Sums of GUE matrices and concentration of hives from correlation decay of eigengaps

Hariharan Narayanan (TIFR) Scott Sheffield (MIT)
Terence Tao (UCLA)

Let us define a relation

$$\lambda \boxplus \mu \to \nu \tag{1}$$

if there exist Hermitian matrices A, B with eigenvalues λ, μ respectively such that $A + B$ has eigenvalues ν.
Weyl asked the question of determining necessary and sufficient conditions on $\lambda, \mu, \nu \in \texttt{Spec}$ for the relation (1) to hold.

Introduction

As conjectured by Horn and proved by Klyachko, Knutson and Tao, the set

$$\mathtt{HORN}_{\lambda \boxplus \mu} := \{\nu \in \mathtt{Spec} : \lambda \boxplus \mu \to \nu\}$$

of possible ν arising from a given choice of λ, μ forms a polytope (known as the *Horn polytope*), given by the trace condition

$$\sum \lambda + \sum \mu = \sum \nu$$

together with a recursively defined set of linear inequalities known as the *Horn inequalities*.

Introduction

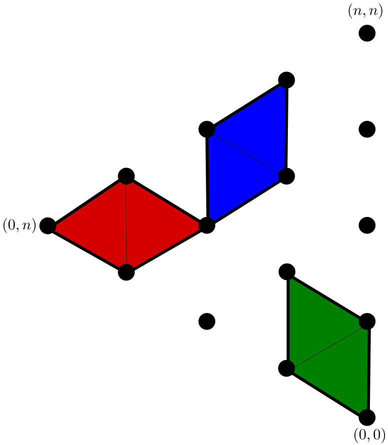


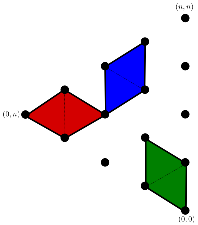
Figure: The triangular region T with $n = 4$, tilted to lie on the equilateral lattice (so that a rhombus is precisely the union of two adjacent unit equilateral triangles).

One of the key tools used in the proof of the Horn conjecture is that of a *hive*.
A *rhombus* is a quadruple $ABCD$ in the lattice $\mathbb{Z}^2$ of one of the following three forms for some $i, j \in \mathbb{Z}^2$:

(i) Blue
$(A, B, C, D) = ((i, j), (i+1, j), (i+2, j+1), (i+1, j+1))$

(ii) Green
$(A, B, C, D) = ((i, j), (i+1, j+1), (i+1, j+2), (i, j+1))$

(iii) Red $(A, B, C, D) = ((i, j), (i, j-1), (i+1, j-1), (i+1, j))$.

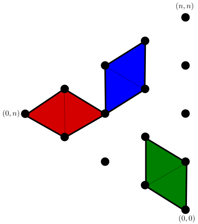


We refer to AC as the *long diagonal* of the rhombus and BD as the *short diagonal*. A function $h\colon \Omega \to \mathbb{R}$ defined on some subset Ω of $\mathbb{Z}^2$ is said to be *rhombus-concave* if one has

$$h(A) + h(C) \leq h(B) + h(D) \tag{2}$$

for all rhombi $ABCD$ in Ω. A *hive* is a rhombus concave function $h\colon T \to \mathbb{R}$ defined on the triangle

$$T := \{(i,j) \in \mathbb{Z}^2 : 0 \leq i \leq j \leq n\}. \tag{3}$$



Introduction

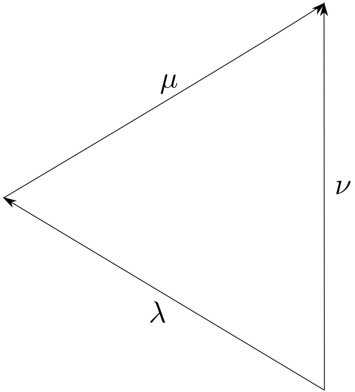


Figure: A schematic depiction of the boundary condition $\lambda \boxplus \mu \to \nu$. Thus, the hive increases according to the tuple λ as one moves from the southern vertex $(0, 0)$ to the western one $(0, n)$, according to the tuple μ as one moves from the western vertex $(0, n)$ to the northern vertex (n, n), and according to the tuple ν as one moves from the southern vertex $(0, 0)$ to the northern vertex (n, n).

Introduction

If $\lambda, \mu, \nu \in \texttt{Spec}$, we say that a hive h has boundary condition $\lambda \boxplus \mu \to \nu$ if one has

$$h(0,i) = \sum_{j=1}^{i} \lambda_j \tag{4}$$

$$h(i,n) = \sum \lambda + \sum_{j=1}^{i} \mu_j \tag{5}$$

$$h(i,i) = \sum_{j=1}^{i} \nu_j \tag{6}$$

for all $0 \leq i \leq n$, and write $\texttt{HIVE}_{\lambda \boxplus \mu \to \nu}$ for the set of all hives with boundary condition $\lambda \boxplus \mu \to \nu$.

Introduction

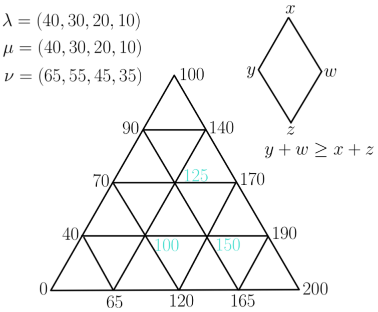


Figure: A hive with boundary condition $(40, 30, 20, 10) \boxplus (40, 30, 20, 10) \to (65, 55, 45, 35)$.

We also adopt the "wildcard convention" that replacing a tuple such as λ, μ, or ν with an asterisk $*$ denotes the operation of taking unions over all values of that tuple, thus for instance

$$\mathtt{HIVE}_{\lambda \boxplus \mu \to *} := \bigcup_{\nu} \mathtt{HIVE}_{\lambda \boxplus \mu \to \nu}$$
$$\mathtt{HIVE}_{\lambda \boxplus * \to \nu} := \bigcup_{\mu} \mathtt{HIVE}_{\lambda \boxplus \mu \to \nu}$$
$$\mathtt{HIVE}_{* \boxplus * \to *} := \bigcup_{\lambda,\mu,\nu} \mathtt{HIVE}_{\lambda \boxplus \mu \to \nu}$$

denote the hives with boundary conditions $\lambda \boxplus \mu \to *$, $\lambda \boxplus * \to \nu$,and $* \boxplus * \to *$ respectively. Note that all of these sets are convex.

Introduction

There is a natural probability measure on the Horn polytope $\mathrm{HORN}_{\lambda \boxplus \mu}$, referred to as the *Horn probability measure*, defined as the eigenvalues of $A + B$ when A, B are chosen independently and uniformly from the space of all Hermitian matrices with eigenvalues λ, μ respectively. This Horn measure turns out to be piecewise polynomial and was computed explicitly by Coquereaux and Zuber to be given by the formula

$$\frac{V(\nu)V(\tau)}{V(\lambda)V(\mu)} |\mathrm{HIVE}_{\lambda \boxplus \mu \to \nu}| \; d\nu$$

for $\lambda, \mu \in \mathrm{Spec}^{\circ}$, where

$$V(\lambda) = V_n(\lambda) := \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j)$$

is the Vandermonde determinant, and τ is the tuple

$$\tau := (n, n-1, \ldots, 1).$$

Introduction

We introduce the relation

$$\mathrm{diag}(\lambda) \to a$$

for $\lambda \in \mathtt{Spec}$ and $a \in \mathbb{R}^n$ to denote the claim that there exists a Hermitian matrix A with eigenvalues λ and diagonal entries $a_1, \ldots, a_n$. The classical *Schur–Horn theorem* asserts that the relation $\mathrm{diag}(\lambda) \to a$ holds if and only if *majorized* by λ in the sense that one has the trace condition

$$\sum a = \sum \lambda$$

and the majorizing inequalities

$$a_{i_1} + \cdots + a_{i_k} \leq \lambda_1 + \cdots + \lambda_k$$

for all $1 \leq i_1 < \cdots < i_k \leq n$; equivalently, a lies in the *permutahedron* formed by the convex hull of the image of λ under the permutation group S_n.

Introduction

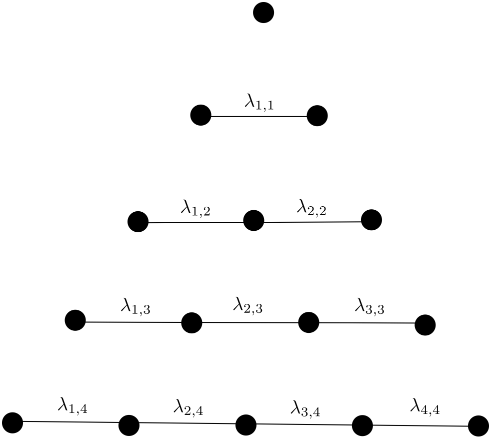


Figure: An $n = 4$ Gelfand–Tsetlin pattern. Each number $\lambda_{i,j}$ in the pattern is greater than or equal to numbers immediately to the northeast or southeast of the pattern; in particular, every row of the pattern is decreasing. Note that such patterns are sometimes depicted as inverted pyramids instead of pyramids in the literature.

Now define a Gelfand–Tsetlin pattern to be a pattern $\gamma = (\lambda_{j,k})_{1 \leq j \leq k \leq n}$ of real numbers obeying the interlacing conditions

$$\lambda_{j,k+1} \geq \lambda_{j,k} \geq \lambda_{j+1,k+1}$$

for $1 \leq j \leq k \leq n-1$. We say that this pattern has boundary condition $\mathrm{diag}(\lambda) \to a$ for some $\lambda \in \mathtt{Spec}_n$ and $a \in \mathbb{R}^n$ if one has

$$\lambda_{j,n} = \lambda_j$$

for $1 \leq j \leq n$ and

$$\sum_{j=1}^{k} \lambda_{j,k} = \sum_{j=1}^{k} a_j$$

for $1 \leq k \leq n$.

We remark that for $\lambda \in \mathtt{Spec}_n^\circ$, $\mathtt{GT}_{\mathrm{diag}(\lambda)\to *}$ is a $\binom{n}{2}$-dimensional polytope, which we call a *Gelfand–Tsetlin polytope*, while $\mathtt{GT}_{\mathrm{diag}(*)\to *}$ is a $\binom{n+1}{2}$-dimensional convex cone, which we call the *Gelfand–Tsetlin cone*.

Introduction

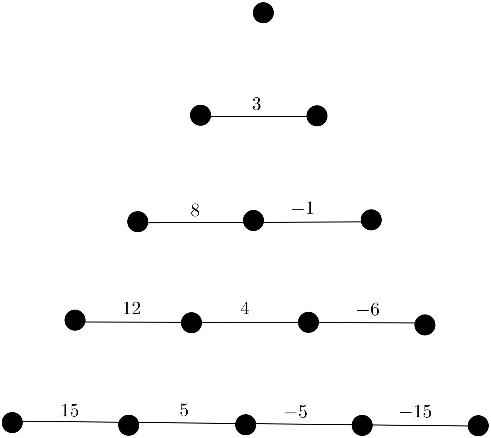


Figure: A Gelfand–Tsetlin pattern with boundary $\mathrm{diag}(15, 5, -5, -15) \to (3, 4, 3, -10)$.

Introduction

We recall the following standard facts about Gelfand–Tsetlin polytopes:

Proposition: Let $\lambda \in \mathtt{Spec}^\circ$.

(i) If $a \in \mathbb{R}^n$, then $\mathrm{diag}(\lambda) \to a$ holds if and only if $\mathtt{GT}_{\mathrm{diag}(\lambda)\to a}$ is non-empty.

(ii) The $\binom{n}{2}$-dimensional volume of $\mathtt{GT}_{\mathrm{diag}(\lambda)\to *}$ is $V(\lambda)/V(\tau)$.

(iii) Let A be a random Hermitian matrix with eigenvalues λ, drawn using the $U(n)$-invariant Haar probability measure. For $1 \leq k \leq n$, let $\lambda_{1,k} \geq \cdots \geq \lambda_{k,k}$ be eigenvalues of the top left $k \times k$ minor of A. Then $(\lambda_{j,k})_{1\leq j\leq k\leq n}$ lies in the polytope $\mathtt{GT}_{\mathrm{diag}(\lambda)\to *}$ with the uniform probability distribution; it has boundary data $\mathrm{diag}(\lambda) \to a$ where $a = (a_{11}, \ldots, a_{nn})$ are the diagonal entries of A.

Introduction

(iv) If $\Lambda \in \mathtt{Spec}$ has *large gaps* in the sense that

$$\min_{1 \leq i < n} \Lambda_i - \Lambda_{i+1} > \lambda_1 - \lambda_n,$$

then there is a volume-preserving linear bijection between $\mathrm{GT}_{\mathrm{diag}(\lambda) \to a}$ and $\mathrm{HIVE}_{\Lambda \boxplus \lambda \to \Lambda + a}$ for any $a \in \mathbb{R}^n$, with a Gelfand–Tsetlin pattern $(\lambda_{j,k})_{1 \leq j \leq k \leq n}$ being mapped to the hive $h \colon T \to \mathbb{R}$ defined by the formula

$$h(i,j) = \Lambda_1 + \cdots + \Lambda_j + \lambda_{1,j} + \cdots + \lambda_{i,j};$$

Introduction

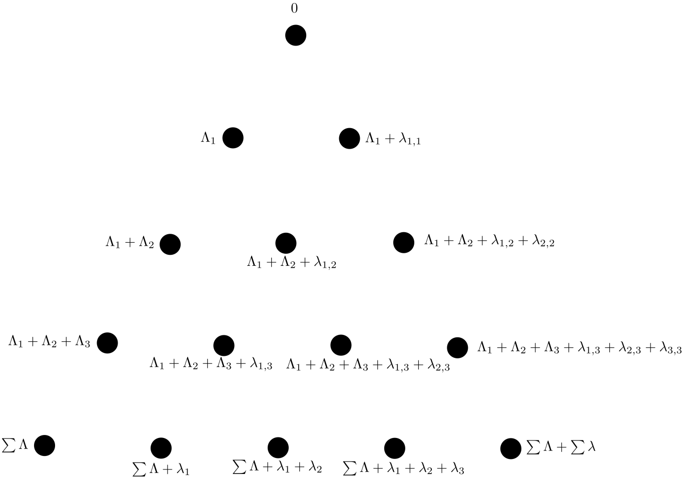


Figure: The hive associated with the Gelfand–Tsetlin pattern in an earlier figure and some large gap tuple Λ.

Introduction

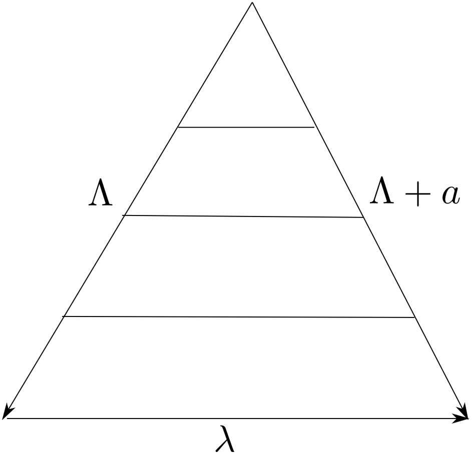


Figure: A schematic depiction of the boundary conditions of the hive in the previous figure. The horizontal "creases" inside the triangle indicate that the rhombus concavity condition is essentially an automatic consequence of the large gaps hypothesis for rhombi that cross these creases.

Introduction

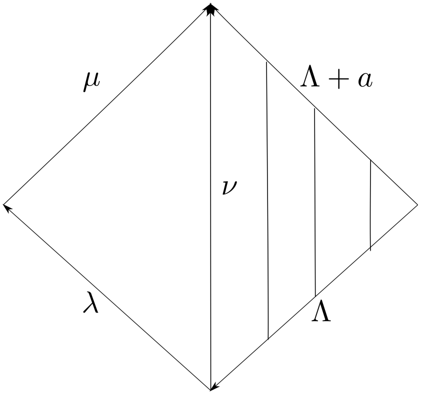


Figure: A schematic depiction of an augmented hive in $\mathtt{AUGHIVE}_{\mathrm{diag}(\lambda \boxplus \mu \to \nu) \to a}$, where we artificially shift by a tuple Λ with large gaps in order to create two hives, instead of a hive and a Gelfand–Tsetlin pattern.

Introduction

It is natural to ask if Lebesgue measure on the polytope $\mathtt{AUGHIVE}_{\mathrm{diag}(\lambda\boxplus\mu\to*)\to*}$ also exhibits concentration. As a first step towards this goal, we are able to establish this for spectra λ, μ that are not deterministic, but are instead drawn from (scalar multiples of) the *GUE ensemble*.

Introduction

To establish normalization conventions, we define a GUE random matrix to be a random Hermitian matrix $M = (\xi_{ij})_{1 \leq i,j \leq n}$ where $\xi_{ij} = \overline{\xi_{ji}}$ for $i < j$ are independent complex gaussians of mean zero and variance 1, ξ_{ii} are independent real gaussians of mean zero and variance 1, independent of the ξ_{ij} for $i < j$.

As is well known, if $\sigma > 0$ and A is a random matrix with $\frac{A}{\sqrt{\sigma^2 n}}$ drawn from the GUE ensemble, then the eigenvalues $\lambda \in \mathtt{Spec}$ of A are distributed with probability density function

$$C_n \sigma^{-\frac{n(n+1)}{2}} \exp\left(-\frac{|\lambda|^2}{2\sigma^2 n}\right) V(\lambda)^2$$

for some constant $C_n > 0$ depending only on n.

Introduction

If $\sigma_\lambda, \sigma_\mu > 0$ are fixed and A, B are independent random matrices with $\frac{A}{\sqrt{\sigma_\lambda^2 n}}, \frac{B}{\sqrt{\sigma_\mu^2 n}}$ drawn from the GUE ensemble, then the the distribution of the eigenvalues of $A + B$ are the pushforward of the measure on the $n(n+1)$-dimensional *augmented hive cone*

$$\mathtt{AUGHIVE}_{\mathrm{diag}(*\boxplus*\to*)\to*} := \bigcup_{\lambda,\mu,\nu,\pi} \left(\mathtt{HIVE}_{\lambda\boxplus\mu\to\nu} \times \mathtt{GT}_{\mathrm{diag}(\nu)\to\pi}\right),$$

where the probability density function of this measure is given by

$$C_{n,\sigma_\lambda,\sigma_\mu} \exp\left(-\frac{|\lambda|^2}{2\sigma_\lambda^2 n} - \frac{|\mu|^2}{2\sigma_\mu^2 n}\right) V(\lambda) V(\mu) \qquad (7)$$

on the slices

$$\mathtt{AUGHIVE}_{\mathrm{diag}(\lambda\boxplus\mu\to*)\to*} := \bigcup_{\nu,\pi}\left(\mathtt{HIVE}_{\lambda\boxplus\mu\to\nu} \times \mathtt{GT}_{\mathrm{diag}(\nu)\to\pi}\right).$$

Introduction

Since GUE matrices have an operator norm of $O(\sqrt{n})$ with overwhelming probability (by which we mean with probability $1 - O(n^{-C})$ for any fixed $C > 0$), the boundary differences λ, μ, ν of an augmented hive (h, γ) drawn from the above measure will be of size $O(n)$ with overwhelming probability, and hence the entries $h(v), v \in T$ of the hive will be of size $O(n^2)$ with overwhelming probability. From this fact (and some crude moment estimates to treat the contribution of the exceptional event), it is not difficult to show the "trivial bound" that the variance $\operatorname{var} h(v)$ of any individual entry $h(v)$ of the hive is bounded by $O(n^4)$.

The main result of this talk is a gain over this trivial bound:

Theorem (Concentration of augmented hives)

Let $\sigma_\lambda, \sigma_\mu > 0$ *be fixed, and let* $(h, \gamma) \in \mathtt{AUGHIVE}_{\mathrm{diag}(*\boxplus*\to*)\to*}$ *be a random augmented hive drawn using the probability measure* (7)*. Then for all* $v \in T$*, we have the variance bound*

$$\mathrm{var}\, h(v) = o(n^4)$$

as $n \to \infty$*, uniformly in* v*.*

Informally, this theorem asserts that randomly selected hives (with GUE boundary data) have an asymptotic limiting profile, at least in a subsequential sense.

Methods of proof

The first step is to exploit the *octahedron recurrence*, which has appeared in the enumerative combinatorics literature several times, which was observed by Knutson, Tao and Woodward to witness an "associativity" property

$$\bigcup_{\nu} \mathtt{HIVE}_{\lambda \boxplus \mu \to \nu} \times \mathtt{HIVE}_{\gamma \boxplus \nu \to \pi} \equiv \bigcup_{\sigma} \mathtt{HIVE}_{\gamma \boxplus \lambda \to \sigma} \times \mathtt{HIVE}_{\sigma \boxplus \mu \to \pi}$$

on hives related to the trivial associativity

$$(A + B) + C = A + (B + C)$$

of the addition operation on Hermitian matrices A, B, C.

Methods of proof

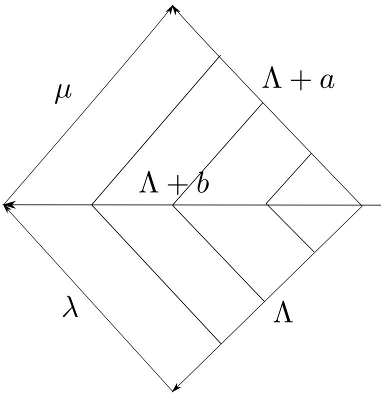


Figure: A schematic depiction of a pair in $\mathtt{GT}_{\mathrm{diag}(\lambda)\to b} \times \mathtt{GT}_{\mathrm{diag}(\mu)\to a-b}$, where an artificial shift by a tuple Λ with large gaps is used to re-interpret this pair as a pair of hives with a common edge.

Methods of proof

In our context (viewing Gelfand–Tsetlin patterns as degenerations of hives), the octahedron recurrence is a piecewise-linear volume-preserving bijection

$$\textbf{oct}: \mathtt{GT}_{\mathrm{diag}(*)\to *} \times \mathtt{GT}_{\mathrm{diag}(*)\to *} \to \mathtt{AUGHIVE}_{\mathrm{diag}(*\boxplus *\to *)\to *}$$

between the two $n(n-1)$-dimensional convex cones $\mathtt{GT}_{\mathrm{diag}(*)\to *} \times \mathtt{GT}_{\mathrm{diag}(*)\to *}$, $\mathtt{AUGHIVE}_{\mathrm{diag}(*\boxplus *\to *)\to *}$.
In fact **oct** is a piecewise-linear volume-preserving bijection

$$\textbf{oct}: \bigcup_b \mathtt{GT}_{\mathrm{diag}(\lambda)\to b} \times \mathtt{GT}_{\mathrm{diag}(\mu)\to a-b} \to \bigcup_\nu \mathtt{AUGHIVE}_{\mathrm{diag}(\lambda\boxplus\mu\to\nu)\to a}$$

for any $\lambda, \mu \in \mathtt{Spec}^\circ$ and $a \in \mathbb{R}^d$.

Methods of proof

Theorem (Excavation form of octahedron recurrence)

Let v be an element of the triangle T. Then there is an explicit finite family $\mathcal{W}_v \colon \mathrm{GT}_{\mathrm{diag}(*)\to *} \times \mathrm{GT}_{\mathrm{diag}(*)\to *} \to \mathbb{R}$ of linear functionals on $\mathrm{GT}_{\mathrm{diag}(*)\to *} \times \mathrm{GT}_{\mathrm{diag}(*)\to *}$, such that whenever $(h, g) = \mathbf{oct}(g_1, g_2)$ is the image of the octahedron recurrence for some $g_1, g_2 \in \mathrm{GT}_{\mathrm{diag}(*)\to *}$, then

$$h(v) = \max_{w \in \mathcal{W}} w(g_1, g_2).$$

The linear functionals are given in terms of lozenge tilings of a certain hexagon $\hexagon_v$ associated to v; this version of the Speyer formula has not explicitly been written previously in the literature, and may be of independent interest.

Methods of proof

Direct calculation reveals that the density function on $\mathtt{AUGHIVE}_{\mathrm{diag}(\lambda\boxplus\mu\to*)\to*}$ is log-concave. The supremum in Theorem 2 can then be handled by the following tool, which may be of independent interest:

Lemma

Let η be an log-concave probability measure in $\mathbb{R}^d$ with finite second moments, and let $\mathcal{W}$ be a family of affine functions $w\colon \mathbb{R}^d \to \mathbb{R}$. Then

$$\operatorname{var}_\eta \left(\sup_{w\in\mathcal{W}_v} w \right) \ll \sup_{w\in\mathcal{W}} (\operatorname{var}_\eta w) \log(2+d).$$

Here of course we use the probabilistic notation $\mathbb{E}_\eta w := \int_{\mathbb{R}^d} w \, d\eta$ and $\operatorname{var}_\eta w := \mathbb{E}_\eta |w|^2 - |\mathbb{E}_\eta w|^2$.

This lemma is a consequence of Cheeger's inequality and recent work of Klartag on the KLS conjecture.

Methods of proof

In view of this lemma, it would now suffice to establish the variance bound

$$\operatorname{var} w(\gamma_1, \gamma_2) = O(n^{4-c})$$

for all $v \in T$ and $w \in \mathcal{W}_v$ and some constant $c > 0$, where (γ_1, γ_2) was the random variable

$$(\gamma_1, \gamma_2) = \left((\lambda_{j,k})_{1 \leq j \leq k \leq n}, (\mu_{j,k})_{1 \leq j \leq k \leq n}\right);$$

the additional factor of n^{-c} is needed to overcome the logarithmic loss in the preceding Lemma. This is a variance estimate for linear statistics of the GUE minor process.

As it turns out, the covariance estimates for eigenvalue gaps of GUE established by Cipolloni, Erdős and Schröder, combined with some further manipulations from the theory of determinantal processes to analyze the minor process, are *almost* enough to obtain this sort of bound; there is however a technical difficulty because these bounds are only established in the bulk of the spectrum and not on the edge. However, the contributions coming from the edge region can be controlled by relatively crude estimates, and after removing these contributions to focus on the bulk contribution we will be able to make the above strategy work.

Poincaré inequalities over log-concave measures

We establish a useful Poincaré inequality over log-concave measures. Namely, we show

Proposition (Poincaré inequality on log-concave measures)

Let η be an log-concave probability measure in $\mathbb{R}^d$ with finite second moments, and define the $d \times d$ inertia matrix M by the formula

$$M := \mathbb{E}_\eta x x^T - (\mathbb{E}_\eta x)(\mathbb{E}_\eta x)^T.$$

Then for any Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$, one has

$$\operatorname{var}_\eta f \ll \left(\mathbb{E}_\eta |\nabla f|^2\right) \|M\|_{\mathrm{op}} \log(2+d).$$

Poincaré inequalities over log-concave measures

Without loss of generality, we may assume that η is an *isotropic measure* in the sense that

$$\mathbb{E}_\eta x = 0; \quad \mathbb{E}_\eta x x^T = I_d.$$

Define the *Cheeger constant* $D_{\mathrm{Che}}(\eta)$ of η (with respect to the Euclidean inner product) by the formula

$$D_{\mathrm{Che}}(\eta) := \inf_{A \subset \mathbb{R}^d} \frac{\int_{\partial A} \rho}{\min(\eta(A), 1 - \eta(A))}$$

where the infimum runs over all open subsets A of $\mathbb{R}^d$ with smooth boundary with $0 < \eta(A) < 1$, and ∂_A is integrated using surface measure.

Proof.

By the Cheeger inequality, one has the Poincaré inequality

$$D_{\mathrm{Che}}(\eta)^2 \operatorname{var}_\eta f \ll \mathbb{E}_\eta |\nabla f|^2$$

so the task reduces to (and is in fact equivalent to) the lower bound

$$D_{\mathrm{Che}}(\eta) \gg \frac{1}{\sqrt{\log(2+d)}}$$

on the Cheeger constant of an isotropic log-concave measure.

But this follows from recent work of Klartag (building upon previous advances by Chen, Klartag-Lehec and Lee-Vempala. □

Poincaré inequalities over log-concave measures

Proposition (Weighted Poincaré inequality on log-concave measures)

Let η be an log-concave probability measure in $\mathbb{R}^d$ with finite second moments. Express $\mathbb{R}^d$ as a Cartesian product $\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_k}$ for some $d_1, \ldots, d_k$ summing to d (so that a vector $x \in \mathbb{R}^d$ is expressed as $(x_1, \ldots, x_k)$ for $x_j \in \mathbb{R}^{d_j}$), and for each $i = 1, \ldots, k$, and define the $d_j \times d_j$ inertia matrix M_j by the formula

$$M_j := \mathbb{E}_\eta x_j x_j^T - (\mathbb{E}_\eta x_j)(\mathbb{E}_\eta x_j)^T.$$

Then for any Lipschitz function $f \colon \mathbb{R}^d \to \mathbb{R}$, and any weights $\alpha_1, \ldots, \alpha_k > 0$, one has

$$\mathrm{var}_\eta f \ll \left(\mathbb{E}_\eta \sum_{j=1}^{k} \alpha_j |\nabla_j f|^2 \right) \left(\sum_{j=1}^{k} \alpha_j^{-2} \|M_j\|_{\mathrm{op}}^2 \right)^{1/2} \log(2+d).$$

Proof.

By pushing forward η by the map $(x_1, \dots, x_k) \mapsto (\alpha_1^{1/2} x_1, \dots, \alpha_k^{1/2} x_k)$ we may normalize $\alpha_j = 1$ for all j, so that $\sum_{j=1}^k \alpha_j |\nabla_j f|^2 = |\nabla f|^2$. By the previous proposition, it thus suffices to establish the bound

$$\|M\|_{\mathrm{op}} \leq \left(\sum_{j=1}^k \|M_j\|_{\mathrm{op}}^2 \right)^{1/2}.$$

□

Proof (continued).

For any $x = (x_1, \ldots, x_k) \in \mathbb{R}^d$, where $x_j \in \mathbb{R}^{d_j}$ for all j, it follows from the positive semi-definiteness of M and the triangle inequality followed by Cauchy–Schwarz that

$$\begin{aligned}(x^T M x)^{1/2} &\leq \sum_{j=1}^{k} (x_j^T M_j x_j)^{1/2} \\ &\leq \sum_{j=1}^{k} \|M_j\|_{\mathrm{op}} |x_j| \\ &\leq \left(\sum_{j=1}^{k} \|M_j\|_{\mathrm{op}}^2 \right)^{1/2} |x|\end{aligned}$$

and the claim follows. □

The octahedron recurrence

Let $\lambda, \mu, \gamma, \pi \in \mathtt{Spec}^\circ$. Knutson-Tao-Woodward constructed a volume-preserving, piecewise-linear *octahedron recurrence*

$$\textbf{oct}\colon \bigcup_{\sigma} \mathtt{HIVE}_{\sigma \boxplus \mu \to \pi} \times \mathtt{HIVE}_{\gamma \boxplus \lambda \to \sigma} \to \bigcup_{\nu} \mathtt{HIVE}_{\lambda \boxplus \mu \to \nu} \times \mathtt{HIVE}_{\gamma \boxplus \nu \to \pi} \tag{8}$$

and an explicit formula for it given using the work of Speyer. We now recall a version of this formula that will be convenient for our purposes. We will identify a pair $(h, h') \in \mathtt{HIVE}_{\lambda \boxplus \mu \to \nu} \times \mathtt{HIVE}_{\gamma \boxplus \nu \to \pi}$ with a single function $\tilde{h}\colon T \cup T' \to \mathbb{R}$ defined on the square $T \cup T' = \{0, \ldots, n\}^2$.

The octahedron recurrence

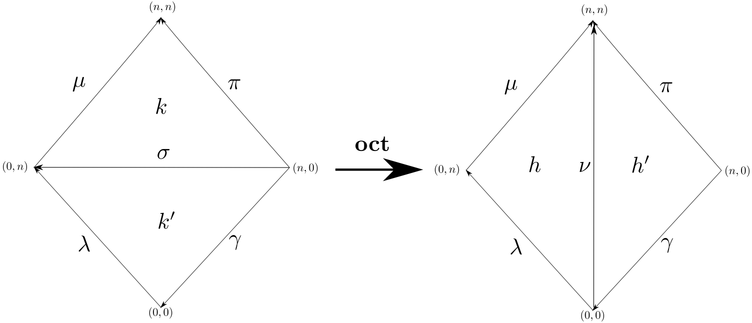


Figure: A schematic depiction of the octahedron recurrence that transforms one pair $(k, k') \in \mathtt{HIVE}_{\sigma \boxplus \mu \to \pi} \times \mathtt{HIVE}_{\gamma \boxplus \lambda \to \sigma}$ of hives into another $(h, h') \in \mathtt{HIVE}_{\lambda \boxplus \mu \to \nu} \times \mathtt{HIVE}_{\gamma \boxplus \nu \to \pi}$. The hives h, h', k, k' have been shifted to lie on triangles T, T', U, U' respectively.

The octahedron recurrence

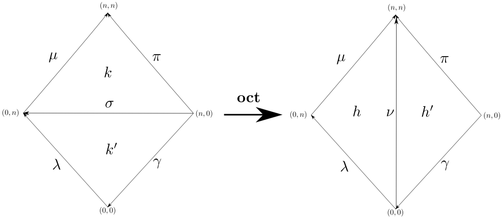
$\tilde{h}$ is defined by the formula

$$\tilde{h}(i,j) := h(i,j) \tag{9}$$

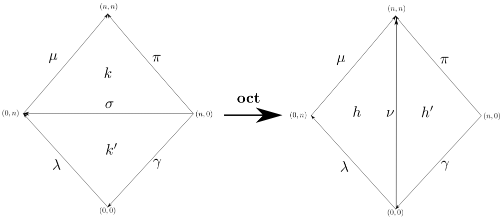
when (i,j) lies in the triangle $T := \{(i,j) : 0 \le i \le j \le n\}$, and

$$\tilde{h}(i,j) := h'(j, n-i+j) - \sum \gamma \tag{10}$$

when (i,j) lies in the opposite triangle
$T' := \{(i,j) : 0 \le j \le i \le n\}$.



Note that both definitions agree on the diagonal $\{(i,i) : 0 \leq i \leq n\}$ due to the boundary values of the hives h, h'. The function $\tilde{h}$ will be rhombus concave on T and on T', but not necessarily concave along rhombi that cross the diagonal separating T and T'.



The octahedron recurrence

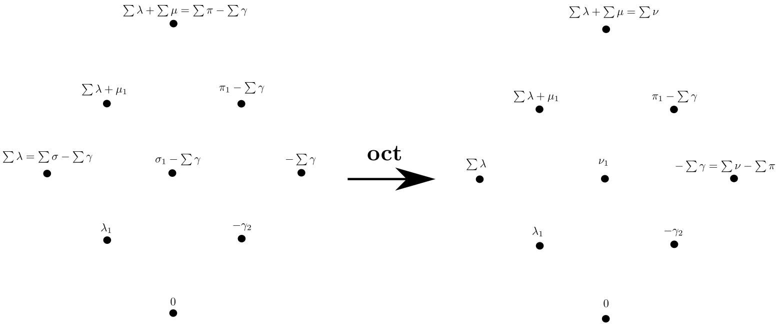


Figure: The $n = 2$ case of the octahedron recurrence. One can determine the value ν_1 in the right image from the data in the left image by the formula $\nu_1 = \max(\sum\lambda + \mu_1 + \gamma_1 - \sigma_1, \lambda_1 + \pi_1 - \sigma_1)$.

The octahedron recurrence

In a similar vein, we identify a pair $(k, k') \in \mathtt{HIVE}_{\sigma \boxplus \mu \to \pi} \times \mathtt{HIVE}_{\gamma \boxplus \lambda \to \sigma}$ with a single function $\tilde{k} \colon U \cup U' \to \mathbb{R}$ defined on the square. Even though $T \cup T'$ and $U \cup U'$ are both technically equal to the same set $\{0, \dots, n\}^2$, it is conceptually better to think of these sets as being distinct (except on the boundary). We will view these two copies of $\{0, \dots, n\}^2$ as the upper and lower faces respectively of a certain tetrahedron $\mathtt{tet}$, and the octahedron recurrence **oct** can be constructed by "excavating" that tetrahedron.

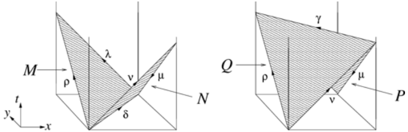
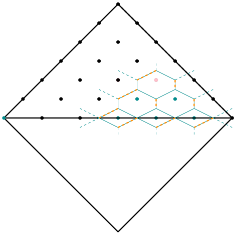


Figure: Lower and upper panels of $\mathtt{tet}$.
Image credits, Henriques and Kamnitzer

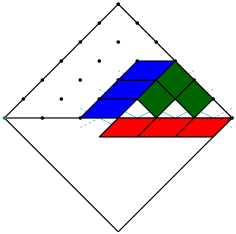
The octahedron recurrence

For vertices $v = (i, j)$ in the interior or $\{0, \ldots, n\}^2$, the octahedron recurrence specifying $\tilde{h}(i, j)$ is more complicated to describe. It was initially defined by recursively "excavating" a real-valued function on a tetrahedron $\{(a, b, c, d) \in \mathbb{Z}^4 : a, b, c, d \geq 0; a + b + c + d = n\}$ with $\tilde{k}$ describing the values on the top two faces, and $\tilde{h}$ the bottom two faces.

An alternate description was given by Speyer, in terms of perfect matchings of an “excavation graph” associated to (i, j). We will use a modification of Speyer’s formula that is more convenient for our purposes, in which the perfect matchings are replaced the dual concept of a *lozenge tiling*. To describe this formula we need some definitions.



(a) Perfect matching



(b) Lozenge (and triangle) tiling

Figure: Correspondence between perfect matchings and lozenge and triangle tilings

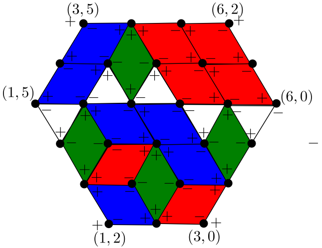
Definition (Lozenges and border triangles)

A *lozenge* is a quadruple $ABCD$ in U or U' that is one of following three forms for some $i, j \in \mathbb{Z}$:

(i) $(A, B, C, D) = ((i, j), (i+1, j-1), (i+2, j-1), (i+1, j))$

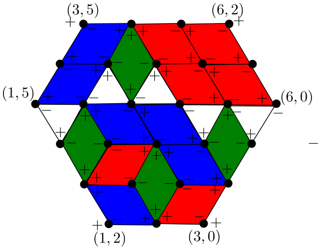
(ii) $(A, B, C, D) = ((i, j), (i, j+1), (i-1, j+2), (i-1, j+1))$

(iii) $(A, B, C, D) = ((i, j), (i+1, j), (i+1, j+1), (i, j+1))$.



Definition (Continued)

Lozenges of type (i) will be called blue if they lie in U and red if they lie in U'; lozenges of type (ii) will be called red if they lie in U and blue if they lie in U'; and lozenges of type (iii) that lie either in U or in U' will be called green. A quadruple of the form (iii) that crosses the diagonal separating U and U' is *not* considered to be a lozenge, but instead splits into two border triangles as defined below.



Definition (Continued)

A *border edge* is an edge AC of the form $(A, C) = ((i, n-i), (i+1, n-i-1))$ for some $0 \leq i < n$; the border edges thus separate U and U'. Each border edge $(A, C) = ((i, n-i), (i+1, n-i-1))$ is bordered by two *border triangles ABC*, defined as follows:

- (Upward triangle) $(A, B, C) = ((i, n-i), (i+1, n-i), (i+1, n-i-1))$.
- (Downward triangle) $(A, B, C) = ((i, n-i), (i, n-i-1), (i+1, n-i-1))$.

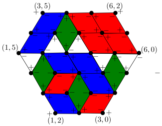


Figure: A typical lozenge tiling of $\hexagon_{(3,2)}$, $n = 6$.

The octahedron recurrence

Given a lozenge $\diamond = ABCD$ and a function $\tilde{k}\colon \{0,\dots,n\}^2 \to \mathbb{R}$ defined as before, we define the *weight* $\mathbf{wt}(\diamond) = \mathbf{wt}(\diamond, \tilde{k})$ to be the quantity

$$\mathbf{wt}(\diamond) := \frac{1}{3}(\tilde{k}(A) + \tilde{k}(C) - \tilde{k}(B) - \tilde{k}(D)).$$

Similarly, given a border triangle $\Delta = ABC$, the weight $\mathbf{wt}(\Delta) = \mathbf{wt}(\tau, \tilde{k})$ is defined as

$$\mathbf{wt}(\Delta) := \frac{1}{3}(\tilde{k}(A) - \tilde{k}(B)).$$

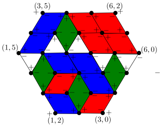


Figure: A typical lozenge tiling of $\hexagon_{(3,2)}$, $n = 6$.

The octahedron recurrence

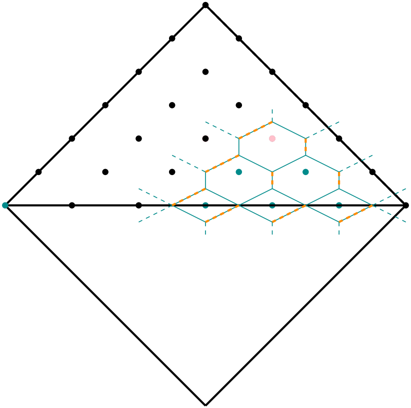


Figure: The two bottom panels have been shown. The green and pink dots correspond to vertices that have been excavated on the two top faces. The pink dot also corresponds to the vertex on the top face at which we are computing the value of the hive. A matching of the graph has been depicted in dashed yellow lines.

The octahedron recurrence

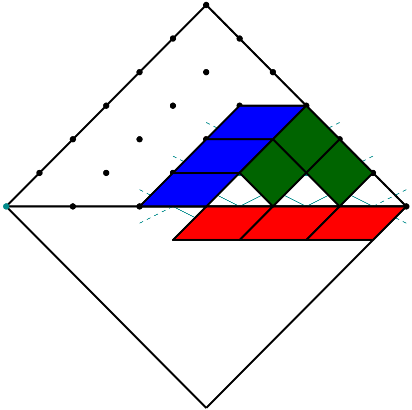


Figure: The matching in the preceding figure has been converted into a lozenge tiling. Lozenges corresponding to edges of the three different orientations have been colored differently.

The octahedron recurrence

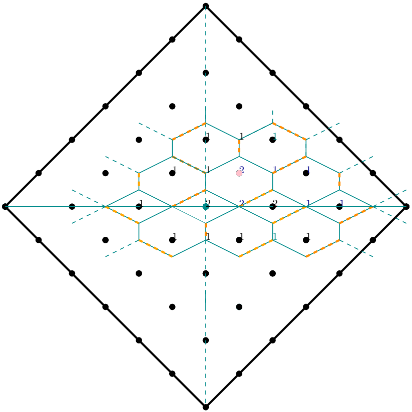


Figure: The pink dot corresponds to the vertex on the top face at which we are computing the value of the hive. The numbers at the sites tell us how many hive excavations have taken place at the respective positions. If there is no number at a site, there are no excavations at the corresponding vertex. A matching of the graph has been depicted in dashed yellow lines.

The octahedron recurrence

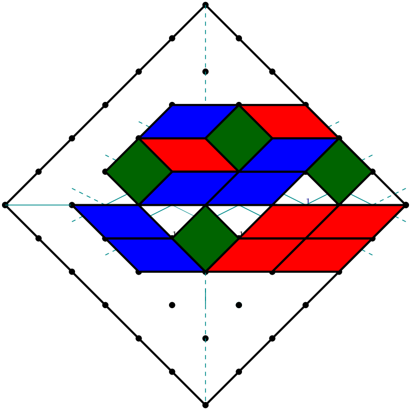


Figure: The matching in the preceding figure has been converted into a lozenge tiling. Lozenges corresponding to edges of the three different orientations have been colored differently.

The octahedron recurrence

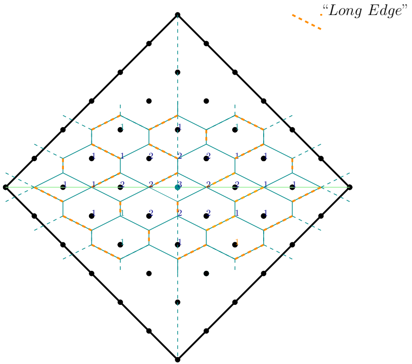


Figure: The green dot corresponds to the vertex on the top face at which we are computing the value of the hive. Since the bottom faces correspond to modified GT patterns, the direction of the "Long Edge" is that in which the honeycomb edges corresponding to the hive that the GT pattern has been converted into are long.

The octahedron recurrence

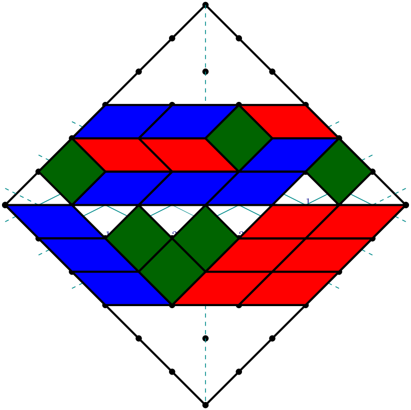


Figure: The matching in the preceding figure has been converted into a lozenge tiling. Lozenges corresponding to edges of the three different orientations have been colored differently.

The octahedron recurrence

Definition (Octahedron recurrence)

If $v = (i, j)$ lies in the interior of $\{0, \dots, n\}^2 = T \cup T'$, then the *excavation hexagon* $\hexagon_v = ABCDEF$ in $\{0, \dots, n\}^2 = U \cup U'$ centered at v is defined as follows:

- If $v \in T$ (i.e., $i \leq j$), then

$$(A, B, C, D, E, F) =$$

$$((0, n), (0, j), (i, j - i), (n + i - j, j - i), (n + i - j, j), (i, n)).$$

- If $v \in T'$ (i.e., $i \geq j$), then

$$(A, B, C, D, E, F) =$$

$$((i - j, n + j - i), (i - j, j), (i, 0), (n, 0), (n, j), (i, n + j - i)).$$

Note that these two definitions agree when $v \in T \cap T'$ (i.e., when $i = j$). The original point $v = (i, j)$ is then the intersection of the diagonals BE and CF. The line AD is called the *equator*; it lies on the border between U and U'.

Definition

The *weight* $\mathbf{wt}(\hexagon_v) = \mathbf{wt}(\hexagon_v, \tilde{k})$ of this hexagon is defined as

$$\mathbf{wt}(\hexagon_v) := \frac{1}{3}(\tilde{k}(B) + \tilde{k}(C) - \tilde{k}(D) + \tilde{k}(E) + \tilde{k}(F)). \qquad (11)$$

Definition

A *lozenge tiling* Ξ of the excavation hexagon $\hexagon_v$ is a partition of the (solid) hexagon into (solid) lozenges and (solid) border triangles, such that each border edge on the equator is adjacent to exactly one border triangle in the tiling. An example of a lozenge tiling is the *standard lozenge tiling* Ξ_0, in which the trapezoid $ABEF$ is tiled by blue lozenges in U and by green lozenges and downward border triangles in U', while the opposite trapezoid $BCDE$ is tiled by green lozenges and upward border triangles in U and by red lozenges in U'.

Definition

The *weight* $w_\Xi = w_\Xi(\tilde{k})$ of such a tiling is defined to be the sum of the weights of all the lozenges $\diamond$ and triangles Δ in the tiling, as well as the weight of the entire hexagon $\hexagon_v$:

$$w_\Xi := \sum_{\diamond \in \Xi} \mathbf{wt}(\diamond) + \sum_{\Delta \in \Xi} \mathbf{wt}(\Delta) + \mathbf{wt}(\hexagon_v). \qquad (12)$$

Note that the w_Ξ depend linearly on $\tilde{k}$, and hence on k, k'. We then define

$$\tilde{h}(v) := \max_{\Xi \text{ tiles } \hexagon_v} w_\Xi$$

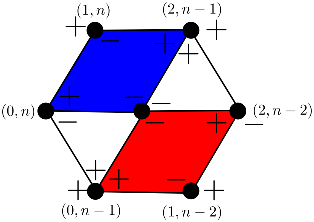


Figure: The two lozenge tilings of $\hexagon_{(1,n-1)}$. The weight coefficients of the lozenges, border triangles, and hexagon are marked with $+$ (for a weight of $+1/3$) and $-$ (for a weight of $-1/3$). The tiling on the left is the standard one.

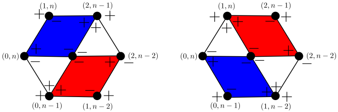
Example

As a simple example, take $v = (1, n-1) \in T$ (assuming $n \geq 2$), then the excavation hexagon $\hexagon_v = ABCDEF$ is given by the unit hexagon centered at v:

$$(A, B, C, D, E, F) =$$

$$((0, n), (0, n-1), (1, n-2), (2, n-2), (2, n-1), (1, n)).$$

This hexagon has two lozenge tilings.



Example (continued)

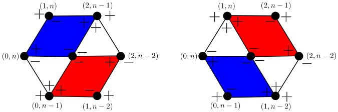
For the tiling on the left, the blue lozenge has weight

$$\frac{1}{3}(\tilde{k}(0,n)+\tilde{k}(2,n-1)-\tilde{k}(1,n)-\tilde{k}(1,n-1))$$

the red lozenge has weight

$$\frac{1}{3}(\tilde{k}(0,n-1)+\tilde{k}(2,n-2)-\tilde{k}(1,n-1)-\tilde{k}(1,n-2))$$

.

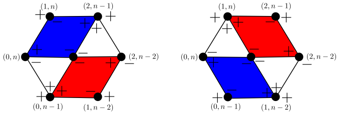


Example (continued)

The upward triangle has weight $\frac{1}{3}(\tilde{k}(2, n-1) - \tilde{k}(1, n-1))$ the downward triangle has weight $\frac{1}{3}(\tilde{k}(0, n-1) - \tilde{k}(0, n))$ and the hexagon has weight

$$\frac{1}{3}(\tilde{k}(0, n-1) + \tilde{k}(1, n-2) - \tilde{k}(2, n-2) + \tilde{k}(2, n-1) + \tilde{k}(1, n)$$

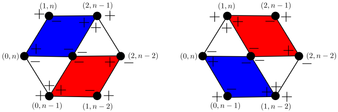
leading to a total weight of $\tilde{k}(2, n-1) + \tilde{k}(0, n-1) - \tilde{k}(1, n-1)$.



Example (continued)

The tiling on the right can similarly be computed to have a total weight of $\tilde{k}(1,n)+\tilde{k}(1,n-2)-\tilde{k}(1,n-1)$ leading to the familiar octahedron relation

$$\tilde{h}(1,n-1) = \max(\tilde{k}(2,n-1)+\tilde{k}(0,n-1),$$
$$\tilde{k}(1,n)+\tilde{k}(1,n-2)) - \tilde{k}(1,n-1).$$



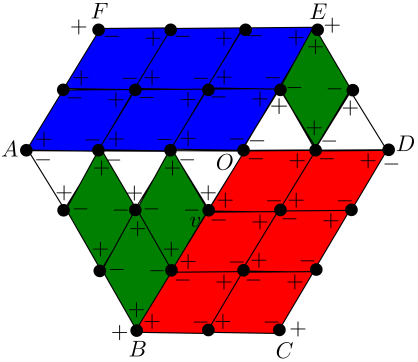


Figure: The standard lozenge tiling of a hexagon *ABCDEF* centered at v. The total weight of this tiling is $\tilde{k}(E) + \tilde{k}(B) - \tilde{k}(O)$, where O is the intersection of the diagonal BE with the equator AD.

Example (continued)

The lozenge tiling in below figure has weight

$$\tilde{k}(4,4)+\tilde{k}(3,3)-\tilde{k}(4,2)-\tilde{k}(2,3)+\tilde{k}(1,3)$$

$$-\tilde{k}(2,2)-\tilde{k}(3,1)+\tilde{k}(4,0)+\tilde{k}(2,1).$$

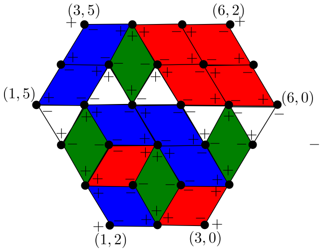


Figure: A typical lozenge tiling of $\hexagon_{(3,2)}$, $n=6$.

The octahedron recurrence

Theorem

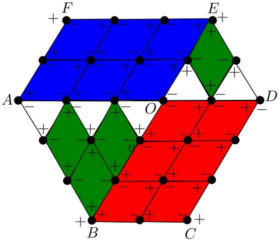
The construction above agrees with the octahedron recurrence described by Knutson-Tao-Woodward.

Proof.

This is basically a matter of comparing notations with the Speyer formula and performing some calculations. □

The octahedron recurrence

As a consequence of this theorem and the results in Knutson-Tao-Woodward [KTW], the octahedron recurrence that we have just defined is indeed a volume-preserving bijection between the polytopes. In [KTW] the stronger assertion that this recurrence is a bijection between integer lattice points is established, but the volume-preserving nature of the bijection then follows by a standard rescaling and limiting argument to pass from the discrete to the continuous setting.



Lemma (Replacing red lozenges with blue)

Let v be an interior point of $\{0, \ldots, n\}^2$, and let Ξ be a lozenge tiling of $\hexagon_v = ABCDEF$. Then one has the identity

$$\sum_{\diamond \in \Xi,\ red} \mathbf{wt}(\diamond) - \sum_{\diamond \in \Xi,\ blue} \mathbf{wt}(\diamond) =$$

$$\frac{1}{3}(-\tilde{k}(A) + \tilde{k}(B) - \tilde{k}(C) + \tilde{k}(D) - \tilde{k}(E) + \tilde{k}(F)).$$

The proof of the main theorem is thus reduced to

Proposition (Reduction to the minor process)

Let $\sigma_\lambda, \sigma_\mu > 0$ be fixed, and let A, B with $\frac{A}{\sqrt{\sigma_\lambda^2 n}}, \frac{B}{\sqrt{\sigma_\mu^2 n}}$ be drawn independently from the GUE ensemble, and let (g, g') be the resulting Gelfand–Tsetlin patterns. Then for any $v \in T$, we have

$$\operatorname{var} \max_{\Xi \text{ tiles } \hexagon_v} 2 \sum_{\diamond \in \Xi, \text{ blue}} \mathbf{wt}(\diamond) + \sum_{\diamond \in \Xi, \text{ green}} \mathbf{wt}(\diamond) + \sum_{\Delta \in \Xi} \mathbf{wt}(\Delta)$$

$$+\mathbf{wt}'(\hexagon_v) = o(n^4)$$

where we identify (g, g') with a pair of hives (k, k') using a large gaps tuple γ as indicated above.

Using eigenvalue rigidity to remove edge contributions

Lemma (Eigenvalue rigidity (Tao-Vu'13))

Let A be a matrix with $A/\sqrt{n}$ having the distribution of GUE. Then for any $1 \leq i \leq n$ we have

$$\mathbb{P}(n^{-1/3} \min(i, n-i+1)^{1/3} |\lambda_i - \sqrt{n}\gamma_i| \geq T) \ll n^{O(1)} exp(-cT^c)$$

for any $T > 0$ and some absolute constant $c > 0$, where the classical location γ_i is the value predicted by the semicircular law:

$$\int_{-\infty}^{\gamma_i} \frac{1}{2\pi}(4-x^2)^{1/2}\, dx = \frac{i}{n}.$$

In particular,

$$\lambda_i, \mathbb{E}\lambda_i = \sqrt{n}\gamma_i + O(n^{1/3} \min(i, n-i+1)^{-1/3} \log^{O(1)} n)$$

with overwhelming probability.

Using eigenvalue rigidity to remove edge contributions

We conclude that for any lozenge tiling Ξ of $\hexagon_v$, we have

$$\sum_{\Delta \in \Xi} \mathbf{wt}(\Delta) = \sum_{\Delta \in \Xi} \mathbb{E}\mathbf{wt}(\Delta) + O(n^{4/3} \log^{O(1)} n)$$

with overwhelming probability. The weight $\mathbf{wt}'(\hexagon_v)$ is a certain linear combination of the eigenvalues λ_i, μ_j with bounded coefficients. By a preceding lemma, we conclude that

$$\begin{aligned} \mathbf{wt}'(\hexagon_v) &= \mathbb{E}\mathbf{wt}'(\hexagon_v) + O\left(\sum_{i=1}^{n} n^{1/3} \min(i, n-i+1)^{-1/3} \log^{O(1)} n\right) \\ &= \mathbb{E}\mathbf{wt}'(\hexagon_v) + O(n \log^{O(1)} n). \end{aligned}$$

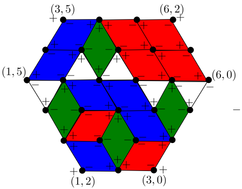
Again, the contribution of the $O(n \log^{O(1)} n)$ error is acceptable, so we may also replace $\mathbf{wt}'(\hexagon_v)$ by $\mathbb{E}\mathbf{wt}'(\hexagon_v)$.

Using eigenvalue rigidity to remove edge contributions

Let $\epsilon > 0$ be a small parameter, and let U_ϵ denote the portion of U that lies at Euclidean distance at least ϵn from the boundary of U. Define U'_ϵ similarly. Using telescoping sums, we show

$$\sum_{\diamond \not\subset U_\epsilon \cup U'_\epsilon} |\mathbf{wt}(\diamond)| \ll \epsilon^{1/3} n^2$$

with overwhelming probability, where the sum is over all blue or green lozenges in U or U' that are not contained in U_ϵ or U'_ϵ.



Henceforth we fix $\epsilon > 0$ and assume n sufficiently large depending on ϵ. By the triangle inequality, it thus suffices to establish the bound

$$\mathrm{var} \sum_{\diamond \in \Xi,\ \text{blue}: \diamond \subset U'_\epsilon} \mathbf{wt}(\diamond) = O(n^{4-c+o(1)}) \qquad (13)$$

and similarly with blue replaced by green, or U'_ϵ replaced by U_ϵ, or both.

We focus on establishing (13), as the other three cases are proven similarly. It suffices to establish the bound

$$\mathrm{var} \sum_{(j,k) \in \Omega} \lambda_{j,k+1} - \lambda_{j,k} = O(n^{4-c+o(1)})$$

whenever Ω is a collection of tuples of integers $1 \leq j \leq k \leq n$ with $j, k-j, n-k \gg \epsilon n$.

By the triangle inequality, it suffices to show that

$$\operatorname{var} \sum_{j \in S_k} \lambda_{j,k+1} - \lambda_{j,k} = O(n^{2-c+o(1)})$$

for each $\epsilon n \ll k \leq n-1$, where S_k is some subset of the bulk region $\{1 \leq j \leq k : j, k-j \gg \epsilon n\}$. Since the minor of a GUE matrix is a rescaled version of a GUE matrix, it suffices to establish this claim for the case $k = n-1$, that is to say (after adjusting ϵ slightly) to show that

$$\operatorname{var} X_S = O(n^{2-c+o(1)})$$

for an arbitrary subset S of $\{2\epsilon n \leq j \leq (1-2\epsilon)n\}$, where X_S denotes the random variable

$$X_S := \sum_{j \in S} \lambda_j - \lambda_{j,n-1}.$$

It is convenient to exclude a small exceptional set to keep the eigenvalues λ_j somewhat under control. From a lemma of Tao and Vu, we already know that there is a constant C_0 such that

$$|\lambda_j - \sigma_\lambda \gamma_j n^{1/2}| \leq n^{1/3} \min(j, n-j+1)^{-1/3} \log^{C_0} n \tag{14}$$

for all $1 \leq j \leq n$ with overwhelming probability. From the Wegner estimate (see Erdős-Schlein-Yau) and enlarging C_0 if needed, we also see that

$$|\lambda_{j+1} - \lambda_j| \geq \exp(-\log^{C_0} n) \tag{15}$$

with overwhelming probability for all $\epsilon n \leq j \leq (1-\epsilon)n$. Thus, if we let E denote the event that (14), (15) both hold for all $\epsilon n \leq j \leq (1-\epsilon)n$, then E holds with overwhelming probability; for future reference we also note the constraints (14), (15) defining E are restricting λ to a certain convex subset of $\texttt{Spec}$. It suffices to show that

$$\mathrm{var}(X_S|E) = O(n^{2-c+o(1)}).$$

We split this by conditioning on the spectrum λ of A. By the law of total variance (noting that the event E is measurable with respect to λ), it suffices to establish the bounds

$$\mathrm{var}(\mathbb{E}(X_S|\lambda)|E) = O(n^{2-c+o(1)}) \tag{16}$$

and

$$\mathbb{E}(\mathrm{var}(X_S|\lambda)|E) = O(n^{2-c+o(1)}). \tag{17}$$

To prove (17), we expand out the left-hand side as

$$\sum_{i,j\in S} \mathbb{E}(\mathrm{cov}(\lambda_i - \lambda_{i,n-1}, \lambda_j - \lambda_{j,n-1}|\lambda)|E)$$

where we use $\mathrm{cov}(X, Y) := \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y)$ to denote the covariance between two random variables X, Y.

We will indicate below how to establish the bound

$$\mathbb{E}(\mathrm{cov}(\lambda_i - \lambda_{i,n-1}, \lambda_j - \lambda_{j,n-1} | \lambda) | E) \ll \frac{n^{o(1)}}{(1 + |i - j|)^2} \qquad (18)$$

for all $2\epsilon n \le i, j \le (1 - 2\epsilon)n$, which implies (17).

Since the event E is restricting λ to a convex set in $\mathbb{R}^n$, so the probability distribution function of λ is still log-concave after conditioning to E. Thus Poincaré estimates such as Proposition 1 become available. As it turns out, a direct application of this proposition gives unfavorable estimates, basically because of long-range correlations between λ_i and λ_j make the operator norm of the inertia matrix large, and also because the known correlation decay estimates are currently only available in the bulk. To resolve this, we do not use the standard basis $e_1, \ldots, e_n$ of $\mathbb{R}^n$, but instead the following basis consisting of three groups:

- The vector $e_1 + \cdots + e_n$.
- The vectors $e_{i+1} - e_i$ for i in the bulk region $\texttt{bulk} := \{i : \epsilon n \leq i < (1-\epsilon)n\}$.
- The vectors $e_{i+1} - e_i$ for i in the edge region $\texttt{edge} := \{i : 1 \leq i < \epsilon n \text{ or } (1-\epsilon)n \leq i < n\}$.

The point is that $\mathbb{E}(X_S|\lambda)$ has different behavior in each of the three groups of vectors. In the direction $e_1 + \cdots + e_n$, the function $\mathbb{E}(X_S|\lambda)$ is in fact constant. This is because once one conditions on λ, the random variable $\lambda_{j,n-1}$ has the distribution of the j^{th} largest eigenvalue of the top left $n-1 \times n-1$ minor of a Hermitian matrix chosen uniformly at random amongst all matrices with eigenvalue λ. Moving λ in the direction $e_1 + \cdots + e_n$ then amounts to shifting λ_j and $\lambda_{j,n-1}$ by the same constant, so the expectation $\mathbb{E}(X_S|\lambda)$ remains unchanged.

As it turns out, $\mathbb{E}(X_S|\lambda)$ is significantly more sensitive to the bulk eigenvalue gaps $\lambda_{i+1} - \lambda_i$ than the edge eigenvalue gaps $\lambda_{j+1} - \lambda_j$. To exploit this, we apply the weighted log-concave Poincaré inequality with suitable choices of weights (sending the weight on the basis vector $e_1 + \cdots + e_n$ to infinity) to conclude that

$$\begin{aligned} \mathrm{var}(\mathbb{E}(X_S|\lambda)|E) \ll \mathbb{E}\left(|\nabla_{\texttt{bulk}}\mathbb{E}(X_S|\lambda)|^2 + n|\nabla_{\texttt{edge}}\mathbb{E}(X_S|\lambda)|^2|E\right) \\ \times \left(\|M_{\texttt{bulk}}\|_{\mathrm{op}} + n^{-1}\|M_{\texttt{edge}}\|_{\mathrm{op}}\right)\log n \end{aligned} \tag{19}$$

where for $\Omega = \texttt{bulk}, \texttt{edge}$ one has

$$|\nabla_\Omega \mathbb{E}(X_S|\lambda)|^2 := \sum_{i\in\Omega} |(\partial_{\lambda_{i+1}} - \partial_{\lambda_i})\mathbb{E}(X_S|\lambda)|^2$$

and M_Ω is the covariance matrix with entries

$$\mathrm{cov}(\lambda_{i+1} - \lambda_i, \lambda_{j+1} - \lambda_j|E)$$

for $i, j \in \Omega$.

To prove (17), it suffices to establish the bound

$$\mathbb{E}\left(|\nabla_{\texttt{bulk}}\mathbb{E}(X_S|\lambda)|^2 + n|\nabla_{\texttt{edge}}\mathbb{E}(X_S|\lambda)|^2|E\right) \ll n^{1+o(1)}. \quad (20)$$

We establish the bound

$$\mathbb{E}(|\partial_{\lambda_i}\mathbb{E}(\lambda_j - \lambda_{j,n-1}|\lambda)|^2|E) \ll n^{o(1)}(1 + n|\gamma_i - \gamma_j|)^{-4} \quad (21)$$

whenever $1 \leq i \leq n$ and $j \in \texttt{bulk}$. Taking square roots and summing over $j \in S$ using the triangle inequality, one obtains

$$\mathbb{E}(|\partial_{\lambda_i}\mathbb{E}(X_S|\lambda)|^2|E) \ll n^{-2+o(1)}$$

for $i \in \texttt{edge}$, and

$$\mathbb{E}(|\partial_{\lambda_i}\mathbb{E}(X_S|\lambda)|^2|E) \ll n^{o(1)}$$

for $i \in \texttt{bulk}$. Summing in i, one obtains (20) and thus (17).

Determinantal process calculations

It thus remains to establish the bounds (18), (21). We will indicate some key ingredients.

We fix λ to be a deterministic element of $\mathtt{Spec}^\circ$, and let A be a Hermitian matrix drawn uniformly at random amongst all matrices with eigenvalues λ. We then let $x_1 \geq \cdots \geq x_{n-1}$ be the eigenvalues of the top left $n-1 \times n-1$ minor of A. In order to establish (18), (21), we would like to understand the mean and covariances of the gaps $\lambda_j - x_j$, as these random variables have the same distribution as $\lambda_j - \lambda_{j,n-1}$ conditioned to this choice of λ. As it turns out, the theory of determinantal processes provide an explicit formula for these quantities:

Proposition (4; First and second moments)

$$\mathbb{E}(\lambda_i - x_i) = \int_{I_i} Q_i(x)\, dx \tag{22}$$

for all $1 \le i \le n$ *and*

$$\operatorname{cov}(\lambda_i - x_i, \lambda_j - x_j) = \left(\int_{I_i} (1 - Q_j(x))\, dx \right) \left(\int_{I_j} Q_i(x)\, dx \right) \tag{23}$$

for all $1 \le i < j < n$*, where* $I_j = I_{j,\lambda}$ *is the interval* $I_j := [\lambda_{j+1}, \lambda_j]$ *and each* $Q_j = Q_{j,\lambda}$ *is the unique degree* $n-1$ *polynomial such that* $Q_j(\lambda_i) = 1_{i \le j}$ *for* $1 \le i \le n$*. More explicitly, by the Lagrange interpolation formula one has*

$$Q_j(x) := \sum_{i \le j} \frac{\prod_{\ell \ne i}(x - \lambda_\ell)}{\prod_{\ell \ne i}(\lambda_i - \lambda_\ell)}. \tag{24}$$

Lemma (Contour integral representation)

Let $P = P_\lambda$ denote the degree n polynomial

$$P(x) := \prod_{k=1}^{n} (x - \lambda_k).$$

Then for any $1 \leq j \leq n$ and σ in the interior of I_j, one has

$$Q_j(x) = -\frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{P(x)}{P(z)(x-z)}\, dz$$

for $x < \sigma$ and

$$1 - Q_j(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{P(x)}{P(z)(x-z)}\, dz$$

for $x > \sigma$.

Proof of the contour integral representation lemma:

Proof.

Observe that the rational function $Q_j(x)/P(x)$ decays at infinity and has poles at $\lambda_i, i \leq j$ with residues $1/P'(\lambda_i)$, thus we have the partial fractions decomposition

$$\frac{Q_j(x)}{P(x)} = \sum_{i \leq j} \frac{1}{P'(\lambda_i)(x - \lambda_i)}.$$

Similarly

$$\frac{1 - Q_j(x)}{P(x)} = \sum_{i > j} \frac{1}{P'(\lambda_i)(x - \lambda_i)}.$$

The claim now follows from the residue theorem. □

Open questions

1. What can be said about the concentration of random real valued augmented hives with general boundary conditions? If they do concentrate, what are the possible subsequential limit shapes? In particular, is the limit unique? In the limit when one of the boundary conditions is more spread out than the other, the limit shape should essentially degenerate to fractional free convolution powers (See Shlyakhtenko and Tao).
2. Do the local statistics of the random augmented GUE hive process converge (either in the bulk or the edge) to a known limit? In the case of the random Gelfand–Tsetlin process, the limit is known to essentially be the Boutillier bead process.
3. Do random integer valued augmented hives with general boundary conditions concentrate? Again, if they do concentrate, what are the possible subsequential limit shapes? In particular, is the limit unique?

Thank you for your attention!

Proof of Proposition 4:

Each x_i lies in I_i, with probability measure

$$(n-1)!\frac{V_{n-1}(x)}{V_n(\lambda_1,\dots,\lambda_n)} 1_{I_1}(x_1)\dots 1_{I_{n-1}}(x_{n-1})dx_1\dots dx_{n-1}. \quad (25)$$

As observed by Metcalfe, this law also has a determinantal form involving the polynomials Q_j as follows. From the fundamental theorem of calculus, the derivatives Q_j' are degree $n-2$ polynomials that obey the mean zero conditions

$$\int_{I_i} Q_j'(x)dx = 1_{i=j} \quad (26)$$

for $1 \leq i,j \leq n-1$, and thus form a basis of the polynomials of degree at most $n-2$.

Proof of Proposition 4:

If one introduces the kernel $K: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by the formula

$$K(x,y) := \sum_{j=1}^{n-1} 1_{I_j}(x) Q_j'(y)$$

then from (26) we conclude that K is a rank $n-1$ projection in the sense that

$$\int_{\mathbb{R}} K(x,y)K(y,z)dy = K(x,z),$$

and $\int_{\mathbb{R}} K(y,y)dy = n-1$ for all $x, z \in \mathbb{R}$.

Lemma (Gaudin's lemma)

Let $K : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *be a (sufficiently nice) function obeying the idempotent relation*

$$K(x,z) = \int_{\mathbb{R}} K(x,y)K(y,z)\, dy$$

and the trace formula

$$\int_{\mathbb{R}} K(x,x)\, dx = n - 1.$$

For each k, let ρ_k *be the correlation function*

$$\rho_k(x_1, \dots, x_k) := \det(K(x_i, x_j))_{1 \le i,j \le k}.$$

Then we have

$$\int_{\mathbb{R}} \rho_{k+1}(x_1, \dots, x_k, x_{k+1})\, dx_{k+1} = (n - k - 1)\rho_k(x_1, \dots, x_k).$$

Proof of Proposition 4:

By the Gaudin lemma, we then have

$$\int_{\mathbb{R}^{n-1}} \det(K(x_i, x_j))_{1 \le i,j \le n-1} dx_1 \dots dx_{n-1} = (n-1)!.$$

On the other hand this determinant is symmetric and supported on the $(n-1)!$ permutations of $I_1 \times \cdots \times I_{n-1}$, hence

$$\int_{I_1 \times \cdots \times I_{n-1}} \det(K(x_i, x_j))_{1 \le i,j \le n-1} dx_1 \dots dx_{n-1} = 1.$$

Proof of Proposition 4:

Because $Q'_1, \ldots, Q'_{n-1}$ is a basis of the polynomials of degree $n-2$, we see that for $(x_1, \ldots, x_{n-1})$ in $I_1 \times \cdots \times I_{n-1}$, the determinant

$$\det(K(x_i, x_j))_{1 \le i,j \le n-1} = \det(Q'_i(x_j))_{1 \le i,j \le n-1}$$

is a scalar multiple of the Vandermonde determinant, while also having a total mass 1; comparing this with the probability measure (25), we see that this measure has the determinantal form

$$\det(K(x_i, x_j))_{1 \le i,j \le n-1}\, dx_1 \ldots dx_{n-1}.$$

Proof of Proposition 4:

In particular (by another application of the Gaudin lemma) the one-point correlation function is $K(x,x)$ and the two point correlation function is $K(x,x)K(y,y) - K(x,y)K(y,x)$. The identity (22) then follows from integration by parts:

$$\begin{aligned} \mathbb{E}\lambda_i - x_i &= \int_{I_i} (\lambda_i - x) K(x,x)\, dx \\ &= \int_{I_i} (\lambda_i - x) Q_i'(x)\, dx \\ &= \int_{I_i} Q_i(x)\, dx. \end{aligned}$$

Proof of Proposition 4:

Proof (continued).

A similar computation gives (23): $\mathrm{cov}(\lambda_i - x_i, \lambda_j - x_j)$

$$\begin{aligned}
&= \int_{I_i}\int_{I_j}(\lambda_i - x)(\lambda_j - y)(K(x,x)K(y,y) - K(x,y)K(y,x))\,dxdy \\
&\quad - \int_{I_i}(\lambda_i - x)K(x,x)dx\int_{I_j}(\lambda_j - y)K(y,y)\,dy \\
&= -\int_{I_1}\int_{I_j}(\lambda_i - x)(\lambda_j - y)K(x,y)K(y,x)\,dxdy \\
&= -\int_{I_i}\int_{I_j}(\lambda_i - x)(\lambda_j - y)Q_i'(y)Q_j'(x)\,dxdy \\
&= \left(\int_{I_i}(1 - Q_j(y))dy\right)\left(\int_{I_j}Q_i(x)\,dx\right). \qquad (27)
\end{aligned}$$

□

Proof:

Now let $1 \leq i, j \leq n$. Using the identity

$$\frac{P(x)}{P(z)(x-z)} = \prod_{1 \leq k \leq n : k \neq i} \frac{x - \lambda_k}{z - \lambda_k} \left(\frac{1}{z - \lambda_i} + \frac{1}{x - z} \right)$$

we have

$$\partial_{\lambda_i} \frac{P(x)}{P(z)(x-z)} = \frac{1}{(z - \lambda_i)^2} \prod_{1 \leq k \leq n : k \neq i} \frac{x - \lambda_k}{z - \lambda_k}$$

and so on differentiating under the integral sign we obtain

$$\partial_{\lambda_i} Q_j(x) = -\frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \frac{1}{(z - \lambda_i)^2} \prod_{1 \leq k \leq n : k \neq i} \frac{x - \lambda_k}{z - \lambda_k} \, dz \qquad (28)$$

whenever σ is in the interior of I_j and $x \neq \sigma$.

Proof:

By continuity the restriction $x \neq \sigma$ can then be dropped. Setting $x = \sigma$, which implies $|z - \lambda_k| \geq |x - \lambda_k|$, we conclude from the triangle inequality that

$$|\partial_{\lambda_i} Q_j(x)| \leq \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{|x - \lambda_i + it|^2} \, dt = \frac{1}{2|x - \lambda_i|}. \qquad (29)$$

Further arguments lead to a proof of (18) and (21) by controlling

$$\mathbb{E}(\mathrm{cov}(\lambda_i - \lambda_{i,n-1}, \lambda_j - \lambda_{j,n-1}|\lambda)|E),$$

and

$$\mathbb{E}(|\partial_{\lambda_i} \mathbb{E}(\lambda_j - \lambda_{j,n-1}|\lambda)|^2|E).$$

We begin with (21). Fix $1 \leq i \leq n$ and $j \in \texttt{bulk}$. By Proposition 4, the left-hand side of (21)

$$\mathbb{E}(|\partial_{\lambda_i}\mathbb{E}(\lambda_j - \lambda_{j,n-1}|\lambda)|^2|E)$$

is

$$\mathbb{E}\left(|\partial_{\lambda_i}\int_{I_{j,\lambda}} Q_{j,\lambda}(x)\,dx|^2\Big|E\right).$$

We divide the interval $I_{j,\lambda} = [\lambda_{j+1}, \lambda_j]$ into the left half $I^l_{j,\lambda} = [\lambda_{j+1}, \frac{\lambda_{j+1}+\lambda_j}{2}]$ and the right half $I^r_{j,\lambda} = [\frac{\lambda_{j+1}+\lambda_j}{2}, \lambda_j]$.

We shall just establish the bound

$$\mathbb{E}\left(|\partial_{\lambda_i} \int_{I^r_{j,\lambda}} Q_{j,\lambda}(x)\, dx|^2 \middle| E \right) \ll n^{o(1)}(1 + n|\gamma_i - \gamma_j|)^{-4}; \qquad (30)$$

similar arguments apply for the left half $I^l_{j,\lambda}$, and the claim (21) will then follow from the triangle inequality.

The quantity $\int_{I^r_{j,\lambda}} Q_{j,\lambda}(x)\, dx$ is unchanged if all of the λ are shifted by the same constant. In particular

$$\sum_{i=1}^{n} \partial_{\lambda_i} \int_{I^r_{j,\lambda}} Q_{j,\lambda}(x)\, dx = 0.$$

Thus it will suffice to establish (30) under the additional hypothesis $i \neq j$, as the excluded case $i = j$ is then handled by the triangle inequality. The point of this reduction is that it generates a separation between λ_i and $I^r_{j,\lambda}$. (For the left half $I^l_{j,\lambda}$, one would instead enforce the hypothesis $i \neq j + 1$.)

Henceforth $i \neq j$. If i is also not equal to $j+1$, we of course have

$$\partial_{\lambda_i} \int_{I^r_{j,\lambda}} Q_{j,\lambda}(x)\, dx = \int_{I^r_{j,\lambda}} \partial_{\lambda_i} Q_{j,\lambda}(x)\, dx.$$

For $i = j+1$, we acquire an additional term of $\frac{1}{2} Q_{j,\lambda}(\frac{\lambda_{j+1}+\lambda_j}{2})$.

Thus, it will suffice to establish the bounds

$$\mathbb{E}\left(|\int_{I^r_{j,\lambda}} \partial_{\lambda_i} Q_{j,\lambda}(x)\, dx|^2 | E \right) \ll n^{o(1)} (1 + n|\gamma_i - \gamma_j|)^{-4} \qquad (31)$$

whenever $i \neq j$, as well as the additional bound

$$\mathbb{E}\left(|Q_{j,\lambda}\left(\frac{\lambda_{j+1} + \lambda_j}{2}\right)|^2 | E \right) \ll n^{o(1)}. \qquad (32)$$

By (14), $I_{j,\lambda}$ is contained in a fixed interval I_j^* of length $n^{o(1)}$ centered around $\sigma_\lambda \sqrt{n}\gamma_j$, thus by Cauchy–Schwarz

$$\left| \int_{I_{j,\lambda}^r} \partial_{\lambda_i} Q_{j,\lambda}(x)\, dx \right|^2 \ll n^{o(1)} \int_{I_j^*} |\partial_{\lambda_i} Q_{j,\lambda}(x)|^2 1_{x \in I_{j,\lambda}^r}\, dx.$$

By the triangle inequality, (31) will then follow from the pointwise bound

$$\mathbb{E}\left(|\partial_{\lambda_i} Q_{j,\lambda}(x)|^2 1_{x \in I^r_{j,\lambda}} | E\right) \ll n^{o(1)}(1 + n|\gamma_i - \gamma_j|)^{-4} \qquad (33)$$

for each $x \in I_j^*$.

First consider the case where $|i-j| \geq \log^{2C_0} n$. Applying (28) with $\sigma = x$ and the triangle inequality, we have

$$\partial_{\lambda_i} Q_j(x) \ll \int_{\mathbb{R}} \frac{1}{(x-\lambda_i)^2} \prod_{1 \leq k \leq n : k \neq i} \frac{|x-\lambda_k|}{|x-\lambda_k+it|}\, dt.$$

From (14), we can compute $|x-\lambda_i| = n^{o(1)}(1+n|\gamma_i-\gamma_j|)$ and $\prod_{1 \leq k \leq n : k \neq i} \frac{|x-\lambda_k|}{|x-\lambda_k+it|} \ll \frac{n^{o(1)}}{1+t^2}$, and (33) follows in this case.

Now suppose $|i-j| < \log^{2C_0} n$, so that the right-hand side of (33) simplifies to $n^{o(1)}$. By (29), we can bound

$$|\partial_{\lambda_i} Q_{j,\lambda}(x)|^2 1_{x \in I^r_{j,\lambda}} \ll \frac{1}{\lambda_j - \lambda_{j+1}} + \frac{1}{\lambda_{j-1} - \lambda_j}$$

(by splitting into the cases $i < j$ and $i > j$). Thus it will suffice to establish the bound

$$\mathbb{E}\left(\frac{1}{(\lambda_j - \lambda_{j+1})^2} | E \right) \ll n^{o(1)}$$

(the claim for $\frac{1}{\lambda_{j-1}-\lambda_j}$ is of course similar).

Letting $K(x, y)$ be the determinantal kernel of the rescaled GUE matrix A it suffices to show that

$$\int_{I_j^*}\int_{I_j^*} \frac{K(x,x)K(y,y) - K(x,y)K(y,x)}{|x-y|^2}\, dxdy \ll n^{o(1)}.$$

But from existing results on the local smooth convergence of this kernel to a rescaled Dyson sine process, the claim follows. This proves (33).

Finally, we need to show (32). Write $x = \frac{\lambda_{j+1}+\lambda_j}{2}$. By the contour integral representation lemma and the Plemelj formula, we can write

$$Q_{j,\lambda}(x) = \frac{1}{2} - \frac{1}{2\pi}\mathrm{p.v.}\int_{\mathbb{R}} \frac{P(x)}{P(x+it)}\frac{dt}{t}.$$

Using (14), (15) (which among other things makes $P(x)/P(x+it)$ very close to 1 for $|t| \leq \exp(-\log^{2C_0} n)$, bounded in magnitude by 1 for all t, and decaying fast for $|t| \geq n^{C_0}$ (say)), one can calculate that this integral is $O(n^{o(1)})$, giving (32). This completes the proof of (21).

Now we show (18). Let $2\epsilon n \leq i, j \leq (1-2\epsilon)n$. If $|i-j| \leq \log^{2C_0} n$ then the claim follows from the crude bounds

$$\lambda_i - \lambda_{i,n-1}, \lambda_j - \lambda_{j,n-1} = O(n^{o(1)})$$

from interlacing and (14), so we may assume by symmetry that $j - i > \log^{2C_0} n$. Applying (23), it suffices to show the pointwise bound

$$\int_{I_j}\int_{I_i}(1 - Q_j(x))Q_i(y)\,dxdy \ll \frac{n^{o(1)}}{(j-i)^2}.$$

Applying Lemma 19, we can write the left-hand side as

$$\frac{1}{4\pi^2}\int_{I_j}\int_{I_i}\int_{x-i\infty}^{x+i\infty}\int_{y-i\infty}^{y+i\infty}\frac{P(x)P(y)}{P(w)P(z)(y-z)(x-w)}\,dwdzdxdy.$$

From the separation of i, j, we have the lower bounds

$$|y-z|, |x-w| \gg j - i.$$

The quantity $|P(x)|/|P(z)|$ is bounded by 1, and from (14) it is also bounded by $O(n^{o(1)}/|\mathrm{Im} z|^2)$ when $\mathrm{Im} z \geq \log^{2C_0} n$. Similarly for $|P(y)|/|P(w)|$. Also, from (14) the intervals I_i, I_j have length $O(n^{o(1)})$, and the claim follows.