates how these GUE eigenvalues relate to the smaller corners.

5.3 Gelfand–Tsetlin Patterns and Markov Structure

An important combinatorial viewpoint: if we only keep track of the eigenvalues (without any concern for eigenvectors), the random array $\{\lambda_j^{(k)}\}$ forms a random Gelfand–Tsetlin pattern with continuous entries. One can show that as k increases, $(\lambda_1^{(k)}, \ldots, \lambda_k^{(k)})$ is a *Markov chain* in k:

$$(\lambda_1^{(1)}) \longrightarrow (\lambda_1^{(2)}, \lambda_2^{(2)}) \longrightarrow \cdots \longrightarrow (\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}).$$

The transition density from (k)-corner eigenvalues to $(k+1)$-corner eigenvalues encodes the interlacing constraints and the GUE invariance. Determinantal structure yields closed-form transition kernels.

6 General β Corners Processes

The GUE case ($\beta = 2$) is the richest in integrable (determinantal) structures, but corners processes exist for all β as well. Specifically, if one considers the β-ensemble in tridiagonal form (the Dumitriu–Edelman approach), then the top-left corners of this tridiagonal matrix yield an entire

nested sequence of β-ensembles for smaller dimensions, though with certain correlated modifications. The full joint distribution of all these corners forms a random triangular array with similar *interlacing* constraints. The structure is no longer purely determinantal for general β, but it can often be described via *multivariate Bessel functions*, *Selberg integrals*, or other integrable-type objects depending on β.

For example:

- In the *Gaussian Orthogonal Ensemble* ($\beta = 1$), the corners process has a Pfaffian structure (due to real symmetry and real eigenvectors).

- In the *Gaussian Symplectic Ensemble* ($\beta = 4$), a related Pfaffian structure appears (with symplectic symmetry).

- For general β, corners processes can often be described by hypergeometric functions of matrix arguments, or can be seen as special cases of the so-called *multivariate hypergeometric orthogonal polynomial ensembles*.

Thus, while $\beta = 2$ remains the simplest and most explicit (due to unitarity and determinantal formulas), the phenomenon of "cutting corners" to get a nested set of minors is pervasive across all β.

6.1   Wishart/Laguerre and Jacobi Corners

Similar statements hold for Wishart (Laguerre) ensembles or Jacobi (MANOVA) ensembles. One can look at partial corners, say the top-left corner of a rectangular Gaussian matrix X, or the principal corners of $X^\dagger X$ (Wishart), or the corners of a random unitary sub-block (Jacobi). The spectra and their interlacing relationships again produce a random triangular array with a structured correlation law. These corner processes are widely studied in multivariate statistics and in representation-theoretic random measures.

7   Local Limits: Bulk and Edge of Each Corner

One might ask how the local eigenvalue statistics for smaller corners compare to those in the full matrix. Indeed, each corner $H^{(k)}$ is a $k \times k$ Hermitian matrix, so in the limit $n \to \infty$ (and possibly $k \to \infty$ in tandem with n), we can look at:

$$\lambda_{\max}^{(k)}, \quad \text{gap statistics in the interior of the spectrum of } H^{(k)}, \ldots$$

An interesting scenario is when k is proportional to n, i.e. $k = \alpha n$ for some $0 < \alpha \leq 1$. For the GUE, one can use known results about *rank-one updates* or the fact that $H^{(k)}$ is close (in a certain sense) to a smaller GUE plus correlated terms. The main takeaway is that:

- The *global* empirical distribution of $H^{(k)}$ converges to the Wigner semicircle (or appropriate portion of it) if $k \to \infty$. In fact, as $k, n \to \infty$ with $k/n \to \alpha$, the top-left corners have a limiting spectral distribution that is the same as the GUE scaled by $\sqrt{n}$, up to small boundary effects.

- The *local* statistics in the bulk remain universal, giving the *sine kernel* limit. Near the edge, we get *Airy* behavior. These corners do not break the usual universality phenomena: local fluctuations around scaled spectral points still follow the same universal kernels.

- There are also interesting *transitional* regimes if k is close to n, or if k is fixed while $n \to \infty$. In the latter case, $H^{(k)}$ does not grow in size, so the distribution of the $k \times k$ corner can converge to that of a simpler random matrix ensemble with additional constraints.

Hence, one sees a consistent story: the entire triangular array $\{\lambda_j^{(k)}\}$ has local limits that are consistent with the well-known universal kernels in random matrix theory.

8 Connections and Applications

8.1 Gelfand–Tsetlin Patterns in Representation Theory

The corner spectra of a GUE matrix can be viewed as generating a random Gelfand–Tsetlin pattern in continuous variables:

$$
\begin{array}{ccccc}
\lambda_1^{(n)} & & & & \\
\lambda_1^{(n-1)} & \lambda_2^{(n-1)} & \ldots & \lambda_{n-1}^{(n-1)} & \\
\vdots & & \ddots & \vdots & \\
\lambda_1^{(1)} & & & &
\end{array}
$$

with $\lambda_j^{(k)} \geq \lambda_{j+1}^{(k+1)} \geq \cdots$. This is directly analogous to the discrete Gelfand–Tsetlin patterns that parametrize irreducible representations of U(n) (or SU(n)). The random matrix approach suggests that these continuous patterns are natural objects carrying determinantal/Pfaffian structures, leading to connections with *asymptotic representation theory* and *integrable probability.*

8.2 Partial Haar Unitaries

If $H = U\Lambda U^\dagger$ with U Haar-distributed on U(n), then the sub-blocks of U (e.g., the top-left $k \times n$ portion) inherit special rotational invariance properties known as *partial Haar unitaries* or *isometries* from the group measure. One can interpret the corners $H^{(k)}$ in terms of these partial unitaries. This viewpoint is used in quantum information (for random states and channels) and in multivariate statistics (for random orthonormal bases).

8.3 Integrable Systems and Discrete Analogs

Finally, corners processes appear in integrable models of lattice systems and random partitions. For instance, certain *plane partitions* or *Young tableaux* ensembles have limiting shapes described by the GUE-corners distribution in scaled coordinates. The broad principle is that any strongly *interlacing* or *Gelfand–Tsetlin* structure with underlying determinantal or Pfaffian formula often is governed by the same universal corners processes seen in random matrix theory.

9 Problems and Exercises

1. **Schur Complement and Interlacing.**
   Given a Hermitian matrix A of size $n \times n$, show that its $(n-1) \times (n-1)$ top-left corner $A^{(n-1)}$ is the Schur complement obtained by removing the last row/column. Use this viewpoint to deduce the interlacing property between the eigenvalues of $A^{(n-1)}$ and A.

2. **Determinantal / Pfaffian Structures for $\beta = 1, 2, 4$.**
   Explain why for $\beta = 1, 4$ (the GOE and GSE), one gets *Pfaffian* structures rather than purely determinantal ones. Sketch how the presence of real symmetry ($\beta = 1$) or symplectic symmetry ($\beta = 4$) modifies the joint law of eigenvalues.

3. **GUE Corners for $n = 2$ and $n = 3$.**
   Explicitly write out (symbolically, or with a small calculation) the joint distribution of $\{\lambda_j^{(k)}\}$ for $k = 1, 2$ (when $n = 2$), and similarly for $n = 3$. Identify how the interlacing $\lambda_1^{(1)} \geq \lambda_1^{(2)} \geq \lambda_2^{(2)}$ appears. Check if you can see any determinant form for correlation functions in these small cases.

4. **Tridiagonal Realization of Corners ($\beta = 2$).**
   Construct a tridiagonal GUE matrix T of size n, then look at the principal $(k \times k)$ top-left submatrix $T^{(k)}$. Compare the distribution of $T^{(k)}$ with that of a smaller GUE(k) matrix. Are they the same or different? If different, precisely how do they differ?

5. **Wishart / Laguerre Corners.**
   Consider the Wishart/Laguerre ensemble $W = X^\dagger X$, where X is an $m \times n$ complex Gaussian matrix. Define $W^{(k)}$ as the top-left $k \times k$ corner. Write out the joint distribution of eigenvalues of $W^{(1)}, \ldots, W^{(n)}$ (assuming $m \geq n$). Describe the interlacing properties and how they relate to the GUE corners for a suitable transformation of W.

6. **Local Limit for a Fixed-Size Corner.**
   For a large $n \times n$ GUE, consider only the top-left $k \times k$ corner for some *fixed* k. Show that in the $n \to \infty$ limit, this corner *converges in distribution* to a simpler random matrix (explain or guess its form). Does this limit matrix have i.i.d. entries? Discuss the effect of the rank-1 update from the rest of the matrix.

7. **Markov Property in the Triangular Array.**
   Prove (or outline why) the sequence of eigenvalue vectors $(\lambda_1^{(k)}, \ldots, \lambda_k^{(k)})$ is a Markov chain in k, for the GUE corners process. Determine the transition kernel in the finite n case or give a reference for its explicit form.

G   Problems (due 2025-03-25)

G.1   Airy kernel

Define the Airy function by

$$Ai(\xi) := \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iU^3/3 + i\xi U} dU = \frac{1}{\pi} \int_0^{\infty} \cos\left(\frac{U^3}{3} + \xi U\right) dU.$$

This integral converges, but only conditionally. To improve convergence, one should instead integrate along a complex contour, from $e^{\frac{5\pi i}{6}}\infty$ to 0 to $e^{\frac{\pi i}{6}}\infty$.

Show that

$$K_{\mathrm{Ai}}(\xi, \eta) = \frac{Ai(\xi)\, Ai'(\eta) - Ai(\eta)\, Ai'(\xi)}{\xi - \eta}.$$

Note that this expression is parallel to the sine kernel,

$$\frac{\sin(x-y)}{\pi(x-y)} = \frac{\sin x \cos y - \cos x \sin y}{\pi(x-y)}, \qquad \cos x = (\sin x)'.$$

These correlation kernels are called *integrable* [IIKS90].

Hint for the problem: observe that

$$\exp\left\{-izx + iwy\right\} = \frac{i}{x - y}\left(\frac{\partial}{\partial z} + \frac{\partial}{\partial w}\right) \exp\left\{-izx + iwy\right\},$$

and use integration by parts in $K_{\mathrm{Ai}}(\xi, \eta)$ from Definition 1.5.

References

[For93] P. J. Forrester, *The spectrum edge of random matrix ensembles*, Nuclear Physics B **402** (1993), 709–728. ↑8

[IIKS90] A.R. Its, A.G. Izergin, V.E. Korepin, and N.A. Slavnov, *Differential equations for quantum correlation functions*, Int. J. Mod. Phys. B **4** (1990), no. 5, 1003–1037. ↑16

[Oko02] A. Okounkov, *Symmetric functions and random partitions*, Symmetric functions 2001: Surveys of developments and perspectives, 2002. arXiv:math/0309074 [math.CO]. ↑3

[TW93] C. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Physics Letters B **305** (1993), no. 1, 115–118. arXiv:hep-th/9210074. ↑8

[TW94] C. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Commun. Math. Phys. **159** (1994), no. 1, 151–174. arXiv:hep-th/9211141. ↑8

L. PETROV, UNIVERSITY OF VIRGINIA, DEPARTMENT OF MATHEMATICS, 141 CABELL DRIVE, KERCHOF HALL, P.O. BOX 400137, CHARLOTTESVILLE, VA 22904, USA
E-mail: lenia.petrov@gmail.com