hexagon:2611.01322v1

Four-ranks of class groups and the spectrum of Legendre symbols

Zhongpai Gao
gaozhongpai@gmail.com

Abstract

We study how much of the narrow 4-rank k of a real quadratic field can be recovered from the complete Legendre matrix of its squarefree radicand d. Let n be the number of prime factors and $N = 2^{n-1-k}$. If exactly N eigenvalues lie outside $[-1, 1]$, the count recovers k. At a prescribed 4-rank, for radicands $d \le X$ with all prime factors 1 modulo 4 and r prime factors above a suitable cutoff, the success probability tends to $\exp(-e^{-c})$ if $N \to \infty$ and $r/N - \log N \to c$, and to 1 or 0 above or below this window. This is unconditional when all prime factors exceed the cutoff. The real Dirichlet critical-line theorem extends it to a prescribed set of $O(\log\log\log X)$ small prime factors. We identify the critical empirical law through a free multiplicative convolution of Poisson laws and construct a family where the small primes change the scale of both extreme eigenvalues. On the broader family, the two edge fluctuations have independent Tracy–Widom limits at each fixed 4-rank.

2020 Mathematics Subject Classification. Primary 11R29, 60B20; secondary 11R45, 46L54, 05C50.
Keywords. Class group, 4-rank, Rédei matrix, Legendre symbol, random matrix, free multiplicative convolution, Tracy–Widom law.

1 Introduction

How much of the narrow 4-rank of $\mathbb{Q}(\sqrt{d})$ can be read from the real eigenvalues of its complete Legendre matrix? Let $d > 1$ be squarefree, with prime factors $p_1 < \cdots < p_n$, none congruent to 3 modulo 4. The real symmetric matrix K_d has zero diagonal and entry (p_i/p_j) for $i < j$. Encoding the signs as bits and setting the diagonal to give zero row sums yields the Rédei matrix R_d over $\mathbb{F}_2$. The Rédei rank formula [10, 12, 19] gives

$$k = \mathrm{rk}_4\,\mathrm{Cl}^+(d) = \dim\ker R_d - 1, \qquad s = n - 1 - k, \qquad N = 2^s.$$

Let M_d count the eigenvalues of K_d outside $[-1, 1]$, with multiplicity. Proposition 2.3 gives $M_d \le N$. When $M_d = N$, the count recovers $k = n - 1 - \log_2 M_d$.

The recovery threshold. The count has a sharp transition at $r \sim N\log N$, where r counts the prime factors above a cutoff y. For $d \le X$ with all prime factors 1 modulo 4 and $|r - \frac{1}{2}\log\log X| \le (\log\log X)^{2/3}$, we may take $y = (\log X)^{64}$. The threshold is unconditional when all factors exceed y. Section 6 uses the real Dirichlet critical-line theorem to allow up to h factors at or below y, with $2^h \le (\log\log X)^J$ for fixed J.

For a prescribed small-prime part and feasible k, sample d uniformly in the corresponding family. As $N \to \infty$, the probability of $M_d = N$ tends to $\exp(-e^{-c})$ if $r/N - \log N \to c$ (Corollary 6.6). If this difference tends to $+\infty$ or $-\infty$, the probability tends to 1 or 0, respectively. The rounding rule using only the largest eigenvalue holds in the denser range $N(s+1) = o(r)$ and can fail at the threshold.

What the spectrum records at other ranks. Theorem 6.4 gives the empirical spectral law at every feasible rank. At the sparse end, both edge fluctuations retain their independent Tracy–Widom limits in the stated rank range. At intermediate ranks, several small primes can belong to one class under the relation $p_i \sim p_j$ if $e_i + e_j \in \ker R_d$. Their multiplicity enters the scale of both extremes. Example 6.5 constructs discriminants in the typical window for which both extreme magnitudes of the complete matrix are asymptotically larger than those of its large-prime submatrix.

For the family with every prime factor above y, Theorem 5.2 gives the spectral laws at every 4-rank using the relative pattern comparison of [7]. The empirical law $\nu_{r/N}$ interpolates between the semicircle law and the dense regime. For $r/N \to \lambda \in (0, \infty)$, Theorem 3.4 gives the law of the squared eigenvalues in terms of a free multiplicative convolution of Poisson laws. On the broader family $\mathcal{D}(X)$ of all the admissible squarefree integers, the comparison from Chan, Koymans, Milovic and Pagano [6] has an absolute error. It gives joint Tracy–Widom limits independent of a fixed 4-rank (Corollary 4.10). The rare large ranks require the relative estimates of Sections 5 and 6.

Method and earlier work. Fouvry and Klüners [10], and subsequent work [12, 6], use the binary corank to study the 4-rank. Here the same symbols also define a real spectral statistic. Lemma 2.1 labels each row by a vector in $\mathbb{F}_2^s$. For growing N, the count M_d equals N precisely when every possible label appears. The recovery threshold is therefore the threshold for seeing every label.

The row labels also control moments of the empirical law and products of signs at the edges. Prior sign-matrix results [15, 14] give semicircle and norm bounds. The correlated-matrix theorems [3, 8] have hypotheses beyond the comparison proved here. The quadratic graph model of [4, Remark 4.21] does not reach the regime with fewer binary coordinates than rows. Section 3.5 states the observed-block edge theorem, whose Chebyshev approximation argument is in the supplement.

2 The quotient model

For $r \geq 2$, choose independent uniform bits b_{ij}, $i < j$, and let $R = L_b$ be their binary graph Laplacian. Its real sign lift is

$$K_{ii} = 0, \qquad K_{ij} = (-1)^{b_{ij}} \quad (i \neq j).$$

Write $\kappa(R) = \dim \ker R - 1$ and let W_r have the unconditioned law of K, the zero-diagonal Rademacher Wigner law.

Put $n = r - 1$ and $s = n - k$. The rank constraint has an exact sampling representation; its only additional restriction is that the sampled vectors span.

Lemma 2.1 (Quotient representation)**.** *Choose $C \in \mathbb{F}_2^{n \times s}$ uniformly among matrices of full column rank, and independently choose a uniform invertible symmetric $s \times s$ matrix B. If $v_1, \ldots, v_n$ are the rows of C and $v_r = \sum_{i=1}^n v_i$, then*

$$R_{ij} = v_i^{\mathsf{T}} B v_j$$

has the uniform binary row-sum-zero law conditioned on $\kappa(R) = k$.

Proof. The principal $n \times n$ minor S determines R uniquely: $R = USU^{\mathsf{T}}$, where U has first n rows I_n and last row $\mathbf{1}^{\mathsf{T}}$. In particular the two ranks coincide. We represent S as CBC^{T}. Every symmetric matrix of rank s has exactly $|\mathrm{GL}_s(\mathbb{F}_2)|$ such representations: the columns of C form a basis of its image, and each basis determines a unique invertible symmetric B. Thus the minor is uniform in its rank stratum. The definition of v_r gives the unique row-sum-zero extension. The empty matrices give the same statement when $s = 0$. □

Occupancies and characters of the same kernel. Put $N = 2^s$, $\lambda = r/N$, and, for $a, b \in \mathbb{F}_2^s$, define

$$U_{ab} = N^{-1/2}(-1)^{a^{\mathsf{T}}Bb}, \quad Z_{ia} = \mathbf{1}_{\{v_i = a\}}, \quad D = Z^{\mathsf{T}}Z, \quad \Delta_{ii} = (-1)^{v_i^{\mathsf{T}}Bv_i}.$$

Character orthogonality gives the common identities

$$U = U^{\mathsf{T}}, \quad U^2 = I, \quad |U_{ab}| = N^{-1/2}, \qquad K = \sqrt{N}\, ZUZ^{\mathsf{T}} - \Delta, \quad \|\Delta\| = 1. \tag{1}$$

The nonzero spectrum of $(K + \Delta)/\sqrt{r}$ agrees with that of $\lambda^{-1/2} D^{1/2} U D^{1/2}$. The diagonal entries of D are the quotient occupancies; write $\ell_r = \max_a D_a$. Their classes are intrinsic: $v_i = v_j$ if and only if $e_i + e_j \in \ker R$, since the full quotient-row matrix has full column rank. Moreover $K + \Delta$ has rank at most N and Frobenius norm r. Thus $\|K\| \geq r/\sqrt{N} - 1$, and at most N eigenvalues lie outside $[-1, 1]$, by the min–max principle. Also $\operatorname{tr} U = \sqrt{N}$ for alternating B and zero otherwise, since $a^{\mathsf{T}}Ba$ is zero or a nonzero linear form. These facts control dense samples; the character estimate below controls sparse ones.

2.1 Averages of products of signs

For a simple graph F on the vertex set $[r]$, write $v(F)$ for the number of its nonisolated vertices, A_F for its adjacency matrix over $\mathbb{F}_2$, and

$$\chi_F(K) = \prod_{ij \in F} K_{ij}, \qquad \eta_{r,s}(v) = 2 \min\{2^{s-n}, 5\, 2^{v-s}\}.$$

Proposition 2.2 (Comparison after fixing part of the matrix)**.** *Fix $A \subseteq [r]$, put $h = |A|$, $M = n - h$, and let $\mathcal{C}$ be any event of positive probability determined by B and the labels $(v_i)_{i \in A}$. In particular it may prescribe any feasible internal sign pattern on A. For a nonempty graph F outside A, with $u = v(F)$ and $h + u < r$,*

$$|\mathbb{E}[\chi_F(K) \mid \kappa(R) = k, \mathcal{C}]| \leq 2^{-s \operatorname{rank}(A_F)/2} + 2\sqrt{5}\, 2^{(u-M)/2}. \tag{2}$$

There is no factor involving $\Pr(\mathcal{C} \mid \kappa = k)^{-1}$. Without the additional condition, the sharper bound is

$$|\mathbb{E}[\chi_F(K) \mid \kappa(R) = k]| \leq 2^{-s \operatorname{rank}(A_F)/2} + \eta_{r,s}(v(F)), \qquad v(F) < r. \tag{3}$$

Consequently, for a multilinear polynomial $P = \sum_F \widehat{P}(F) \chi_F$ supported on at most $L < r$ vertices in each nonconstant term,

$$|\mathbb{E}[P \mid \kappa(R) = k] - \mathbb{E}_{\mathrm{Wig}} P| \leq \{2^{-s} + \eta_{r,s}(L)\} \sum_{F \neq \varnothing} |\widehat{P}(F)|. \tag{4}$$

With $\mathcal{C}$, the same assertion holds for terms outside A and $h + L < r$, with bracket $2^{-s} + 2\sqrt{5}\, 2^{(L-M)/2}$.

Proof. Initially sample the n input rows independently, and write E for full column rank. Any proper subset of the r labels consists of independent linear forms in those inputs. Rebase the inputs so that $A \cup V(F)$ are coordinate rows; this preserves E and includes the forced final label whenever it is selected.

Fix B and the labels in A, whose span has dimension a. Their remaining M independent uniform labels give full span exactly when their projections modulo that span cover a space of dimension $t = s - a$. A feasible exposure has $t \leq M$. Its spanning probability is $\prod_{i=0}^{t-1}(1 - 2^{i-M}) > 1/4$, and its failure probability is less than 2^{t-M} (zero for $t = 0$). For $u \leq t$, the selected u projections are independent except with probability less than 2^{u-t}, or less than $4\, 2^{u-t}$ after E. Given independence,

their laws agree by GL_t symmetry. The components in the exposed span remain independent uniforms. Thus the selected labels change in total variation by at most

$$\min\{2^{t-M}, 5\,2^{u-t}\} \le \sqrt{5}\,2^{(u-M)/2}.$$

For $u > t$ the second bound is trivial. Averaging over the actual posterior of the exposed labels and B costs no conditioning mass.

Before spanning, the character is a quadratic phase with polar matrix $A_F \otimes B$, of rank $s\,\mathrm{rank}(A_F)$. For a quadratic form Q on $\mathbb{F}_2^d$ with polar form $\mathcal{B}$, differencing gives

$$\left|2^{-d}\sum_x (-1)^{Q(x)}\right|^2 = 2^{-d}\sum_{z\in\operatorname{rad}\mathcal{B}} (-1)^{Q(z)},$$

which is zero or $2^{-\operatorname{rank}\mathcal{B}}$. This proves the first term, uniformly in B. Twice the total-variation bound proves (2). For $A = \varnothing$, $t = s$ and $M = n$ give (3) before the geometric-mean bound. Finally every nonempty A_F has even rank at least two, while its Wigner character has mean zero. Summing coefficients proves the polynomial comparisons. □

2.2 One matrix at a time

The identities (1) use no randomness. They hold for every symmetric binary matrix R with zero row sums, with s its rank: the principal minor of R factors as CBC^{T} by the counting argument in the proof of Lemma 2.1. We record their consequences in the form used for class groups. Call two indices *linked* if $e_i + e_j \in \ker R$. This is an equivalence relation, and its classes are the sets of rows with a common label.

Proposition 2.3 (Rank and spectrum of a single matrix)**.** *Let R be a symmetric $r \times r$ matrix over $\mathbb{F}_2$ with zero row sums, let $k = \dim\ker R - 1$, and let K be its sign matrix. Then*

1. *at most $2^{\,r-1-k}$ eigenvalues of K lie outside $[-1, 1]$;*
2. *$\|K\| \ge r\,2^{-(r-1-k)/2} - 1$, equivalently*
$$k \le r - 1 - 2\log_2 \frac{r}{\|K\| + 1};$$
3. *$\|K\| \ge \ell - 1$, where ℓ is the size of the largest class of linked indices.*

In particular, for d squarefree with r prime factors, none congruent to 3 modulo 4,

$$\mathrm{rk}_4\,\mathrm{Cl}^+(d) \le r - 1 - 2\log_2 \frac{r}{\|K_d\| + 1}. \tag{5}$$

Proof. Put $M = K + \Delta = \sqrt{N}\,ZUZ^{\mathsf{T}}$ with $N = 2^{\,r-1-k}$. All entries of M are ± 1, so its Frobenius norm is r, and its rank is at most N. Since $\|\Delta\| = 1$, Weyl's inequality gives the first assertion. A matrix of rank at most N has operator norm at least its Frobenius norm divided by $\sqrt{N}$; hence $\|K\| \ge \|M\| - 1 \ge r/\sqrt{N} - 1$. For the third, the rows in one class have a common label v, so M equals the constant $(-1)^{v^{\mathsf{T}}Bv}$ on the corresponding principal block. The normalized indicator vector x of the class has $|x^{\mathsf{T}}Mx| = \ell$. The last assertion is the second one for $R = R_d$; see Section 4 for the identification of the corank. □

Equality holds in (5) when all symbols equal 1: then $k = r - 1$ and $K_d = J - I$. For typical d the right side is about $r - \log_2 r$, far from the truth. Theorem 5.2 shows that the bound becomes sharp, up to rounding, once $2^{\,r-1-k}\log 2^{\,r-1-k} = o(r)$.

3 Spectral laws

We use the quotient sample of Section 2 throughout, with $H = K/\sqrt{r}$ and $m_p(H) = r^{-1}\operatorname{tr} H^p$. A common walk expansion first gives the empirical law; the same kernel then gives occupancy spikes and, through compressed traces, GOE edges. All estimates are uniform in the invertible symmetric form B.

3.1 The distribution of all eigenvalues

Proposition 3.1 (Walk comparison)**.** *For positive integers $p_1, \dots, p_t$ with $J = \sum_a p_a < r$ and $N \geq r$,*

$$\left| \mathbb{E} \prod_a m_{p_a}(H) - \mathbb{E}_{\mathrm{Wig}} \prod_a m_{p_a}(W_r/\sqrt{r}) \right| \leq J^J \frac{r}{N} + 2r^{J/2-t} \min\{2^{s-n}, 5\,2^{J-s}\}. \tag{6}$$

Proof. Expand the t closed walks. Let G be their full support and F their odd-multiplicity edges. Write v, c_G for the vertex and component counts of G, and v_F, e_F, c_F for those of F. The degrees of F are even. Its binary incidence matrix D_F has rank $v_F - c_F$ and satisfies $A_F = D_F D_F^{\mathsf{T}}$. Sylvester's rank inequality gives, for $d = \operatorname{rank} A_F$,

$$2v - J - 2c_G \leq 2v_F - e_F - 2c_F \leq d, \qquad v - J/2 - t - d/2 \leq c_G - t. \tag{7}$$

The first inequality counts the edges needed to connect the components of F and the remaining vertices: $J \geq e_F + 2(v - v_F + c_F - c_G)$. All edges outside F are used at least twice, and $c_G \leq t$. If $F \neq \varnothing$, then $d \geq 2$. Before full-rank conditioning, Proposition 2.2 bounds each equality pattern's normalized contribution by $r^{v-J/2-t}N^{-d/2} \leq (r/N)^{d/2} \leq r/N$. There are at most J^J patterns. Empty odd supports contribute one in both ensembles and cancel. The spanning correction, charged on at most r^J labelled walks, gives the second term of (6), including walks through v_r. □

An equality pattern of a closed walk records which of its vertices coincide, with vertices labelled in their order of first appearance. Consecutive vertices, including the last and first, must be distinct. For $j \geq 1$ and a pattern w of length $2j$, write $v(w)$ for its vertex count and $d(w)$ for the binary adjacency rank of its odd-multiplicity edge graph. Define

$$c_{2j}(\lambda) = \sum_{w:\, v(w)-j-1=d(w)/2} \lambda^{d(w)/2}, \qquad c_{2j+1}(\lambda) = 0, \quad c_0(\lambda) = 1. \tag{8}$$

Thus $c_2 = 1$, $c_4 = 2 + \lambda$ and $c_6 = 5 + 6\lambda + \lambda^2$.

Theorem 3.2 (Eigenvalue distributions at every binary rank)**.** *For each $0 \leq \lambda < \infty$, (8) is the moment sequence of a unique symmetric probability measure ν_λ. The measure ν_0 is the semicircle law on $[-2, 2]$; put $\nu_\infty = \delta_0$. This family is weakly continuous on $[0, \infty]$. With $s = r - k - 1$, $N = 2^s$ and $\mathcal{L}_r = r^{-1}\sum_i \delta_{\lambda_i(K)/\sqrt{r}}$, one has*

$$d_{\mathrm{BL}}(\mathcal{L}_r, \nu_{r/N}) \longrightarrow 0 \quad \textit{in probability}$$

for every corank sequence, uniformly in the invertible quotient form B. If $r/N \to \lambda < \infty$, every fixed empirical moment also converges in probability and expectation. At $r/N \to \infty$ only weak convergence to δ_0 is asserted: the second empirical moment remains $(r-1)/r$.

Proof. If $r/N \to 0$, (6) tends to zero at every fixed order: use 2^{s-n} for $s \le n/2$ and the local bound otherwise. One- and two-trace moments give the semicircle moments and vanishing variances; higher orders give uniform integrability. If $r/N \to \infty$, (1) leaves at most N eigenvalues outside $[-1,1]$, proving weak convergence to δ_0.

Now suppose $r/N \to \lambda \in (0,\infty)$, so $s = O(\log r)$. Use the walk notation and component-sensitive budget (7). A one-trace even pattern survives only when $v - j - 1 = d/2$; every other pattern loses at least $r^{-1/2}$. For a surviving nonempty F, let D_F be its incidence matrix and $a = \operatorname{rank} D_F = v_F - c_F$. Equality gives $d = 2a - e_F$, and hence

$$\dim(\ker D_F \cap \operatorname{im} D_F^{\mathsf{T}}) = a - d = e_F - a = \dim \ker D_F.$$

Thus the binary cycle space is contained in its orthogonal complement. Each cycle has even length, so F is bipartite. Sum the independent quotient vectors on one side of its bipartition first: the exact character expectation is $N^{-d/2}$ for every invertible B. This proves (8); odd lengths have no equality because d is even. In two traces, a pattern joining their supports has $c_G < 2$ and vanishes. The surviving disjoint patterns factor, so every fixed empirical moment has vanishing variance. The dependent final vector is included by the proper-subset argument; full-span conditioning adds only $O_J(r^{J/2-t}N2^{-r}) = o(1)$.

There are at most $(2j)^{2j}$ patterns and $d/2 \le j-1$ in a surviving one-trace pattern. Consequently

$$c_{2j}(\lambda) \le (2j)^{2j} \max(1, \lambda^{j-1}).$$

The same rank bound at higher fixed orders bounds the corresponding finite-sample moments on compact positive r/N ranges. Tightness, uniform integrability and Carleman's criterion give a unique limiting probability measure and convergence of every fixed moment. Existence for each prescribed $\lambda > 0$ follows, for example, along $r = \lfloor \lambda 2^s \rfloor$. All estimates are uniform in B.

The moment polynomials and the locally uniform bounds prove weak continuity at every finite positive λ. Their constant term counts the empty-odd-support patterns with $v = j+1$, namely tree walks traversing each edge twice. This is the Catalan number, giving continuity at zero and the semicircle law. For $\lambda > 1$, the rank bound also gives $\nu_\lambda(\{0\}) \ge 1 - 1/\lambda$: at most N/r of the finite-sample eigenvalues lie outside $[-r^{-1/2}, r^{-1/2}]$. Hence $\nu_\lambda \Rightarrow \delta_0$ as $\lambda \to \infty$. Every sequence r/N has a convergent subsequence in the compactified half-line. The preceding limits and continuity prove the moving-target bounded-Lipschitz assertion. At infinity there is no second-moment convergence, as the deterministic second moment shows. □

Corollary 3.3 (Sparse edge locations)**.** *If $s/(\log r \log\log r) \to \infty$, then the two extreme eigenvalues of $K/\sqrt{r}$ tend in probability to 2 and -2.*

Proof. In (6) take $J = 2m$ with $m = \lceil a \log r \rceil$. The Wigner bound $\mathbb{E}\operatorname{tr}(W_r/\sqrt{r})^{2m} \ll r4^m$ follows from [18, Theorem 3] by power-mean interpolation. The hypothesis pays $J \log J = O_a(\log r \log\log r)$ and the spanning correction is exponentially small. Markov's inequality, with a large depending on $\varepsilon > 0$, gives $\Pr(\|K\|/\sqrt{r} > 2 + \varepsilon) \to 0$. The bulk law supplies eigenvalues near both edges from inside. □

3.2 The limiting distributions

For $\lambda > 0$ let Π_λ be the Poisson law of mean λ. Let d_1, d_2 be free self-adjoint elements of a tracial noncommutative probability space $(\mathcal{A}, \tau)$ whose moments are those of Π_λ [13, Lectures 6 and 14]. We write

$$m_j(\Pi_\lambda \boxtimes \Pi_\lambda) = \tau\big((d_1 d_2)^j\big),$$

the moments of the free multiplicative convolution of Π_λ with itself. The theorem below shows in particular that they are the moments of a unique probability measure on $[0,\infty)$, which we denote by $\Pi_\lambda \boxtimes \Pi_\lambda$.

Theorem 3.4 (The measures ν_λ)**.** *For every $\lambda > 0$ and $j \geq 1$,*

$$c_{2j}(\lambda) = \lambda^{-(j+1)}\, m_j(\Pi_\lambda \boxtimes \Pi_\lambda). \tag{9}$$

Equivalently, if X has law ν_λ, then λX^2 has law

$$\lambda^{-1}\,\Pi_\lambda \boxtimes \Pi_\lambda + (1-\lambda^{-1})\delta_0.$$

Moreover

$$(\Pi_\lambda \boxtimes \Pi_\lambda)(\{0\}) = e^{-\lambda}, \qquad \nu_\lambda(\{0\}) = 1 - \frac{1-e^{-\lambda}}{\lambda}.$$

For $\lambda < 1$ the coefficient of δ_0 is negative; the atom of $\Pi_\lambda \boxtimes \Pi_\lambda$ compensates, since $e^{-\lambda} \geq 1-\lambda$.

Proof. Fix an invertible symmetric B and the matrix U of (1). For a partition π of the cyclically ordered positions $1,\dots,2j$ put

$$T_\pi = \sum_a \prod_{i=1}^{2j} (-1)^{a_i^{\mathsf{T}} B a_{i+1}},$$

where $a = (a_1,\dots,a_{2j})$ runs over the tuples in $\mathbb{F}_2^s$ that are constant on the blocks of π, and $a_{2j+1} = a_1$. The quotient of the $2j$-cycle by π is a closed walk on $|\pi|$ vertices, possibly with loops. Let $d(\pi)$ be the binary adjacency rank of its odd-multiplicity edges between distinct vertices. The sum T_π is a quadratic character sum in $|\pi|$ variables of $\mathbb{F}_2^s$ whose polar form has rank $s\,d(\pi)$; loops change the quadratic form but not its polar form. The differencing identity in the proof of Proposition 2.2 gives $|T_\pi| \leq N^{|\pi| - d(\pi)/2}$.

Let w be the walk obtained by deleting the ℓ loop steps. It has the same vertices and the same odd edges, and length $2j-\ell$. The budget (7) with one component gives $2|\pi| - (2j-\ell) - 2 \leq d(\pi)$, that is,

$$|\pi| - 1 - j - \tfrac{1}{2}d(\pi) \leq -\tfrac{1}{2}\ell. \tag{10}$$

Call π *surviving* if it has no loop and equality holds. The surviving partitions are exactly the equality patterns counted in (8), and for them the proof of Theorem 3.2 gives $T_\pi = N^{|\pi| - d(\pi)/2}$. Every other π has $N^{-1-j}|T_\pi| \leq N^{-1/2}$; without loops the bound is N^{-1}, because $d(\pi)$ is even.

A surviving partition respects parity. Let G be the support graph of the walk and F its odd edges. In the equality case every edge outside F is traversed exactly twice, and their number is $v - v_F + c_F - 1$ in the notation of (7). These edges therefore join the components of F and the remaining vertices as a tree, and add no cycle: G and F have the same cycle space. The proof of Theorem 3.2 shows that F is bipartite. Hence G is bipartite, the walk alternates between its two sides, and positions in one block of π have the same parity.

Two diagonal matrices. Let D_1, D_2 be independent $N \times N$ diagonal matrices with independent Π_λ entries. All cumulants of Π_λ equal λ, and joint cumulants of entries at different sites vanish. The moment–cumulant formula gives

$$\frac{1}{N}\mathbb{E}\,\mathrm{tr}(D_1U)^{2j} = N^{-1-j}\sum_\pi \lambda^{|\pi|}T_\pi, \qquad \frac{1}{N}\mathbb{E}\,\mathrm{tr}(D_1UD_2U)^{j} = N^{-1-j}\sum_{\pi\ \text{respecting parity}} \lambda^{|\pi|}T_\pi.$$

By the two preceding paragraphs both converge, as $N\to\infty$, to

$$\sum_{\pi\ \text{surviving}} \lambda^{|\pi|} = \lambda^{j+1}c_{2j}(\lambda),$$

since a surviving π has $|\pi| = j + 1 + d(\pi)/2$.

Freeness. We claim that D_1 and UD_2U are asymptotically free in expectation. Since $U^2 = I$, a polynomial in UD_2U is $Ug(D_2)U$. Let $f_1, g_1, \dots, f_m, g_m$ be polynomials with $\int f_i \, d\Pi_\lambda = \int g_i \, d\Pi_\lambda = 0$. Then

$$\frac{1}{N}\mathbb{E}\,\mathrm{tr}\prod_{i=1}^{m} f_i(D_1)\,U g_i(D_2)U = N^{-1-m}\sum_a \mathbb{E}\Big[\prod_i f_i(D_{1,a_{2i-1}})\Big]\mathbb{E}\Big[\prod_i g_i(D_{2,a_{2i}})\Big]\prod_{i=1}^{2m}\sqrt{N}\,U_{a_i a_{i+1}}.$$

The first expectation vanishes unless every site occurring among the odd positions occurs at least twice there, and similarly for the even positions. Such tuples have at most m distinct sites, so there are $O_m(N^m)$ of them. The expectations are bounded and $|\sqrt{N}\,U_{ab}| = 1$. The mean trace is therefore $O(N^{-1})$. For an odd alternating word, cycle the two end factors together and write their product as its centred part plus its mean. The centred part gives an even alternating word, and the mean gives a shorter odd word. Induction starts with one centred factor. Since D_1 and UD_2U each have the exact mean moments of Π_λ, their joint mean moments converge to those of free elements d_1, d_2 with law Π_λ. In particular

$$\frac{1}{N}\mathbb{E}\,\mathrm{tr}(D_1UD_2U)^j \longrightarrow \tau\big((d_1d_2)^j\big) = m_j(\Pi_\lambda \boxtimes \Pi_\lambda).$$

Comparison with the previous step proves (9).

An operator model and the atom. Realize d_1, d_2 as free positive operators affiliated with a finite von Neumann algebra with faithful normal trace τ, each with distribution Π_λ; for instance in the free product of two copies of $L^\infty(\Pi_\lambda)$. They have moments of all orders, so $A = \overline{d_2^{1/2}d_1^{1/2}}$ in the algebra of τ-measurable operators defines $T = A^*A$, with $\tau(T^j) = \tau((d_1d_2)^j)$; the moment formula for free elements extends from the bounded truncations $d_i\mathbf{1}_{\{d_i \le n\}}$ by Hölder's inequality. The distribution ρ of T is therefore a probability measure on $[0,\infty)$ with moments $m_j(\Pi_\lambda \boxtimes \Pi_\lambda) = \lambda^{j+1}c_{2j}(\lambda)$. These grow at most like $\lambda^{j+1}(2j)^{2j}\max(1,\lambda^{j-1})$ by the proof of Theorem 3.2. The Stieltjes Carleman sum $\sum_j m_j(\rho)^{-1/(2j)}$ diverges, so these moments determine ρ. This is the measure $\Pi_\lambda \boxtimes \Pi_\lambda$.

Let e_i be the kernel projection of d_i, so $\tau(e_i) = e^{-\lambda}$. The inverse square root of d_i on its support is bounded, since its nonzero spectrum lies in $\{1, 2, \dots\}$. If k is the kernel projection of T, then $Ak = 0$. Multiplying by that inverse for d_2 shows that the range of $d_1^{1/2}k$ lies in both e_2 and $1 - e_1$. The projections e_2 and $1 - e_1$ are free, with traces $e^{-\lambda}$ and $1 - e^{-\lambda}$. For free projections p, q the element pqp has an atom at 1 of mass $\max(0, \tau(p) + \tau(q) - 1)$ [23, Example 3.6.7], and the spectral projection of pqp at 1 is the meet $p \wedge q$. Here the two traces add up to exactly one, so the meet has trace zero and vanishes, the trace being faithful. Hence $d_1^{1/2}k = 0$, that is $k \le e_1$. Conversely $Te_1 = 0$. So $k = e_1$ and $\rho(\{0\}) = e^{-\lambda}$.

The law of the square. The measure $\lambda^{-1}\rho + (1 - \lambda^{-1})\delta_0$ has total mass one and mass $\lambda^{-1}e^{-\lambda} + 1 - \lambda^{-1} \ge 0$ at zero, so it is a probability measure. Its jth moment is $\lambda^j c_{2j}(\lambda)$ for $j \ge 1$, which is the jth moment of λX^2. Both are determined by their moments, hence equal. Evaluating at zero gives $\nu_\lambda(\{0\})$. □

Remark 3.5 (Classes of linked rows)**.** *The atom has a direct meaning. The matrix $K + \Delta$ has rank at most the number of distinct labels. For r independent uniform labels with $r/N \to \lambda$ this number has mean $N(1 - (1 - 1/N)^r)$ and variance at most its mean; the dependent last label and the spanning condition change it by at most one and by a bounded factor in probability. Thus the number of classes of linked rows, divided by r, tends in probability to $(1 - e^{-\lambda})/\lambda$. The theorem says that the rank deficiency accounts for the whole atom, in the sense that the proportion of eigenvalues of $K/\sqrt{r}$ in $[-\epsilon, \epsilon]$ beyond it tends to zero when first $r \to \infty$ and then $\epsilon \to 0$. The step proving freeness uses only $U^2 = I$ and $|U_{ab}| = N^{-1/2}$; the earlier steps use the quadratic form B.*

Proposition 3.6 (Generating function and size of the moments)**.** *Fix $\lambda > 0$ and write $G(w) = \sum_{n\geq 0} m_n(\Pi_\lambda) w^n$ for the moment series of the Poisson law.*

1. *There is a unique formal power series $w(z) = \lambda z + \cdots$ with*

$$w(z)^2\, G(w(z)) = z\,\{G(w(z)) - 1\},$$

and

$$1 + \sum_{j\geq 1} \lambda^{j+1} c_{2j}(\lambda)\, z^j = G(w(z)).$$

2. *For every $j \geq 1$,*

$$\frac{2}{\lambda}\, m_j(\Pi_\lambda) - \lambda^{j-1} \;\leq\; c_{2j}(\lambda) \;\leq\; \lambda^{-(j+1)}\, m_{2j}(\Pi_\lambda). \tag{11}$$

3. *The support of ν_λ is unbounded.*

Proof. For a law μ with all moments and nonzero mean put $\psi_\mu(w) = \sum_{n\geq 1} m_n(\mu) w^n$, let χ_μ be its inverse under composition, and $S_\mu(u) = \chi_\mu(u)(1+u)/u$. For free elements the S-transform of the product is the product of the S-transforms, as an identity of formal power series [13, Lecture 18]. With $\mu = \Pi_\lambda$ and $\rho = \Pi_\lambda \boxtimes \Pi_\lambda$ this reads $\chi_\rho(u) = \chi_\mu(u)^2(1+u)/u$. Put $u = \psi_\rho(z)$ and $w = \chi_\mu(u)$. Then $\psi_\mu(w) = u$ and $z = w^2(1+u)/u$, which is the displayed equation because $1 + \psi_\mu = G$. It determines the coefficients of w recursively from $w = \lambda z + \cdots$. Theorem 3.4 converts $1 + \psi_\rho(z) = G(w(z))$ into the second display.

For the lower bound in (11) return to the proof of Theorem 3.4. Let ω be any partition of the j odd positions, and let π consist of the blocks of ω and the j even positions as singletons. Summing over the label at an even position between two odd positions with labels a, a' gives $\sum_c (-1)^{c^{\mathsf T} B(a+a')} = N \mathbf{1}_{a=a'}$, since B is invertible. Going around the cycle forces all odd labels to coincide, so $T_\pi = N^{j+1}$ exactly, and $N^{-1-j}\lambda^{|\pi|} T_\pi = \lambda^{j+|\omega|}$. Hence π is surviving. The same holds with the roles of odd and even positions exchanged. The two families share only the partition into singletons, whose term is λ^{2j}. Every surviving term is positive, and $\sum_\omega \lambda^{|\omega|} = m_j(\Pi_\lambda)$. Therefore

$$\lambda^{j+1} c_{2j}(\lambda) \geq 2\lambda^j m_j(\Pi_\lambda) - \lambda^{2j}.$$

For the upper bound, Hölder's inequality for Schatten norms and the Cauchy–Schwarz inequality give

$$\frac{1}{N}\mathbb{E}\,\mathrm{tr}(D_1 U D_2 U)^j \leq \mathbb{E}\Big[\Big(\frac{1}{N}\,\mathrm{tr}\, D_1^{2j}\Big)^{1/2}\Big(\frac{1}{N}\,\mathrm{tr}\, D_2^{2j}\Big)^{1/2}\Big] \leq m_{2j}(\Pi_\lambda),$$

and the left side tends to $\lambda^{j+1} c_{2j}(\lambda)$.

By Jensen's inequality $m_j(\Pi_\lambda) \geq \lambda^j$, so the lower bound gives $c_{2j}(\lambda) \geq \lambda^{-1} m_j(\Pi_\lambda)$. Since Π_λ has unbounded support, $m_j(\Pi_\lambda)^{1/j} \to \infty$, and therefore $c_{2j}(\lambda)^{1/(2j)} \to \infty$. □

The factor 2 comes from the two families of partitions in the proof.

3.3 Large eigenvalues from repeated labels

Theorem 3.7 (Extreme eigenvalues and label counts)**.** *Suppose $N = 2^s \to \infty$ and $\lambda_r = r/N \geq (\log N)^{-\beta}$ for fixed $0 < \beta < 1$. Let $b_N > \lambda_r$ be the unique solution of*

$$b_N \log(b_N/\lambda_r) - b_N + \lambda_r = \log N.$$

For the actual maximum occupancy ℓ_r of Section 2,

$$\frac{\ell_r}{b_N} \longrightarrow 1, \qquad \frac{\lambda_{\max}(K)}{\sqrt{rb_N}} \longrightarrow 1, \qquad \frac{\lambda_{\min}(K)}{\sqrt{rb_N}} \longrightarrow -1 \quad \textit{in probability.}$$

The conclusions are uniform in the invertible symmetric quotient form B.

Proof. Use the kernel U of (1). We first replace the actual sample by r independent uniform rows, with occupancy matrix D and $\ell = \max_a D_a$; the source corrections are paid below. It suffices to locate the extremes of $D^{1/2}UD^{1/2}$.

We use the following consequence of Tropp's principal restriction bound [20, Theorem 1.1]. For a Bernoulli coordinate projector P of density p and $q \geq 2\log N$,

$$\left(\mathbb{E}\|PUP\|^q\right)^{1/q} \leq C\{qN^{-1/2} + \sqrt{pq} + p\} + N^{-1/2}. \tag{12}$$

Indeed the off-diagonal part of U has maximum entry $N^{-1/2}$, maximum column norm at most one and operator norm at most two. If $p \leq (\log N)^{-10}$, choosing $q = A\log N$ gives $\|PUP\| \leq C_A(\log N)^{-4}$ with failure at most N^{-A}.

Poisson occupancies. For now assume $(\log N)^{-\beta} \leq \lambda_r \leq (\log N)^2$ and let the D_a be independent Poisson variables of mean $\lambda = \lambda_r$. All operator events below have failure $O(N^{-3})$; constants may depend on β and can be fixed uniformly on this entire range of λ.

First suppose $\lambda \leq \sqrt{\log N}$. Set $t = \lceil C_0(\lambda + \log\log N)\rceil$ with fixed C_0 large, $H = \{a : D_a > t\}$, $L = H^c$. Poisson Chernoff gives $p := \Pr(a \in H) \leq (\log N)^{-10}$. With the stated failure bound,

$$\alpha := |H|/N \leq (\log N)^{-8}, \qquad \frac{\log N}{4(1+\beta)\log\log N} \leq \ell \leq \frac{10\log N}{\log\log N}, \qquad \|U_{HH}\| \leq C(\log N)^{-4}. \tag{13}$$

The cardinality bound is scalar Chernoff. For the lower load bound, monotonicity reduces to mean $(\log N)^{-\beta}$; the probability of one box having the displayed load is $N^{-1/4+o(1)}$. Independence then pays the failure. The upper Poisson tail is uniform for $\lambda \leq \sqrt{\log N}$. The restriction bound is (12). Conditional on H, the low occupancies are independent, bounded by t, with mean μ_t and variance at most 2λ. By Cauchy–Schwarz, $|\mu_t/\lambda - 1| \leq 2\sqrt{p/\lambda} = o(1)$ uniformly. For the variance, center first at the unconditioned mean: $\mathrm{Var}(D_a \mid D_a \leq t) \leq \lambda/\Pr(D_a \leq t)$. Put $u_a = P_H U e_a$, so $\|u_a\|^2 = \alpha$ exactly. Then

$$Q = P_H U D_L U P_H = \sum_{a\in L} D_a u_a u_a^{\mathsf{T}}, \qquad \mathbb{E}[Q \mid H] = \mu_t(I_H - U_{HH}^2).$$

The centered summands have norm at most $t\alpha$ and total variance at most $2\lambda\alpha$. Matrix Bernstein [21, Theorem 6.1(ii)], in both signs at $\delta = \lambda/\log N$, gives the conditional failure bound

$$2N\exp\left[-\frac{\delta^2}{2(2\lambda\alpha + t\alpha\delta/3)}\right] \leq 2N\exp\{-c_1(\log N)^{6-\beta}\}.$$

Here $t/\log N \to 0$ uniformly on this range. Thus $Q = \lambda(I_H + o(1))$ in operator norm. For $C_H = D_H^{1/2}U_{HL}D_L^{1/2}$ this yields

$$(1-o(1))\lambda D_H \preccurlyeq C_H C_H^{\mathsf{T}} \preccurlyeq (1+o(1))\lambda D_H.$$

The largest occupancy belongs to H eventually, so $\|C_H\| \sim \sqrt{\lambda\ell}$. The two diagonal blocks of $D^{1/2}UD^{1/2}$ have norms at most $\ell\|U_{HH}\|$ and t. Both are $o(\sqrt{\lambda\ell})$: the first costs at most $C(\log N)^{-4+(1+\beta)/2}/\sqrt{\log\log N}$ relatively, while

$$\frac{t^2}{\lambda\ell} \ll_\beta \frac{\log\log N}{\sqrt{\log N}} + \frac{(\log\log N)^3}{(\log N)^{1-\beta}} \longrightarrow 0.$$

The off-diagonal block matrix has extremes $\pm\|C_H\|$, proving both signed estimates in this range.

Now suppose $\sqrt{\log N} \le \lambda \le (\log N)^2$. Set $\delta = (C_2 \log\log N/\lambda)^{1/2} = o(1)$, and take $H = \{a : |D_a - \lambda| > \delta\lambda\}$. Scalar Chernoff with C_2 large again gives $p \le (\log N)^{-10}$. Outside an event of probability $O(N^{-3})$,

$$h := |H| \le N(\log N)^{-8}, \quad \|U_{HH}\| \le \varepsilon_N = C(\log N)^{-4}, \quad \ell \le C_3(\lambda + \log N).$$

Let D_0 agree with D on H and equal λI on L. Since $D_L/\lambda = I + O(\delta)$, and L is nonempty, $\ell \ge (1-\delta)\lambda$ and

$$\|D^{1/2}UD^{1/2} - D_0^{1/2}UD_0^{1/2}\| \le C\delta\sqrt{\lambda\ell}. \tag{14}$$

There is a useful exact block description of U. Write $A = U_{HH}$, $S = (I - A^2)^{1/2}$ and $W = S^{-1}U_{HL}$. Then $WW^{\mathsf{T}} = I_H$ and $U_{LL}W^{\mathsf{T}} = -W^{\mathsf{T}}A$. In the orthogonal decomposition $H \oplus \operatorname{ran} W^{\mathsf{T}} \oplus \ker W$,

$$U = \begin{pmatrix} A & S & 0 \\ S & -A & 0 \\ 0 & 0 & V \end{pmatrix}, \qquad V = V^{\mathsf{T}}, \quad V^2 = I, \quad \operatorname{tr} V = \operatorname{tr} U.$$

Replace the first two blocks by $\left(\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right)$ to obtain an involution U_0 with $\|U - U_0\| \le 2\varepsilon_N$. As $D_0 = \lambda I$ on the low space, the spectrum of $D_0^{1/2}U_0D_0^{1/2}$ consists exactly of $\pm\sqrt{\lambda D_a}$, $a \in H$, and the spectrum of λV. Both signs of V occur for large N, because $\dim V = N - 2h = N - o(N)$ and $|\operatorname{tr} V| \le \sqrt{N}$. Its extremes are consequently

$$\pm\max\{\lambda, \sqrt{\lambda \max_{a\in H} D_a}\} = \pm(1 + O(\delta))\sqrt{\lambda\ell}.$$

For $H = \varnothing$ interpret the inner maximum as zero. The remaining perturbation has norm at most $C\varepsilon_N\ell = o(\sqrt{\lambda\ell})$, since $\ell/\lambda \ll 1 + \log N/\lambda \ll \sqrt{\log N}$. Together with (14), Weyl proves both extreme estimates on this second range.

The sum of the Poisson occupancies has mean r and equals r with probability $\asymp r^{-1/2} \gg N^{-1/2}(\log N)^{-1}$. Conditioning therefore preserves the preceding high-probability operator statements and gives exactly the multinomial law of the iid row sample.

For $\lambda_r \ge (\log N)^2$, work directly with the multinomial sample. Chernoff with $\eta = (\log N/\lambda_r)^{1/4} = o(1)$ gives $\max_a |D_a/\lambda_r - 1| = o_{\Pr}(1)$. Hence $D^{1/2}UD^{1/2} = \lambda_r U + o_{\Pr}(\lambda_r)$ and $\ell \sim \lambda_r$. Since U has both signs, this proves both extreme estimates here as well. In all ranges the factor $\lambda_r^{-1/2}$ leaves the kernel extremes equal to $\pm(1 + o_{\Pr}(1))\sqrt{r\ell}$ before normalization.

The load bounds show that the largest occupancy tends to infinity throughout the stated range. Deleting the sign diagonal changes the norm by at most one. Replacing one iid row vector by the required sum changes its row and column by operator norm at most $2\sqrt{r}$, and changes the largest occupancy by at most one. These errors are $o(\sqrt{r\ell})$. Finally full-column-rank conditioning has failure at most $N2^{-r+1} = o(1)$, since $r \ge N/(\log N)^{\beta}$. All estimates are uniform in B. *Deterministic normalization.* First use the fixed iid row sample; its last-row correction changes the maximum by at most one. Put $F(x) = x\log(x/\lambda_r) - x + \lambda_r$, so $F(b_N) = \log N$. First suppose $\lambda_r \le C\log N$; then $b_N = O_C(\log N)$. For fixed $0 < \varepsilon < 1$, convexity gives

$$F((1+\varepsilon)b_N) \ge (1+\varepsilon)\log N.$$

Indeed $b_N F'(b_N) = b_N \log(b_N/\lambda_r) \ge F(b_N)$. A binomial Chernoff union bound proves the upper occupancy bound. If $(1-\varepsilon)b_N \le \lambda_r$, the lower bound follows from $\ell_r \ge \lambda_r$ for a fixed sample. Otherwise write $(1-\varepsilon)b_N = (1-a)\lambda_r + ab_N$, where $a \le 1 - \varepsilon$. Convexity now gives $F((1-\varepsilon)b_N) \le$

$(1-\varepsilon)\log N$. For $u = \lfloor (1-\varepsilon)b_N \rfloor$, Stirling's formula gives $N\Pr(D_1 = u) \geq N^{\varepsilon + o(1)}$. Moreover, exactly,

$$\frac{\Pr(D_1 = u, D_2 = u)}{\Pr(D_1 = u)^2} = \frac{(r-u)!^2}{r!(r-2u)!}\,\frac{(1-2/N)^{r-2u}}{(1-1/N)^{2r-2u}} = 1 + o(1).$$

The error is $O_C((\log N)^{2+\beta}/N)$ uniformly here. The second moment of the number of boxes of occupancy u therefore proves the lower bound.

When $\lambda_r/\log N \to \infty$, Chernoff gives $\max_a |D_a/\lambda_r - 1| \to 0$ in probability, whereas the defining equation implies $b_N/\lambda_r \to 1$. These two alternatives cover every subsequence. Changing the last row and conditioning on full span have the negligible costs already proved. □

In particular, if $\lambda_r/\log N \to \gamma \in (0,\infty)$, then $b_N/\lambda_r \to u_\gamma > 1$, where $u_\gamma \log u_\gamma - u_\gamma + 1 = 1/\gamma$. If $\lambda_r/\log N \to \infty$, then $b_N/\lambda_r \to 1$. For λ_r bounded above and away from zero, $b_N \sim \log N/\log\log N$. If $\lambda_r \sim C(\log N)^{-\alpha}$ with $C > 0$ and $0 < \alpha < 1$, then $b_N \sim \log N/((1+\alpha)\log\log N)$ (apply the theorem with any $\beta > \alpha$). The special cases follow directly from the defining equation.

Corollary 3.8 (Nearly equal label counts)**.** *If $N(s+1) = o(r)$, precisely N eigenvalues of K lie outside $[-1,1]$ with probability tending to one. Uniformly over their indices, they are $(1+o_{\Pr}(1))r/\sqrt{N}$ or $-(1+o_{\Pr}(1))r/\sqrt{N}$, with respective multiplicities*

$$(m_+, m_-) = \begin{cases} ((N+\sqrt{N})/2, (N-\sqrt{N})/2), & B \text{ alternating}, \\ (N/2, N/2), & B \text{ nonalternating}. \end{cases}$$

This includes bounded N and $s = 0$; the other $r - N$ eigenvalues stay in $[-1,1]$.

Proof. The first $n = r-1$ rows are iid before full-span conditioning, whose cost is at most four; the last row adds one occupancy. Thus Chernoff gives

$$\Pr\Big(\max_a |D_a - r/N| > un/N + 1\Big) \leq 8Ne^{-u^2 n/(3N)} \qquad (0 < u < 1).$$

With $u = (N(s+1)/r)^{1/4}$ this implies $D = (r/N)(I + o_{\Pr}(1))$, including positive occupancies. The shared kernel identity yields $\sqrt{N}D^{1/2}UD^{1/2} = (r/\sqrt{N})U + o_{\Pr}(r/\sqrt{N})$. Since $U^2 = I$, its trace gives the stated multiplicities. Weyl pays $\|\Delta\| = 1$; as $r/\sqrt{N} \to \infty$, all N nonzero eigenvalues stay outside $[-1,1]$, and the deterministic rank bound excludes any others. □

3.4 Fluctuations of the large eigenvalues

Corollary 3.8 locates the N eigenvalues outside $[-1,1]$ when $N(s+1) = o(r)$. We now describe their deviation from $\pm r/\sqrt{N}$. Let $P_\pm = (I \pm U)/2$ be the spectral projections of U, with ranges $V_\pm$ of dimensions $m_\pm$. Write $D = (r/N)(I+E)$, so that E is diagonal with entries $E_a = (D_a - r/N)/(r/N)$ and trace zero.

Theorem 3.9 (Bounded number of labels)**.** *Fix $s \geq 0$, $N = 2^s$ and a nondegenerate symmetric form B. Let g be a centred Gaussian vector in $\mathbb{R}^N$ with covariance $I - N^{-1}J$, where J is the all-ones matrix. Let $\theta_1 \geq \cdots \geq \theta_{m_+}$ be the eigenvalues of $P_+ \operatorname{diag}(g) P_+$ on V_+ and $\eta_1 \geq \cdots \geq \eta_{m_-}$ those of $P_- \operatorname{diag}(g) P_-$ on V_-. As $r \to \infty$, the eigenvalues $\xi_1 \geq \cdots \geq \xi_{m_+}$ of K above 1 and $\zeta_1 \leq \cdots \leq \zeta_{m_-}$ below -1 satisfy*

$$\Big(\frac{\xi_i - r/\sqrt{N}}{\sqrt{r}}\Big)_{i \leq m_+},\ \Big(\frac{\zeta_i + r/\sqrt{N}}{\sqrt{r}}\Big)_{i \leq m_-} \implies (\theta_i)_{i \leq m_+},\ (-\eta_i)_{i \leq m_-}$$

jointly in distribution. For $N = 1$ the limit of the positive shift is zero. For $N = 2$ the limit of $(\lambda_{\max}(K) - r/\sqrt{2})/\sqrt{r}$ is Gaussian with mean zero and variance $1/4$. For $m_+ \geq 2$ the limit of the largest eigenvalue is the largest eigenvalue of an $m_+ \times m_+$ Gaussian symmetric matrix and is not Gaussian in general.

Proof. For $N = 1$ one has $K = J - I$ and the claimed shift tends to zero. For fixed $N \geq 2$ the first $r - 1$ labels span $\mathbb{F}_2^s$ with probability tending to one, and the last label changes one occupancy by one. The multinomial central limit theorem gives $E = \sqrt{N/r}\,\operatorname{diag}(g_r)$ with $g_r \Rightarrow g$. By (1) the nonzero eigenvalues of $K + \Delta$ are those of $\sqrt{N}\,D^{1/2}UD^{1/2} = (r/\sqrt{N})A$, where

$$A = (I + E)^{1/2}U(I + E)^{1/2} = U + H, \qquad H = \tfrac{1}{2}(EU + UE) + O(\|E\|^2).$$

The eigenvalues of U are ± 1. For $\|H\| < 1/2$, the Schur complement with respect to $V_+ \oplus V_-$ shows that the m_+ largest eigenvalues of A differ from those of $I + P_+HP_+$ on V_+ by $O(\|H\|^2)$, and similarly at -1. Since $UP_\pm = \pm P_\pm$,

$$P_+HP_+ = P_+EP_+ + O(\|E\|^2), \qquad P_-HP_- = -P_-EP_- + O(\|E\|^2).$$

Multiplying by $r/\sqrt{N}$, the first-order terms are $\sqrt{r}$ times the eigenvalues of $P_\pm \operatorname{diag}(g_r)P_\pm$, and the remainder is $O_{\Pr}(\sqrt{N})$. Weyl's inequality for Δ costs at most one. Divide by $\sqrt{r}$ and use continuity of eigenvalues.

For $N = 2$ the only form is $B = (1)$. Then V_+ is spanned by $(\cos\frac{\pi}{8}, \sin\frac{\pi}{8})$ and $g = (g_0, -g_0)$ with $\operatorname{Var} g_0 = 1/2$, so $\theta_1 = g_0 \cos\frac{\pi}{4}$ has variance $1/4$. □

Proposition 3.10 (Growing number of labels)**.** *Suppose $N \to \infty$ and $N(s + 1) = o(r)$. With probability tending to one,*

$$\left(\tfrac{1}{2} - o(1)\right)\sqrt{2r \log N} \;\leq\; \lambda_{\max}(K) - \frac{r}{\sqrt{N}} \;\leq\; \left(1 + o(1)\right)\sqrt{2r \log N},$$

and the same holds for $-\lambda_{\min}(K) - r/\sqrt{N}$. The bounds are uniform in B.

Proof. The Chernoff bound in the proof of Corollary 3.8 reads $\Pr(\max_a |D_a - r/N| > un/N + 1) \leq 8Ne^{-u^2 n/(3N)}$ for $0 < u < 1$. Take $u = C\sqrt{N(s+1)/r}$ with $C^2 > 3 \log 2$. The right side is then at most $8 \cdot 2^s e^{-C^2(s+1)n/(3r)}$, which tends to zero because $N \to \infty$. Hence $\|E\| = O_{\Pr}(\sqrt{N(s+1)/r}) = o_{\Pr}(1)$. The occupancies of the first $r - 1$ labels are binomial before the spanning condition, which costs $o(1)$. Their mean is $(r - 1)/N = r/N + O(1)$, so we write $\mu = r/N$ for the deviation estimate. For a deviation $u = (1 \pm \epsilon)\sqrt{2\mu \log N}$ the binomial upper-tail logarithm is $-(1 + o(1))u^2/(2\mu)$. The upper bound is Chernoff's inequality. For the lower bound tilt the binomial to mean $\mu + u + 3\sqrt{\mu}$ and retain the interval of radius $2\sqrt{\mu}$ about that mean. Its tilted probability is bounded below, and undoing the tilt adds $O(u/\sqrt{\mu} + 1)$ to the logarithmic cost. The quadratic expansion of binomial relative entropy applies since $u = o(\mu)$ and $N \to \infty$. For the upper threshold apply a union bound. For the lower threshold the expected number of exceedances tends to infinity and its variance is at most its mean. Indeed, conditioning on one occupancy, the other has a binomial law with a number of trials decreasing in the conditioned value. The last label changes the maximum occupancy by at most one. Hence $\max_a E_a = (1 + o_{\Pr}(1))\sqrt{2N \log N/r}$.

As in the previous proof, $\lambda_{\max}(A) \leq 1 + \lambda_{\max}(P_+EP_+) + O(\|E\|^2)$, and a compression of E has largest eigenvalue at most $\max_a E_a$. For the lower bound let a maximize E_a and take $x = P_+e_a/\|P_+e_a\|$. Then $x^{\mathsf{T}}Ax = 1 + x^{\mathsf{T}}Ex + O(\|E\|^2)$. Here $\|P_+e_a\|^2 = (1 + U_{aa})/2$, $(P_+)^2_{ac} = 1/(4N)$ for $c \neq a$, and $\sum_c E_c = 0$, so

$$x^{\mathsf{T}}Ex = \frac{(P_+)^2_{aa}E_a - E_a/(4N)}{(1 + U_{aa})/2} = \left(\tfrac{1}{2} + O(N^{-1/2})\right)E_a.$$

Multiply by $r/\sqrt{N}$. The second-order term is $O_{\Pr}(\sqrt{N}(s+1))$, which is $o(\sqrt{r\log N})$, and Δ costs at most one.

At -1 the same site is used. Since $P_-HP_- = -P_-EP_- + O(\|E\|^2)$, one has $\lambda_{\min}(A) \geq -1 - \lambda_{\max}(P_-EP_-) - O(\|E\|^2) \geq -1 - \max_a E_a - O(\|E\|^2)$. For the other direction take $x = P_-e_a/\|P_-e_a\|$ with the same maximizing a. Now $\|P_-e_a\|^2 = (1-U_{aa})/2$ and $(P_-)^2_{ac} = 1/(4N)$ for $c \neq a$, so again $x^{\mathsf{T}}Ex = (\frac{1}{2} + O(N^{-1/2}))E_a$ and $x^{\mathsf{T}}Ax = -1 - x^{\mathsf{T}}Ex + O(\|E\|^2)$. Both extremes are thus governed by the most occupied label, as in Theorem 3.7. □

The proposition leaves a factor two. Treating P_+ as a free projection of trace one half suggests that the lower constant is the true one; we do not prove this.

3.5 Gaussian edge laws for observed blocks

The GOE edge process describes the limiting eigenvalues near one end of the spectrum of a real symmetric Gaussian matrix. Its largest point has law TW_1. Put $\mathfrak{c}(m) = m^{1/3}(\log m)^{3/2}$.

Theorem 3.11 (Joint limits for observed blocks)**.** *Let K have the law conditioned on binary rank $s = r - k - 1$. Fix a deterministic set $A \subseteq [r]$ and any positive-probability event $\mathcal{C}$ determined by B and the labels in A; arbitrary feasible conditions on the internal signs of A are allowed. For fixed b, take deterministic sets $I_1, \dots, I_b \subseteq [r] \setminus A$ with $|I_i \cap I_j| \leq 1$ for $i \neq j$. Put $m_j = |I_j|$ and assume $m_j \to \infty$ and $\sum_j \mathfrak{c}(m_j) = o(s)$. Conditional on $\mathcal{C}$, the $2b$ processes*

$$\Xi_j^{\pm} = \sum_{\xi \in \operatorname{Spec} K_{I_j,I_j}} \delta_{m_j^{1/6}(\pm\xi - 2\sqrt{m_j})}$$

converge jointly to independent GOE edge processes, and their largest points to independent TW_1 variables. The conclusions are uniform over feasible $\mathcal{C}$, however small its probability, and over the invertible quotient form B. Block sizes need not be comparable. Without the additional condition, the limits hold jointly with the global approximations in Theorems 3.2 and 3.7 whenever their hypotheses apply.

The proof in [7, Section S5] compares Chebyshev approximations to high trace powers using Proposition 2.2.

4 Quadratic fields

The preceding results concern a finite ensemble. We now use arithmetic equidistribution to transfer its GOE law to quadratic fields. The comparison retains its absolute error before any rank selection.

Let

$$\mathcal{D}(X) = \{1 < d \leq X : d \text{ is squarefree and } p \mid d \Rightarrow p \not\equiv 3 \pmod 4\}.$$

Use uniform measure on $\mathcal{D}(X)$ and on each stated conditioning event. "Arithmetic probability" refers to these spaces as $X \to \infty$. Set $K_1 = 0$; bounded sizes do not affect any limit.

Write $\mathrm{Cl}^+(d)$ for the narrow ideal class group of $\mathbb{Q}(\sqrt{d})$. The class group records which ideals fail to have a single generator; its narrow version also keeps track of signs. Its 4-rank counts the cyclic factors whose orders are divisible by 4. For $d \in \mathcal{D}(X)$ write its prime factors as $p_1 < \cdots < p_{r(d)}$, including $p_1 = 2$ when d is even. Define

$$K_d(i,i) = 0, \qquad K_d(i,j) = K_d(j,i) = \left(\frac{p_i}{p_j}\right) \quad (i < j).$$

The denominator is always odd; if $p_i \neq 2$, then both primes are 1 (mod 4), so this symmetric convention agrees with quadratic reciprocity. Its mod-2 Rédei lift $R_d = (c_{ij})$ is characterized by

$$(-1)^{c_{ij}} = K_d(i,j) \quad (i \neq j),$$

with diagonal chosen so that $R_d\mathbf{1} = 0$. Deleting the first row and column of R_d gives the Legendre matrix of [6, Section 5]: its diagonal entry in column j encodes the product of the symbols (p_i/p_j) over $i \neq j$, including $i = 1$. Conversely, that principal minor uniquely determines the full symmetric row-sum-zero matrix. Thus CKMP's matrix model and the full Laplacian convention used here are canonically bijective, also when $p_1 = 2$.

Put $\mathcal{S}_r(X) = \{d \in \mathcal{D}(X) : r(d) = r\}$.

4.1 Conditioning a comparison

The arithmetic examples provide different estimates for their error. We compare them using only the events being studied, such as events unchanged by relabelling rows. Total variation is the largest difference in probability assigned to the same event. The following rule controls what happens after restricting to a smaller family.

Proposition 4.1 (Normalized comparison)**.** *Let S, T be finite positive measures on the observable sigma-algebra, and put $p = T(\Omega) > 0$, $q = S(\Omega)$, and $b = |S - T|(\Omega)$, the total absolute error before normalization. If $b < p$, then $q > 0$ and*

$$|q - p| \leq b, \qquad d_{\mathrm{TV}}(S/q, T/p) \leq b/p.$$

In particular, for probability laws P, Q and a measurable weight $0 \leq h \leq 1$, use $S = hP$, $T = hQ$ and $b(h) = \int h\,|dP - dQ|$, with $p = Qh$. The following are sufficient error estimates. Here, for $f = dP/dQ$, define $D_{\chi^2}(P\|Q) = \int (f-1)^2\,dQ$.

input	*$b(h)$ bound*	*sufficient condition*
$d_{\mathrm{TV}}(P,Q) \leq \epsilon$	2ϵ	$\epsilon = o(p)$
$D_{\chi^2}(P\|Q) \leq D$	$\sqrt{pD}$	$D = o(p)$
$\lvert dP/dQ - 1\rvert \leq \epsilon$	ϵp	$\epsilon = o(1)$.

All statements hold after restriction to the sigma-algebra of the observables actually being transferred.

Proof. The mass bound is immediate. Since $q \geq p - b > 0$,

$$\left|\frac{p}{q}S - T\right|(\Omega) \leq |S - T|(\Omega) + |p - q| \leq 2b.$$

Division by $2p$ proves the assertion. For the three budgets use, respectively, the variation norm, Cauchy–Schwarz in the form $\int h|f-1|\,dQ \leq \sqrt{Qh\int h(f-1)^2 dQ}$, and the pointwise density bound. Thus every application uses its true normalizing mass, and a small conditioning event is charged only to the extent required by the available error. □

Measurable observation maps cannot increase total variation. Thus the same budget applies to a spectral measure, a collection of blocks or a graph-zeta statistic, once the corresponding reference law is known. The analytic input remains family-specific: CKMP supplies the absolute error below, while the companion proves a relative-pattern estimate on the uniformly relabelled rough-prime family.

4.2 Residue symbols on boxes

All CKMP numbering refers to the published version [6]. We first check its pattern estimate for the full matrix, including the dyadic row; the ensuing averaging must retain an absolute error.

Lemma 4.2 (Pattern comparison)**.** *In [6, Theorem 5.9], replace the printed hypotheses $k_2 > D_1$ and $\log t_{k_1} < t_1^{c_2}$ by $D_1 < s_{k_0+1}$ and $\log t_{k_1} < t_{k_0+1}^{c_2}$. The first k_0 coordinates may be singleton primes not* $3 \bmod 4$*, including* 2*, with Kronecker symbols at denominator* 2*. For fixed positive $\delta_1 < c_7 - c_8$ and $2\delta_2 < \frac{1}{4} - c_2c_3 - 3c_4 - c_5$, the normalized full-pattern L^1 error is $O(k_2^{-\delta_1} + t_{k_0+1}^{-\delta_2})$, once D_1 and k_2 are sufficiently large independently.*

On every acceptable arithmetic box B, these parameters can be chosen with $k_2 = \lfloor r/4 \rfloor$ and $D_1 = \exp((\frac{1}{2}\log\log X)^{c_9})$. Consequently every simultaneous-permutation-invariant event $\mathcal{A}_r$ satisfies

$$\left| \Pr_{d \in B}(K_d \in \mathcal{A}_r) - \Pr(W_r \in \mathcal{A}_r) \right| \le \varepsilon_B, \qquad \varepsilon_B \ll k_2^{-\delta_1} + t_{k_0+1}^{-\delta_2}.$$

The same bound holds for invariant observables of modulus at most one.

Proof. On pp. 27–28 of [6], the normalized variance and small-fibre terms are at most

$$k_1 \left\{ (2^{|P|+k_0+1}/k_2)^{1/2} + k_2^{-c_7} \right\} \ll k_2^{-(c_7-c_8)},$$

using $k_1 < k_2^{c_8}$ and $2^{|P|+k_0} < k_2^{1-2c_7}$. This combinatorial calculation permits the fixed prime 2. Moving the singleton primes into $\widetilde{P}$, each surviving fibre has size at least

$$2^{-|\widetilde{P}|} k_2^{-c_7} |X_i| \ge 2^{-k_0} e^i t_i (\log t_i)^{-c_3} \ge e^{i-k_0} t_i (\log t_i)^{-c_3}.$$

Reindexing by $i' = i - k_0$ gives Proposition 5.7's density condition; its other hypotheses become the stated first endpoint and upper cutoff, $|\widetilde{P}| \le \log t_i - i'$, and spacing indices $j_i - k_0$. Definition 5.3 allows $2 \in \widetilde{P}$, while all remaining coordinates are odd. The allowed prime products, hence the Siegel-less condition, are unchanged. Proposition 5.7 bounds the last term by $O(t_{k_0+1}^{-\delta_2})$; no step compares k_2 with D_1.

For completeness, the acceptable boxes meet these hypotheses as follows. Put $\mu = \frac{1}{2}\log\log X$ and take $k_1 = \lfloor (c_2/4)\log D_1 \rfloor$. Choose $c_9 < c_8 < c_7 < 1/2$ with $2c_9 \log 2 < 1 - 2c_7$, then $c_3 > 5 + c_7/c_9$, $c_2 > 0$ small, $c_1 = c_2/4$, $c_4 = c_2/10$, $c_5 = c_2$, $c_6 > 3$, and $c_{10} > 0$ sufficiently small. Thus $c_2c_3 + 3c_4 + c_5 < 1/4$. The full singleton cutoff of Definition 5.12 has $D_1 < s_{k_0+1}$ and $k_0 \le 2c_9 \log\mu + O(1)$. Regularity gives $\log t_{k_1}, \log t_{k_1+1} = D_1^{c_2/2+o(1)}$ and $\log t_{k_0+1} = \mu^{O(1)}$, proving both cutoff separations. For $\ell_i = \log\log t_i$ it gives $\ell_i \ge \max(c_9 \log\mu, i)$ below $r/3$ and $\ell_i \ge r/2$ above it. Combined with the interval prime number theorem,

$$|X_i| \gg \frac{t_i e^{k_0 - i}}{\log D_1 \log t_i}, \qquad (\log t_i)^{c_3 - 1} \gg e^{2i - k_0} \mu^{c_7 + c_9},$$

this proves the density bound. For $j_i = 1 + i + \lfloor c_5 \log t_i \rfloor$, $j_{k_0+1} > k_1$; if $j_i \le r$, then $\ell_i = O(\log\mu)$, $j_i \gg \log\mu$, and regularity or monotonicity gives $\ell_{j_i} > c_6 \ell_i$, as required for spacing.

Apply the pattern estimate with all pairs $\mathcal{M} = \{(i,j) : i < j\}$ and no auxiliary constraints. It gives

$$\sum_{a : \mathcal{M} \to \{\pm 1\}} \left| 2^{-|\mathcal{M}|} - \frac{1}{k_2! |B|} \sum_{\sigma \in \mathcal{P}(k_2)} |B(\sigma, a)| \right| \le \varepsilon_B.$$

Summing over the event gives the claimed comparison: permutation invariance makes its count independent of σ. The same L^1 inequality applies to bounded invariant observables. □

Remark 4.3 (The dyadic row)**.** *For odd $p \equiv 1 \bmod 4$, $(p/2) = (2/p)$. Thus the dyadic row can be permuted, and the principal-minor bijection above gives the full row-sum-zero model also for even d.*

4.3 Arithmetic averaging and rank selection

Proposition 4.4 (Assembly with a varying density)**.** *Let r lie in the window of [6, Equation (4.3)], and let W be its good set from the proof of Theorem 5.16; spacing, regularity and the Siegel-less condition make every box meeting W acceptable. Suppose $V_r(N) \subseteq \mathcal{S}_r(N)$ and $0 \le \delta_r \le 1$ satisfy, on every box meeting W,*

$$\big||V_r(N) \cap \mathcal{B}| - \delta_r|\mathcal{B}|\big| \le \epsilon_N|\mathcal{B}|. \tag{15}$$

Writing $\Phi_r(N) = |\mathcal{S}_r(N)|$, one has uniformly in δ_r

$$\big||V_r(N)| - \delta_r\Phi_r(N)\big| \le \epsilon_N\Phi_r(N) + |\mathcal{S}_r(N) \setminus W| + \Phi_r(N) - \Phi_r\big(N(1 - C/\log D_1)\big).$$

Proof. Partition $\mathcal{S}_r(N)$ into S_k according to the number of prime factors below D_1; put $V_k = V_r(N) \cap S_k$, $\Phi_k = |S_k|$. On CKMP's parameter set T_k take $d\nu_k = dp_1 \cdots dp_k\, ds_{k+1} \cdots ds_r/(s_{k+1} \cdots s_r)$, where dp_i counts the allowed primes, including unit mass at 2, and ds_i is Lebesgue measure. Set

$$I_k(Y) = \int_{T_k} |Y \cap \mathcal{B}(\mathbf{t})|\, d\nu_k, \qquad \omega_{r,k} = \prod_{i=k+1}^{r} \log\left(1 + \frac{1}{e^{i-k}\log D_1}\right).$$

Each integer has incidence weight at most $\omega_{r,k}$, with equality on $G_k = W \cap S_k \cap \{(5.5) \text{ holds}\}$ [6, p. 31]. For $E_k = |S_k \setminus G_k|$,

$$0 \le \omega_{r,k}|Y| - I_k(Y) \le \omega_{r,k}E_k \quad (Y \subseteq S_k), \tag{16}$$

and the boundary estimate following (5.5) gives

$$\sum_k E_k \le |\mathcal{S}_r(N) \setminus W| + \Phi_r(N) - \Phi_r\big(N(1 - C/\log D_1)\big). \tag{17}$$

Integration of (15) yields

$$|I_k(V_k) - \delta_r I_k(S_k)| \le \epsilon_N I_k(S_k) \le \epsilon_N\omega_{r,k}\Phi_k. \tag{18}$$

For $\omega = \omega_{r,k}$ the deficits $a_k = \omega|V_k| - I_k(V_k)$ and $b_k = \omega\Phi_k - I_k(S_k)$ satisfy $0 \le a_k \le b_k \le \omega E_k$. Hence $|a_k - \delta_r b_k| \le \max(\delta_r, 1 - \delta_r)\omega E_k$, so

$$\big||V_k| - \delta_r\Phi_k\big| \le \epsilon_N\Phi_k + \max(\delta_r, 1 - \delta_r)E_k. \tag{19}$$

Sum over k and use (17). □

Put $T_X = \log\log X$, $\Phi(X) = |\mathcal{D}(X)|$, $w_{X,r} = |\mathcal{S}_r(X)|/\Phi(X)$, and $\mathcal{R}_X = \{r : |r - T_X/2| < (T_X/2)^{2/3}\}$. Let μ_X be the arithmetic law of $(r(d), K_d)$ and let $\nu_X = \sum_{r\ge1} w_{X,r}\mathcal{L}(r, W_r)$, using the zero-matrix convention at size one. Write $\|\mu - \nu\|_{\mathrm{inv}}$ for the supremum event discrepancy over simultaneous-permutation-invariant sections at each size.

Proposition 4.5 (Quantitative comparison)**.** *For absolute $C, \rho > 0$,*

$$\|\mu_X - \nu_X\|_{\mathrm{inv}} \le CT_X^{-\rho}. \tag{20}$$

The invariant events may depend on both X and r.

Proof. CKMP's Theorems 4.1, 5.14 and the above parameters give uniform box error $O(T_X^{-\rho_0})$ and discarded counts G_r with $\sum_{r\in\mathcal{R}_X} G_r \ll \Phi(X)T_X^{-\rho_0}$, for some $\rho_0 > 0$. These estimates concern full sign patterns and hence are independent of the chosen event. Apply the two preceding results with $\delta_r = \Pr(W_r \in \mathcal{A}_{r,X})$. The discrepancy in size r is at most $O(T_X^{-\rho_0}\Phi_r(X) + G_r + B_r)$, where $B_r = \Phi_r(X) - \Phi_r(X(1 - C_1/\log D_1))$. The endpoint estimate on p. 32 gives $\sum_r B_r \ll \Phi(X)/\log D_1$, and Equation (4.2) gives $\eta_X := \sum_{r\notin\mathcal{R}_X} w_{X,r} \ll e^{-cT_X^{1/3}}$. Sum these bounds and take $\rho \le \min(\rho_0, c_9)$. No division by $\Phi_r(X)$ is used. □

Lemma 4.6 (Symmetric-corank probabilities)**.** *For a uniform symmetric binary $n \times n$ matrix, its corank probabilities satisfy, uniformly for $0 \le k \le n$, $a_- 2^{-k(k+1)/2} \le p_{n,k} \le a_+ 2^{-k(k+1)/2}$ for absolute $a_\pm > 0$. In particular $\Pr(\kappa \ge K) \ll 2^{-K(K+1)/2}$, and each fixed k has a positive limiting probability.*

Proof. Counting radicals and invertible quotients gives, for $m = n - k$,

$$p_{n,k} = 2^{-k(k+1)/2} \prod_{j=1}^{\lceil m/2\rceil} (1 - 2^{1-2j}) \frac{\prod_{j=m+1}^{n}(1 - 2^{-j})}{\prod_{j=1}^{k}(1 - 2^{-j})}$$

[6, proof of Theorem 5.16]. Empty products equal one; the corresponding infinite products are positive. The bounds, tail sum and fixed-k limits follow. □

Corollary 4.7 (Event transfer)**.** *If invariant events $\mathcal{A}_r$ have Wigner probabilities tending to L, their arithmetic probabilities tend to L; the same holds for uniformly bounded invariant observables. For fixed k, the corresponding conditional assertion holds whenever the Wigner limit given $\kappa(R) = k$ exists, with arithmetic conditioning $\mathrm{rk}_4\,\mathrm{Cl}^+(d) = k$.*

Proof. Apply the comparison with $r \asymp T_X$ off mass $\eta_X = o(1)$. Level sets give bounded observables; fixed-corank conditioning divides by a positive limiting mass, using $\mathrm{rk}_4\,\mathrm{Cl}^+(d) = \kappa(R_d)$ also for even d [6, proof of Theorem 5.16]. □

Theorem 4.8 (Rank-weighted arithmetic universality)**.** *For arbitrary $h_X(r,k) \in [0,1]$, set*

$$p_X = \sum_{r\ge 1} w_{X,r} \sum_{k=0}^{r-1} h_X(r,k) p_{r-1,k}.$$

If $T_X^\rho p_X \to \infty$, the probability law on $\mathcal{D}(X)$ proportional to $h_X(r(d), \mathrm{rk}_4\,\mathrm{Cl}^+(d))$ has the joint independent GOE edge-process and independent TW_1 extreme-value limits of Theorem 3.11, with its $r^{1/6}$ scaling and centering at $2\sqrt{r}$. For any bounded continuous test F of either pair and $|a_X(r,k)| \le 1$,

$$\mathbb{E}_h[a_X F] - \mathbb{E}_h[a_X]\mathbb{E}F_\infty \longrightarrow 0.$$

Proof. Put $h = h_X(r, \kappa(R))$ and $\epsilon_X = CT_X^{-\rho}$. Layer cake gives $|\mu_X(h) - p_X| \le \epsilon_X$. Proposition 4.1, on the invariant algebra, therefore makes the true normalizer positive and gives

$$\|\mu_X^h - \nu_X^h\|_{\mathrm{inv}} \le \frac{2\epsilon_X}{p_X} = o(1). \tag{21}$$

Choose $K_X = \lceil \sqrt{2(\rho+2)\log T_X/\log 2}\rceil$. The previous lemma bounds the ν_X^h-mass of $r \notin \mathcal{R}_X$ or $k > K_X$ by

$$O\left(\frac{e^{-cT_X^{1/3}} + T_X^{-\rho-2}}{p_X}\right) = o(1).$$

For the remaining indices $r \asymp T_X$ and $k = O(\sqrt{\log T_X})$, Theorem 3.11 is uniform: a violating subsequence would satisfy its quotient-growth hypothesis. Average with the exact posterior weights $w_{X,r}h_X(r,k)p_{r-1,k}/p_X$, then apply (21) and the Rédei identity. The argument remains valid after multiplication by bounded a_X, giving the asserted decorrelation. □

Corollary 4.9 (Arithmetic spectral laws)**.** *Unconditionally, $K_d/\sqrt{r(d)}$ has limiting semicircle measure and norm tending to 2 in arithmetic probability, and its two edge processes and extremes have the independent GOE and* TW_1 *limits above. The edge conclusions also hold after conditioning on* $\operatorname{rk}_4 \mathrm{Cl}^+(d) = k_X(r(d))$ *for any nonnegative integer profile with*

$$\max_{r \in \mathcal{R}_X} k_X(r)(k_X(r)+1) \le c_0 \log\log\log X,$$

where $c_0 > 0$ is absolute. This includes every fixed 4-rank.

Proof. For the unconditioned empirical measure and norm, apply event transfer to the Wigner laws [5]; take $h_X \equiv 1$ for the edges. Choose $c_0 < 2\rho/\log 2$ and $\alpha = c_0 \log 2/2 < \rho$. The corank-mass bound gives $p_X \gg (1-\eta_X)T_X^{-\alpha}$ for $h_X(r,k) = \mathbf{1}_{\{k=k_X(r)\}}$. The weighted theorem applies. □

Corollary 4.10 (The edge is independent of a bounded 4-rank)**.** *Let $\alpha_k = \lim_n p_{n,k}$ be the limiting corank probabilities of Lemma 4.6, and let F_1 be the distribution function of* TW_1*. For every fixed $k \ge 0$ and real t, t', the proportion of $d \in \mathcal{D}(X)$ with*

$$\operatorname{rk}_4 \mathrm{Cl}^+(d) = k, \qquad \frac{\lambda_{\max}(K_d) - 2\sqrt{r}}{r^{-1/6}} \le t, \qquad \frac{-\lambda_{\min}(K_d) - 2\sqrt{r}}{r^{-1/6}} \le t', \qquad r = r(d),$$

tends to $\alpha_k F_1(t)F_1(t')$ as $X \to \infty$. In particular, among the discriminants whose largest Legendre eigenvalue lies in a fixed Tracy–Widom window of positive limiting probability, the 4-rank has the same limiting distribution as in $\mathcal{D}(X)$.

Proof. The event $\operatorname{rk}_4 \mathrm{Cl}^+(d) = k$ has limiting density $\alpha_k > 0$ by Corollary 4.7 and Lemma 4.6. Conditionally on it, Corollary 4.9 with the constant profile $k_X \equiv k$ gives independent TW_1 limits for the two normalized extremes. Multiply. The last assertion is Bayes' rule. □

5 Rough discriminants at every 4-rank

To condition on ranks whose model probabilities lie below the absolute error in (20), we use a relative comparison.

Put $L = \log\log X$ and $e = \binom{r}{2}$. Fix $0 < c < C$ and take an integer $cL \le r \le CL$. This includes $|r - L/2| \le L^{2/3}$. Set $Y_r = 31L + 12r\log 2$. For fixed $B > 1$ assume

$$\log y \le L^B, \qquad \frac{\log y - Y_r}{\log L} \longrightarrow \infty. \tag{22}$$

Let $\mathcal{D}^{\mathrm{rough}}(X; y)$ be the set of squarefree $d \le X$ with exactly r prime factors, all exceeding y and congruent to 1 modulo 4, and let a_X be its cardinality. In the typical window, $y = (\log X)^{64}$ is admissible. Let $\widetilde{\mu}_{X,y}$ be the law of K_d on this family after independently permuting the prime labels uniformly, and let U_r be the law of W_r.

Theorem 5.1 (Relative pattern comparison [7, Theorem S1.1])**.** *Under* (22)*, uniformly in the indicated range of r,*

$$\epsilon_{X,y} := \max_{\tau} |2^e\, \widetilde{\mu}_{X,y}(\tau) - 1| \longrightarrow 0,$$

the maximum being over all symmetric sign patterns τ with zero diagonal. Consequently, for every event C with $p = U_r(C) > 0$,

$$\widetilde{\mu}_{X,y}(C) = p\{1 + O(\epsilon_{X,y})\}, \qquad d_{\mathrm{TV}}(\widetilde{\mu}_{X,y}(\cdot \mid C), U_r(\cdot \mid C)) \le \frac{\epsilon_{X,y}}{1 - \epsilon_{X,y}},$$

uniformly in C, however small p is.

The proof in [7] expands the indicator of a pattern into products of Legendre symbols over independent prime variables, bounds them with the bilinear estimate of Friedlander and Iwaniec [9], and removes the cutoff $d \le X$ and repeated primes without a factor equal to the number of patterns. The consequence for events is Proposition 4.1 with a pointwise density bound.

For $N = 2^s$ and $\lambda_r = r/N$ let $b_N > \lambda_r$ be the solution of $b_N \log(b_N/\lambda_r) - b_N + \lambda_r = \log N$, as in Theorem 3.7.

Theorem 5.2 (The Legendre spectrum at every 4-rank)**.** *Assume* (22). *Let $k = k_X$ be integers with $0 \le k_X \le r-1$, put $s = r - 1 - k_X$ and $N = 2^s$, and let d be uniform among the elements of $\mathcal{D}^{\mathrm{rough}}(X; y)$ with $\mathrm{rk}_4\,\mathrm{Cl}^+(d) = k_X$.*

1. *The number of such d is $a_X\, p_{r-1,k_X}\{1+o(1)\}$, where $p_{r-1,k} \asymp 2^{-k(k+1)/2}$ is the corank probability of Lemma 4.6. It tends to infinity.*
2. *With $\mathcal{L}_d = r^{-1}\sum_i \delta_{\lambda_i(K_d)/\sqrt{r}}$, one has $d_{\mathrm{BL}}(\mathcal{L}_d, \nu_{r/N}) \to 0$ in probability.*
3. *If $s/\{r^{1/3}(\log r)^{3/2}\} \to \infty$, the two point processes $\sum_i \delta_{r^{1/6}(\pm\lambda_i(K_d) - 2\sqrt{r})}$ converge jointly to independent GOE edge processes. In particular $r^{1/6}(\lambda_{\max}(K_d) - 2\sqrt{r})$ and $r^{1/6}(-\lambda_{\min}(K_d) - 2\sqrt{r})$ converge to independent TW_1 variables.*
4. *If $s/(\log r \log\log r) \to \infty$, then $\lambda_{\max}(K_d)/\sqrt{r} \to 2$ and $\lambda_{\min}(K_d)/\sqrt{r} \to -2$ in probability.*
5. *If $N \to \infty$ and $r/N \ge (\log N)^{-\beta}$ for a fixed $0 < \beta < 1$, then*

$$\frac{\lambda_{\max}(K_d)}{\sqrt{r b_N}} \longrightarrow 1, \qquad \frac{\lambda_{\min}(K_d)}{\sqrt{r b_N}} \longrightarrow -1 \quad \text{in probability.}$$

6. *If $N(s+1) = o(r)$, with N bounded or not, then with probability tending to one exactly N eigenvalues of K_d lie outside $[-1,1]$, each of them is $\pm(1+o(1))\, r/\sqrt{N}$, and*

$$\mathrm{rk}_4\,\mathrm{Cl}^+(d) = r - 1 - \log_2 \#\{i : |\lambda_i(K_d)| > 1\} = r - 1 - \mathrm{round}\Big(2\log_2 \frac{r}{\lambda_{\max}(K_d)}\Big).$$

7. *If $N = 1$, the normalized shift of the largest eigenvalue tends to zero. If $N \ge 2$ is fixed, $(\lambda_{\max}(K_d) - r/\sqrt{N})/\sqrt{r}$ converges in distribution to the mixture, over a uniform invertible symmetric B, of the laws of θ_1 in Theorem 3.9. If $N \to \infty$ and $N(s+1) = o(r)$, then $\lambda_{\max}(K_d) - r/\sqrt{N}$ lies between $(\frac{1}{2} - o(1))\sqrt{2r\log N}$ and $(1+o(1))\sqrt{2r\log N}$ with probability tending to one.*

Proof. For these discriminants $\mathrm{rk}_4\,\mathrm{Cl}^+(d)$ is the excess corank κ of the binary Laplacian of K_d, as in Section 4. The event $C = \{\kappa = k_X\}$ is invariant under simultaneous permutation of rows and columns, and $U_r(C) = p_{r-1,k_X} > 0$. Theorem 5.1 gives the count in the first assertion. It tends to infinity because $a_X \ge X\exp\{-O_{B,c,C}(L)\}$ by the population estimate in [7, Section S3], while $p_{r-1,k} \gg 2^{-r^2} = \exp\{-O(L^2)\}$.

The same theorem bounds the total variation distance between the conditional law of the relabelled K_d and $U_r(\cdot \mid \kappa = k_X)$ by $o(1)$. The latter is the law of the sign matrix of a uniform symmetric binary matrix with zero row sums and corank k_X, the model of Section 2. The spectrum does not depend on the labelling. Each remaining assertion says either that a permutation-invariant event has probability tending to one or that a function of the spectrum converges in distribution. Both kinds of statement survive an error $o(1)$ in total variation, so it suffices to quote the model. The second assertion is Theorem 3.2, which holds for every corank sequence. The third is Theorem 3.11 with $A = \varnothing$, $b = 1$ and $I_1 = [r]$. The fourth and fifth are Corollary 3.3 and Theorem 3.7. The last one is Theorem 3.9, averaged over the uniform form B of Lemma 2.1, and Proposition 3.10.

For the sixth, Corollary 3.8 gives exactly N eigenvalues outside $[-1, 1]$, all equal to $\pm(1 + o(1))\, r/\sqrt{N}$, and at least one of them is positive since $m_+ \geq 1$. Hence $\log_2 \#\{i : |\lambda_i(K_d)| > 1\} = s$ and $2\log_2(r/\lambda_{\max}(K_d)) = s + o(1)$. □

Remark 5.3 (Intermediate ranks)**.** *When $s = \log_2 r + O(1)$, the two extremes have magnitude of order $\sqrt{r \log r / \log\log r}$. The sparse edge location is known throughout the range in item 4. Its fluctuation law remains open where that range extends beyond item 3.*

6 Small prime factors, under the Riemann hypothesis for real characters

Theorem 5.2 excludes prime factors below y. Two small primes have a fixed symbol, and for a fixed small prime p and a prime q of size about y there is no second variable to average over. A comparison with relative error then needs $\sum_q (q/p)$ over a short range of primes to be smaller than the number of terms by a power of $\log X$, although q is only of size $\exp((\log\log X)^{O(1)})$. This is a power saving in the prime number theorem for progressions. We assume it in the following form.

Hypothesis GRH_2**.** *For every real primitive Dirichlet character χ, the function $L(s, \chi)$ has no zero with real part larger than $1/2$.*

Under GRH_2, for a real nonprincipal character χ modulo q, real t and $z \geq 2$,

$$\sum_{p \leq z} \chi(p) p^{-it} \ll \sqrt{z}\, \log^2(qz(|t| + 2)), \tag{23}$$

Apply the explicit formula under GRH to the primitive inducing character in $L(s + it, \chi)$ [11, Theorem 5.15]. Its analytic conductor is $O(q(|t| + 2))$, and its logarithmic derivative has coefficients $\Lambda(n)\chi(n)n^{-it}$ at the unramified primes. Removing prime powers, using partial summation gives the prime sum for the primitive character. The imprimitive character omits at most $\omega(q)$ primes, which gives (23). The real primitive case used here is covered by Alpöge's Dirichlet critical-line theorem [1, 2]. We state the input explicitly to identify the dependence of this section.

6.1 The comparison

Keep $L = \log\log X$, $cL \leq r \leq CL$, $e = \binom{r}{2}$ and the cutoff (22). In particular r can lie in the typical window $|r - L/2| \leq L^{2/3}$. Fix $J \geq 1$ and let $h = h(X)$ be integers with $2^h \leq L^J$. Let $\mathcal{D}^{\mathrm{sm}}(X; y, h)$ be the set of squarefree $d \leq X$, all of whose prime factors are congruent to 1 modulo 4, with exactly r prime factors exceeding y and at most h prime factors not exceeding y. Write $d = ab$, where $a = a(d)$ is the product of the prime factors not exceeding y, and put $g = \omega(a)$. Order the prime

factors of a by size and label those of b by an independent uniform permutation. Then

$$K_d = \begin{pmatrix} K_a & C_d \\ C_d^{\mathsf{T}} & K_b \end{pmatrix},$$

where K_a is determined by a, the $g \times r$ matrix C_d holds the symbols between small and large prime factors, and K_b is the Legendre matrix of the large ones. We fix the small-prime product a when making the relative comparison.

Theorem 6.1 (Relative comparison with small prime factors)**.** *Assume* GRH_2 *and* (22)*. There is* $\epsilon_X^{\mathrm{sm}} \to 0$ *such that, for every squarefree* a *with at most* h *prime factors, all at most* y *and congruent to* 1 *modulo* 4*, and for* d *uniform in* $\mathcal{D}^{\mathrm{sm}}(X; y, h)$ *conditioned on* $a(d) = a$*,*

$$\max_{\sigma,\tau} \Big| 2^{\,gr+e} \Pr(C_d = \sigma,\ K_b = \tau \mid a(d) = a) - 1 \Big| \le \epsilon_X^{\mathrm{sm}}.$$

The maximum is over all $\sigma \in \{\pm 1\}^{g \times r}$ *and all symmetric sign patterns* τ *of size* r *with zero diagonal. The error is uniform in* a *and* r *in the stated ranges.*

Proof. We follow the proof of [7, Theorem S1.1] with X/a in place of X, and indicate the changes. Since $\log a \le h \log y = o(\log X)$, the parameter $L_a = \log\log(X/a)$ satisfies $L_a = L + o(1)$ uniformly. The population estimates of [7, Section S3] apply with $r \asymp L_a$ and the same y: the lower cutoff changes by $31(L_a - L) = o(1)$ and the upper bound still holds with $B + 1$. As there, each rough coordinate p_i is uniform on the primes congruent to 1 modulo 4 in a cell $(z, \lambda z]$ with $z \ge y$ and $\lambda = 1 + r^{-\ell}$. Write μ_i for this measure and $P_i \gg z/(r^\ell \log z)$ for its population. Put $N_0 = \exp(1)X/a$.

Residue classes. For a nonempty set S of prime factors of a let χ_S be the Jacobi symbol modulo $a_S = \prod_{p \in S} p$. On odd primes the restriction $\mathbf{1}_{p \equiv 1 \bmod 4}\chi_S(p)$ is $\{\chi_S(p) + \chi_4(p)\chi_S(p)\}/2$. Both characters are real and nonprincipal, of modulus dividing $4a_S$. By (23), partial summation and the lower bound for P_i, for $|t| \le T$ and $0 \le \varsigma \le 1$,

$$\left| \mathbb{E}_{\mu_i} \chi_S(p)\, p^{-\varsigma - it} \right| \le \vartheta := C_\ell\, r^\ell y^{-1/2} \log^3(XT). \tag{24}$$

Given the ith column σ_i of σ, put

$$W_i(p) = \prod_{q \mid a} (1 + \sigma_{qi}\,(p/q)) = 1 + \sum_{S \ne \varnothing} \sigma_{S,i}\, \chi_S(p).$$

It equals 2^g if p has the prescribed symbols and 0 otherwise. Hence $\mathbb{E}_{\mu_i} W_i = 1 + \delta_i$ with $|\delta_i| \le 2^g \vartheta$. Let $\nu_i = W_i \mu_i/(1 + \delta_i)$, the uniform measure on the primes of the cell with the prescribed symbols.

The bilinear estimate on the restricted cells. The estimate of Friedlander and Iwaniec [9, Lemma 2] allows arbitrary coefficients of modulus at most one. For functions f, f' of modulus at most one,

$$\left| \mathbb{E}_{\nu_i \otimes \nu_j} f(p) B(p,p') f'(p') \right| = \frac{\left| \mathbb{E}_{\mu_i \otimes \mu_j} (fW_i)(p) B(p,p') (f'W_j)(p') \right|}{(1+\delta_i)(1+\delta_j)} \le \frac{4^g}{(1 - 2^g\vartheta)^2}\, \|T\|^{\mu}_{\infty \to 1},$$

because $0 \le W_i \le 2^g$. The $L^1 \to L^\infty$ norm is still at most one. Interpolation as in [7, Equation (6)] gives, for the kernels on $\nu_i \otimes \nu_j$,

$$\eta' \ll_\ell 2^g\, r^\ell y^{-1/12} (\log X)^{19/12} + 2^g / P_{\min},$$

where the last term is the cost of the artificial diagonal on the restricted cell, whose population is at least $2^{-g-1} P_{\min}$. Since $2^g \le L^J$, we have $r^2 2^r \eta' \to 0$ under (22).

The pattern count. For patterns (σ,τ) put $Z_\tau=\prod_{i<j}(1+\tau_{ij}B(p_i,p_j))$. The indicator of the atom is $2^{-gr-e}\prod_i W_i\cdot Z_\tau$. Write $\mathbf{1}$ for the cutoff $ab\le X$ and I for distinct coordinates. Then

$$\mathbb{E}_\mu \mathbf{1}I\Big(\prod_i W_i\,Z_\tau-1\Big)=\prod_i(1+\delta_i)\cdot\mathbb{E}_\nu\mathbf{1}I(Z_\tau-1)+\mathbb{E}_\mu\mathbf{1}I\Big(\prod_i W_i-1\Big).$$

The first expectation is the quantity bounded in [7, Equations (7)–(9)], now for the independent coordinates with laws ν_i. The weighted kernel count [7, Lemma S2.1] holds for arbitrary independent probability spaces. Use the Perron cutoff $\lfloor X/a\rfloor+1/2$; it introduces only contractive product weights. When two coordinates belong to the same cell, different prescribed columns have disjoint supports. Equal columns give the same restricted cell, so the collapse identity applies there. With $T=N_0^2 2^{\,e+gr}$ this gives

$$\big|\mathbb{E}_\nu\mathbf{1}I(Z_\tau-1)\big|\ll\log X\cdot r^2 2^r\eta'+\frac{r^2 2^{r+g}}{P_{\min}}+N_0^{-1}.$$

For the second expectation, apply the same Perron formula. The integrand is a product over i of $\mathbb{E}_{\mu_i}W_i(p)p^{-\varsigma-it}$, and by (24) each factor differs from $\mathbb{E}_{\mu_i}p^{-\varsigma-it}$ by at most $2^g\vartheta$. Both have modulus at most $1+2^g\vartheta$. Hence the two products differ by at most $(1+2^g\vartheta)^r-1\le 2r2^g\vartheta$, the integral contributes $O(\log X\cdot r2^g\vartheta)$, and the truncation error is $O(2^{gr}N_0/T)=O(N_0^{-1})$. Removing I costs at most $\binom{r}{2}\{(1+2^g\vartheta)^r\,2^{g+1}+1\}/P_{\min}$, since $\mathbb{E}_\mu\mathbf{1}_{p_i=p_j}\prod_k W_k=\prod_k(1+\delta_k)\,\nu(p_i=p_j)$ and two coordinates coincide with ν-probability at most $2^{g+1}/P_{\min}$.

Assembly. Sum over the boxes of cells as in [7, Equation (11)]; the ratio of the total box mass to the number of b is at most L^{C_0} by the population bounds in [7, Section S3]. Altogether

$$\epsilon_X^{\mathrm{sm}}\ll L^{C_0}\Big\{\log X\big(r^2 2^r\eta'+r2^g\vartheta\big)+\frac{r^2 2^{r+g}}{P_{\min}}+N_0^{-1}\Big\}.$$

The term with η' is $2^g\le L^J$ times the bound of [7, Equation (2)], which tends to zero faster than every power of L under (22). The term with ϑ is $\exp\{-\frac12\log y+O(L)\}$, and the others are smaller. □

6.2 Rows with prescribed labels in the model

In the reference law of Theorem 6.1 the block K_a is fixed and all other signs are independent and uniform. Conditioned on the corank, this is the model of Section 2 with a prescribed internal pattern on the set A of the first g rows. The next lemma is contained in the proof of Proposition 2.2.

Lemma 6.2 (Exposed rows)**.** *Let K have the law of Section 2 on $g+r$ rows, conditioned on binary rank s and on an event $\mathcal{C}$ of positive probability determined by B and the labels in a set A of g rows. Conditionally on B and on the labels in A, the labels of the remaining rows, apart from one of them, are independent and uniform on $\mathbb{F}_2^s$ conditioned on an event of probability larger than $1/4$. The excluded label is the sum of all the other labels, including those in A.*

6.3 The spectrum at every feasible 4-rank

Theorem 6.3 (Large prime factors, with small ones present)**.** *Assume* GRH_2 *and* (22)*. Let $a=a_X$ be squarefree with $g\le h$ prime factors, all at most y and congruent to 1 modulo 4, and let $k=k_X$ be integers such that the set of $d\in\mathcal{D}^{\mathrm{sm}}(X;y,h)$ with $a(d)=a$ and $\mathrm{rk}_4\,\mathrm{Cl}^+(d)=k$ is not empty. Let d be uniform in this set and put*

$$s=g+r-1-k,\qquad N=2^s.$$

Then the conclusions of Theorem 5.2 numbered two to six hold for the Legendre matrix K_b of the prime factors exceeding y, with this N. In particular, if $N(s+1) = o(r)$, then with probability tending to one exactly N eigenvalues of K_b lie outside $[-1,1]$ and

$$\operatorname{rk}_4 \mathrm{Cl}^+(d) = \omega(d) - 1 - \log_2 \#\{i : |\lambda_i(K_b)| > 1\} = \omega(d) - 1 - \operatorname{round}\Big(2\log_2 \frac{r}{\lambda_{\max}(K_b)}\Big).$$

If $N \to \infty$ and $N(s+1) = o(r)$, the bounds of Proposition 3.10 hold for $\lambda_{\max}(K_b)$. All statements are uniform in a, and therefore hold for d uniform in $\mathcal{D}^{\mathrm{sm}}(X; y, h)$ conditioned on its 4-rank and on $\omega(a(d)) = g$, for every $g \le h$.

Proof. For these discriminants the 4-rank is the excess corank κ of the binary Laplacian of K_d. By Theorem 6.1 the law of (C_d, K_b) given $a(d) = a$ has density $1 + O(\epsilon_X^{\mathrm{sm}})$ with respect to independent uniform signs. By Proposition 4.1 with a pointwise density bound, the law given in addition $\kappa(K_d) = k$ is within $o(1)$ in total variation of the reference law conditioned on the same event; the event has positive reference probability because the arithmetic set is not empty. The reference law is the model on $g + r$ rows conditioned on rank s and on the internal pattern K_a of the rows in $A = \{1, \dots, g\}$, and $K_b = K_{A^c}$. Section S6 of [7] uses Lemma 6.2 to compare the large-prime block with the quotient model. It gives the empirical and edge limits, the count of large eigenvalues, and the growing-outlier bounds. Uniform relabelling of the large-prime indices preserves the spectrum. The eigenvalue count gives s, and $k = \omega(d) - 1 - s$. □

6.4 The complete Legendre matrix

The complete matrix has the same empirical limit as K_b. The small-prime rows can change both extreme eigenvalues through the multiplicities of their labels.

Theorem 6.4 (Spectrum of the complete matrix)**.** *Keep the assumptions and notation of Theorem 6.3, put $n = r + g$, and write $E_+(K) = \lambda_{\max}(K)$ and $E_-(K) = -\lambda_{\min}(K)$.*

1. *At every feasible 4-rank k, the empirical eigenvalue law of $K_d/\sqrt{r}$ approaches $\nu_{r/N}$ in bounded Lipschitz distance in probability.*

2. *If $s/(\log r \log\log r) \to \infty$, then both $E_\pm(K_d)/\sqrt{n}$ tend to 2 in probability. If $r^{1/3}(\log r)^{3/2} = o(s)$, the two edge point processes centred at $\pm 2\sqrt{n}$ and scaled by $n^{1/6}$ converge jointly to independent GOE edge processes.*

3. *Suppose $N \to \infty$ and $r/N \ge (\log N)^{-\beta}$ for fixed $0 < \beta < 1$. Partition the small-prime indices by the relation $i \sim j$ if $e_i + e_j \in \ker R_d$, and let $m_A(d)$ be the largest class size (zero if $g = 0$). Then*

$$E_\pm(K_d) = (1 + o_{\Pr}(1))\sqrt{r \max\{b_N,\ r/N + m_A(d)\}}.$$

   In particular, with probability tending to one, each $E_\pm(K_d)$ is between $(1 - o(1))\sqrt{r b_N}$ and $(1 + o(1))\sqrt{r \max\{b_N, r/N + g\}}$. If $g = o(b_N)$, then both extremes have the same first-order magnitudes as for K_b.

4. *If $N(s+1) = o(r)$, then with probability tending to one there are exactly N eigenvalues outside $[-1,1]$, all of magnitude $(1 + o(1))n/\sqrt{N}$. In particular,*

$$k = n - 1 - \log_2 \#\{i : |\lambda_i(K_d)| > 1\} = n - 1 - \operatorname{round}\left(2\log_2 \frac{n}{E_+(K_d)}\right).$$

For growing N, Proposition 3.10 holds with n in place of r. At fixed N, centre the outliers at $\pm n/\sqrt{N}$ and divide the shifts by $\sqrt{n}$. Their law is asymptotic to the mixture of the Gaussian compression laws of Theorem 3.9, with each form B weighted by the number of g-tuples of labels producing K_a. The two collections of shifts are viewed as finite point measures, and the weights may depend on $a = a_X$.

These assertions are uniform over the prescribed small-prime parts and feasible ranks, with the indicated growth conditions on r and s.

Proof. The density comparison in the proof of Theorem 6.3 holds for the full sign pattern. The difference of K_d and $0_g \oplus K_b$ has rank at most $2g$. The rank inequality for empirical distribution functions gives the first assertion since $g = O(\log r)$. To pass to bounded Lipschitz distance, truncate using the bounded second moments of the two empirical laws. The extra mass at zero of $0_g \oplus K_b$ is g/n.

For the remaining assertions, use Lemma 6.2. In the sparse range, the exposed labels are linearly independent with probability tending to one, uniformly in the prescribed pattern. Products of signs involving unexposed rows then have the same comparison with independent signs as in the proof for K_b. The Wigner matrix in this comparison has the prescribed block on the first g rows. Section S6 of [7] proves the comparison and verifies the two edge limits for this reference.

In the range of item 3 the equivalence classes are precisely the sets of equal labels by (1) and the argument following it. The r independent draws have maximum occupancy asymptotic to b_N. On the at most g labels that appear among the small primes, their occupancies are $r/N + o_{\Pr}(b_N)$ uniformly. Adding the exposed rows thus gives maximum occupancy asymptotic to $\max\{b_N, r/N + m_A(d)\}$. Section S6 of [7] checks how the calculation of the two extremes changes when the extra rows concentrate on these labels.

In the dense range $gN = o(r)$: if $s \le \frac{1}{2}\log_2 r$, use $N \le \sqrt{r}$, and otherwise use $g = O(s+1)$. The relative occupancy deviations, now centred at n/N, are $o_{\Pr}(1)$. Equation (1) and Weyl's inequality give exactly N outliers and the stated recovery formula. To compare shifts, couple the full sample with r independent labels by replacing one draw with the required sum and adding the exposed rows. The quotient matrices differ in norm by $O_{\Pr}((g+1)\sqrt{N})$, since their occupancies are comparable to r/N in every class. For growing N this is $o(\sqrt{r \log N})$ by the same split on s. At fixed N its order is $o(\sqrt{r})$, and the multinomial central limit theorem gives the Gaussian compression laws for each B. Before conditioning on span, each exposed label tuple has the same probability. The remaining labels fail to span with probability at most $N2^{-r+1}$ uniformly in the exposed tuple, so the mixture weights are as stated. □

Example 6.5 (Small primes can change the extreme-eigenvalue scale)**.** *Take $r = \lfloor L/2 \rfloor$, $s = \lfloor \log_2 r \rfloor$ and $g = 3s$, allowed with $J = 4$. Prescribe the small-prime pattern in three groups of size s. In the binary encoding the cross block between the first two groups is I_s, and all off-diagonal entries touching the third group are zero. Complete the pattern by taking the Gram signs for $B = I_s$, with two copies of the standard basis and s zero labels. In every rank-s realization the first two groups are bases, by their cross block. All labels of the third group must therefore be zero. Each nonzero label has multiplicity at most two, so $m_A = s$ for large s. Here $r/N \in [1, 2)$ and $b_N \sim \log r / \log\log r$. The extreme magnitudes of the complete matrix are asymptotic to $\sqrt{rs}$, whereas for the large-prime submatrix they are asymptotic to $\sqrt{rb_N}$.*

For an arithmetic realization, take $y = (\log X)^{64}$. Under GRH_2 *one can choose the g primes successively in $(y/2, y]$, congruent to* 1 *modulo* 4 *and satisfying the prescribed symbols with earlier primes. The prime number theorem modulo* 4 *gives the principal term, and expansion of the further restrictions into real characters gives* (23) *for their contribution. The relative error at each step is*

$O(2^g y^{-1/2}(g+1)^2 \log^3 y) = o(1)$, *so there are enough primes to avoid repetitions. Reordering these primes by size preserves the pattern up to permutation. The prescribed pattern extends to binary rank* s *by assigning zero labels to all large-prime rows: the two bases in this realization have sum zero together. There are* r *distinct primes congruent to* 1 *modulo* 4 *in* $(y, y^2]$ *for large* X*, and their product with* a *is at most* $y^{g+2r} < X$*. The relative comparison in Theorem 6.1 therefore gives a positive count for the prescribed rank stratum.*

Corollary 6.6 (Threshold for recovery by counting eigenvalues)**.** *Let* d *be sampled as in Theorem 6.4, or unconditionally from the rough family of Theorem 5.2 at a prescribed* 4*-rank, with* $g = 0$ *and* $n = r$*. Suppose* $N \to \infty$ *and let* M_d *be the number of eigenvalues of* K_d *outside* $[-1, 1]$*, counted with multiplicity. Then*

$$\Pr(M_d = N) \longrightarrow \begin{cases} 1, & r/N - \log N \to +\infty, \\ \exp(-e^{-c}), & r/N - \log N \to c \in \mathbb{R}, \\ 0, & r/N - \log N \to -\infty. \end{cases}$$

Whenever $M_d = N$*, the identity* $N = 2^{n-1-k}$ *gives the recovery formula*

$$\mathrm{rk}_4 \mathrm{Cl}^+(d) = n - 1 - \log_2 M_d.$$

The limits are uniform over the admissible small-prime parts.

Proof. The relative comparison transfers the event $M_d = N$ from the rank-s quotient model with the prescribed small-prime block, as in the proof of Theorem 6.3. When all prime factors exceed y, Theorem 5.1 gives the transfer unconditionally, with no exposed rows.

If a label is missing, the rank of $K + \Delta$ in (1) is less than N. Its perturbation by $-\Delta$ therefore has fewer than N eigenvalues outside $[-1, 1]$. If every label occurs, the least singular value of the quotient matrix $\sqrt{N} D^{1/2} U D^{1/2}$ is at least $\sqrt{N}$, so Weyl's inequality gives exactly N such eigenvalues for $N > 4$.

The count cannot succeed for $N > n$. Otherwise $N \le n$ and the first $r - 1$ free labels span with probability $1 - o(1)$, uniformly in the exposed labels. Above the threshold a union bound shows that these draws already occupy all labels. Below it the number of empty labels outside the exposed set has mean $(N - |S|)(1 - 1/N)^{r-1} \to \infty$, where S is the set of exposed labels and $|S| \le g$. The emptiness indicators have nonpositive pairwise covariances, so with probability tending to one at least two labels remain empty. The final label can fill at most one.

At the threshold the number of empty labels after the independent draws converges to a Poisson variable of mean e^{-c}: its jth factorial moment is $(N)_j (1 - j/N)^{r-1} \to e^{-jc}$ for each fixed j. The $g = O(\log N)$ exposed labels cover an empty label with probability at most $g(1 - 1/N)^{r-1} = o(1)$. Given that a label x is empty, each free draw is uniform on $\mathbb{F}_2^s \setminus \{x\}$. Every nontrivial character has mean of absolute value $1/(N-1)$ for that law. Fourier inversion bounds the conditional probability that the final sum label equals x by $N^{-1} + (N-1)^{-(r-1)}$, regardless of the exposed labels. Summing over empty labels shows that the final label fills one with probability $o(1)$. The probability that the Poisson limit is zero is $\exp(-e^{-c})$. □

The largest-eigenvalue rounding rule has a different range. At the threshold with $g = 0$, Theorem 5.2, item 5, gives $2 \log_2(n/E_+(K_d)) - s \to -\log_2 e$ in probability since $b_N/(r/N) \to e$. Rounding then gives $s - 1$.

Acknowledgements and declarations

No external funding was received for this work. The author declares no competing interests.

OpenAI Codex was used to develop ideas, explore and check arguments, perform symbolic and finite computations, check references, and draft and revise the text. The author is responsible for all mathematical and bibliographic claims. Finite computations do not prove the asymptotic statements. Anthropic Claude also assisted with project-internal reviews and numerical checks. These checks are not external peer review or formal verification.

References

[1] L. Alpöge, *A remark on the Riemann hypothesis*, 20 November 2026, announcement and Lean source.

[2] A. V. Sutherland, *Verification files for the Lean proof of the Riemann hypothesis*, verification report.

[3] J. Alt, L. Erdős, T. Krüger and D. Schröder, *Correlated random matrices: band rigidity and edge universality*, Ann. Probab. **48** (2020), 963–1001. DOI: 10.1214/19-AOP1379.

[4] K. Bangachev and G. Bresler, *Random algebraic graphs and their convergence to Erdős–Rényi*, Random Structures Algorithms **66** (2025), e21276. DOI: 10.1002/rsa.21276. Numbered references use arXiv:2305.04802, version 2 (2023).

[5] Z. D. Bai and Y. Q. Yin, *Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix*, Ann. Probab. **16** (1988), 1729–1741. DOI: 10.1214/AOP/1176991594.

[6] S. Chan, P. Koymans, D. Milovic, and C. Pagano, *The 8-rank of the narrow class group and the negative Pell equation*, Forum Math. Sigma **10** (2022), e46. DOI: 10.1017/fms.2022.40.

[7] Z. Gao, *Supplement to Four-ranks of class groups and the spectrum of Legendre symbols*, 2026.

[8] L. Erdős, J. Henheik and V. Riabov, *Cusp universality for correlated random matrices*, arXiv:2410.06813, version 2 (2024).

[9] J. Friedlander and H. Iwaniec, *Ternary quadratic forms with rational zeros*, J. Théor. Nombres Bordeaux **22** (2010), 97–113. doi:10.5802/jtnb.706.

[10] É. Fouvry and J. Klüners, *On the negative Pell equation*, Ann. of Math. (2) **172** (2010), 2035–2104. DOI: 10.4007/annals.2010.172.2035.

[11] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. Colloq. Publ. **53**, American Mathematical Society, 2004.

[12] P. Koymans and C. Pagano, *Effective convergence of coranks of random Rédei matrices*, Acta Arith. **212** (2024), 337–358. DOI: 10.4064/aa230318-4-1.

[13] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, London Math. Soc. Lecture Note Ser. **335**, Cambridge University Press, 2006. DOI: 10.1017/CBO9780511735127.

[14] I. Soloveychik and V. Tarokh, *On the spectral norms of pseudo-Wigner and related matrices*, arXiv:1708.04291 (2017).

[15] I. Soloveychik, Y. Xiang and V. Tarokh, *Pseudo-Wigner matrices*, arXiv:1701.05544, version 5 (2018).

[16] S. Sachdeva and N. K. Vishnoi, *Faster Algorithms via Approximation Theory*, Found. Trends Theor. Comput. Sci. **9** (2013), 125–210. DOI: 10.1561/0400000065. Section 3.

[17] S. Sodin, *Several applications of the moment method in random matrix theory*, Proc. Int. Congr. Math., Seoul 2014, Vol. III, 451–475. arXiv:1406.3410.

[18] A. Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys. **207** (1999), 697–733. DOI: 10.1007/s002200050743.

[19] P. Stevenhagen, *Rédei reciprocity, governing fields and negative Pell*, Math. Proc. Cambridge Philos. Soc. **172** (2022), 627–654. DOI: 10.1017/S0305004121000335.

[20] J. A. Tropp, *Norms of random submatrices and sparse approximation*, C. R. Math. Acad. Sci. Paris **346** (2008), 1271–1274. DOI: 10.1016/j.crma.2008.10.008.

[21] J. A. Tropp, *User-friendly tail bounds for sums of random matrices*, Found. Comput. Math. **12** (2012), 389–434. DOI: 10.1007/s10208-011-9099-z.

[22] C. A. Tracy and H. Widom, *Matrix kernels for the Gaussian orthogonal and symplectic ensembles*, Ann. Inst. Fourier **55** (2005), 2197–2207. Published article.

[23] D. V. Voiculescu, K. J. Dykema and A. Nica, *Free Random Variables*, CRM Monograph Series **1**, American Mathematical Society, 1992.

Supplementary proofs

We give the uniform relative comparison for complete Legendre-sign patterns when the number of rough prime factors is of order $\log\log X$. The proof keeps the product cutoff and removes repeated prime coordinates before conditioning on the rank. We also give the proofs of the joint edge laws for observed blocks and for the complete matrix with prescribed small-prime signs.

Theorem references prefixed M point back to the main article. The supplementary sections and their results keep the S numbering of the separate supplement; all links in this combined version are internal.

S1 Relative comparison on a rough-prime family

Put $L = \log\log X$ and $e = \binom{r}{2}$. Fix $0 < c < C$ and $B > 1$ and let $cL \le r \le CL$ be an integer. All logarithms are natural. Set $Y_r = 31L + 12r\log 2$ and assume

$$\log y \le L^B, \qquad \frac{\log y - Y_r}{\log L} \longrightarrow \infty. \tag{1}$$

Let $\mathcal{D}^{\mathrm{rough}}(X;y)$ be the squarefree $d \le X$ with exactly r prime factors, all exceeding y and congruent to one modulo four, and write $a_X = |\mathcal{D}^{\mathrm{rough}}(X;y)|$. The parameters of this set include r. The window $|r - L/2| \le L^{2/3}$ is admissible, and $y = (\log X)^{64}$ satisfies the cutoff in that window. Even for typical r, every prime factor is required to exceed y. Choose d uniformly and label its prime factors by an independent uniform permutation. Let $\widetilde{\mu}_{X,y}$ be the resulting law of its zero-diagonal sign matrix K_d, and let U_r be the law of the zero-diagonal Rademacher Wigner matrix. For a sign matrix K, write $\kappa(K)$ for the excess corank of its binary graph Laplacian. It equals the narrow 4-rank for these discriminants.

For fixed $\ell \ge 1$, set $\lambda = 1 + r^{-\ell}$ and mark each labelled prime by J_i when $p_i \in (y\lambda^{J_i}, y\lambda^{J_i+1}]$. Write $\mathsf{P} = \mathcal{L}(\mathbf{J}, K_d)$ and $\rho = \mathcal{L}(\mathbf{J})$; the output marks need not be sorted.

Theorem S1.1 (Relative pattern comparison)**.** *Uniformly for $cL \le r \le CL$ and $Y_r \le \log y \le L^B$,*

$$\begin{aligned}\Delta_{X,y,\ell} &:= \sum_{\mathbf{j}} \max_{\tau} |2^e \mathsf{P}(\mathbf{j},\tau) - \rho(\mathbf{j})| \\ &\ll_{\ell,B,c,C} \exp\{-\tfrac{1}{12}\log y + \tfrac{31}{12}L + r\log 2 + O_{\ell,B,c,C}(\log L)\} \\ &\quad + \exp\{-\log y + L + r\log 2 + O_{\ell,B,c,C}(\log L)\} + a_X^{-1}.\end{aligned} \tag{2}$$

Under (1), $\Delta_{X,y,\ell} L^J \to 0$ for every fixed J. In particular

$$\epsilon_{X,y} := \max_{\tau} |2^e \widetilde{\mu}_{X,y}(\tau) - 1| \le \Delta_{X,y,1} = o(1). \tag{3}$$

For every graph event C with $p = U_r(C) > 0$,

$$\begin{aligned}\widetilde{\mu}_{X,y}(C) &= p\{1 + O(\epsilon_{X,y})\}, \\ d_{\mathrm{TV}}(\widetilde{\mu}_{X,y}(\cdot \mid C), U_r(\cdot \mid C)) &\le \epsilon_{X,y}/(1-\epsilon_{X,y}), \\ d_{\mathrm{TV}}(\mathsf{P}(\cdot \mid C), \rho \otimes U_r(\cdot \mid C)) &\le \Delta_{X,y,\ell}/(1-\Delta_{X,y,\ell}).\end{aligned}$$

These bounds are uniform even when p tends to zero arbitrarily fast. If $r \le L$ and $y \ge (\log X)^{64}$, then $\Delta_{X,y,\ell} \ll_{\ell,B,c,C} (\log X)^{-2}$. If $y = \exp(L^2)$ is admissible, then $\epsilon_{X,y} \le \exp\{-L^2/12 + O_{\ell,c,C}(L)\}$.

The graph event may include both an exact corank and a prescribed internal pattern. Uniform relabelling is part of the arithmetic law: the theorem makes no statement about labels ordered by prime size.

S2 Counting signs on independent prime sets

The following edge-deletion estimate [4, proof of Theorem 4.10] is used before the norm cutoff and the distinct-prime restriction. All operator norms below use probability measures.

Lemma S2.1 (Weighted kernel count)**.** *Let $x_1, \dots, x_m$ be independent finite probability spaces, and let $B_{ij}(x_i, x_j) \in \{\pm 1\}$ have averaging-operator norms at most η. For $\Phi = \prod_i \phi_i(x_i)$, $|\phi_i| \le 1$, and signs τ_{ij}, put $Z_\tau = \prod_{i<j}(1 + \tau_{ij} B_{ij})$. Then*

$$|\mathbb{E}\Phi(Z_\tau - 1)| \le D_m^{(1)} := (1 + 2^{m-2}\eta)^{\binom{m}{2}} - 1, \tag{4}$$

$$\sum_{\varnothing \ne F \subseteq E(K_m)} \left| \mathbb{E}\Phi \prod_{ij \in F} B_{ij} \right|^2 \le D_m := (1 + 2^{m-2}\eta^2)^{\binom{m}{2}} - 1. \tag{5}$$

For $m = 1$ both bounds are zero. Complex product weights are allowed.

Proof. For a graph H set $C_H = \mathbb{E}\prod_{ij \in H}(1 + B_{ij})$. Remove an edge ij to obtain H_0, and remove both its endpoints to obtain H'. With the other coordinates fixed, write the remaining factors as $g f_i f_j$, where g uses H' and f_i, f_j use edges incident to the endpoints. They are nonnegative, with $f_i^2 = 2^{d_i - 1} f_i$ and $f_j^2 = 2^{d_j - 1} f_j$. The operator bound and Cauchy–Schwarz give

$$|C_H - C_{H_0}| \le \eta 2^{(d_i + d_j - 2)/2} \mathbb{E}[g\sqrt{(\mathbb{E}_i f_i)(\mathbb{E}_j f_j)}] \le \eta 2^{m-2} \sqrt{C_{H'} C_{H_0}}.$$

Induction and telescoping yield $|C_H - 1| \le (1 + 2^{m-2}\eta)^{|E(H)|} - 1$. Insert the endpoint weights before applying the operator bound, and remove all other weights after taking absolute values. Their moduli cannot increase the same positive counts. Telescoping now gives (4). Replacing B_{ij} by $\tau_{ij} B_{ij}$ changes neither its norm nor positivity of its factors.

On two independent copies x, x', use the product weights $\Phi(x)\overline{\Phi(x')}$ and the paired kernels

$$B_{ij}(x_i, x_j) B_{ij}(x_i', x_j').$$

Their norms are at most η^2. Expanding the product and subtracting the empty term $|\mathbb{E}\Phi|^2$ gives exactly the sum on the left of (5); the first bound proves the second. □

S3 The norm cutoff and distinct primes

Proof of Theorem S1.1. The bilinear estimate. We keep track of the logarithmic costs of the bilinear input. The estimate of Friedlander–Iwaniec [2, Lemma 2], for complex coefficients of modulus at most one supported on odd integers, is

$$\left| \sum_{a \le M, b \le N} \alpha_a \beta_b (a/b) \right| \ll MN(M^{-1/6} + N^{-1/6})(\log(3MN))^{7/6}.$$

On prime sets 1 mod 4 in the λ-cells with lower endpoints at least y and upper endpoints at most $N = \exp(1)X$, fixed-progression PNT and probability normalization give

$$\|T\|_{\infty \to 1} \ll_\ell r^{2\ell} y^{-1/6} (\log X)^{19/6}.$$

Since $\|T\|_{1\to\infty} \le 1$, interpolation bounds the L^2 operator norm by its square root. On equal prime sets replace the Jacobi diagonal zero by $B(p,p) = 1$, at operator cost $1/P$ for prime count P. The resulting symmetric sign kernels therefore satisfy

$$\eta \ll_\ell r^\ell y^{-1/12}(\log X)^{19/12} + P_{\min}^{-1}, \qquad P_{\min} \gg_\ell y/(r^\ell \log y). \tag{6}$$

These estimates are uniform even for very unequal intervals. The fixed-progression PNT error is smaller than the cell population: $r^\ell \log z \exp(-c\sqrt{\log z}) = o(1)$ uniformly for cell endpoints $z \ge y$. It supplies a lower bound only, and is not added to the transfer error. At $\log y \ge Y_r$ its population error is already bounded by a polynomial in L times $\exp(-c\sqrt{L})$; thus this estimate is uniform throughout the enlarged cutoff range.

For each pattern τ, set $Z_\tau = \prod_{i<j}(1+\tau_{ij}B(p_i,p_j))$. It equals 2^e on that atom and zero elsewhere. Lemma S2.1 applies with the bound (6), including contractive complex product weights.

Put $x_* = \lfloor X \rfloor + 1/2$, $\sigma = 1/\log(2N)$, and $T = N^2 2^e$. Uniformly for integer products $1 \le d \le N$, truncated Perron inversion gives

$$\mathbf{1}_{d\le X} = \frac{1}{2\pi}\int_{-T}^{T} \frac{(x_*/d)^{\sigma+it}}{\sigma+it}\,dt + O(N/T).$$

Integration by parts proves the tail bound using $|\log(x_*/d)| \gg N^{-1}$ and $(x_*/d)^\sigma \ll 1$. For $d = \prod_i p_i$, the phases $p_i^{-\sigma-it}$ are contractive. Since $|Z_\tau - 1| \le 2^e$, (4) yields

$$|\mathbb{E}\mathbf{1}_{d\le X}(Z_\tau - 1)| \ll \log(2T/\sigma) D_r^{(1)} + N^{-1}. \tag{7}$$

Here $\log(2T/\sigma) = O(\log X)$. There is no factor equal to the number of graph atoms in this inversion.

Removing repeated primes. The incident Legendre symbols agree when two prime coordinates coincide. We use this identity before estimating the collision contribution. They can coincide only in the same cell. On $p_i = p_j$, the atom vanishes unless $\tau_{ij} = 1$ and $\tau_{ik} = \tau_{jk}$ for every $k \ne i, j$, with unordered edge indices understood. In the compatible case, delete label j and denote the induced pattern by τ^{ij}. The exact identity is

$$Z_\tau\big|_{p_i=p_j} = 2^{r-1} Z_{\tau^{ij}}. \tag{8}$$

The factor is the loop factor 2 times $r-2$ duplicate edge factors, using $(1+u)^2 = 2(1+u)$ for $u = \pm 1$. Conditional on this equality, the remaining $r-1$ coordinates are still independent and uniform on their respective prime sets. Their kernels satisfy (6), even if more coordinates coincide. Thus, for a common cell of size P,

$$\mathbb{E}\mathbf{1}_{p_i=p_j} Z_\tau \le \frac{2^{r-1}}{P}(1 + D_{r-1}^{(1)}).$$

Write I_{inj} for pairwise distinct coordinates. Positivity lets us drop the norm cutoff in this upper bound. A union bound, also charging the removed baseline 1, gives

$$\begin{aligned}|\mathbb{E} I_{\mathrm{inj}}\mathbf{1}_{d\le X}(Z_\tau - 1)| &\ll \log(2T/\sigma) D_r^{(1)} + N^{-1} \\ &\quad + \frac{\binom{r}{2}}{P_{\min}}\{2^{r-1}(1 + D_{r-1}^{(1)}) + 1\}.\end{aligned} \tag{9}$$

Multiple coincidences are retained in the collapsed counts. The identity uses the same Legendre kernel on incident edges and the diagonal $+1$; it need not hold for unrelated label-dependent sign kernels.

For each nondecreasing list of r cells whose lower endpoints have product at most X, use the full independent product box. If cell j occurs c_j times, give each tuple weight $1/\prod_j c_j!$. A squarefree product under the cutoff has total weight one. Every full-box product is at most $\lambda^r X \le N$. Let $\mathcal{M}$ be the total full-box mass. Before deletion a product $d = \prod_p p^{v_p}$ has weight $1/\prod_p v_p! \le 1$. Hence $\mathcal{M} \le F_r(N; y)$, the sum of these weights over all such products at most N. Put $H_z = \sum_{y<p\le z,\ p\equiv 1(4)} p^{-1}$. Products at most $\sqrt{N}$ contribute at most $\sqrt{N}$. Every larger product has a coordinate exceeding $N^{1/(2r)}$. Choose its position and use the prime-counting upper bound $O(rN/(c\log N))$ after fixing the other product c; it applies since $N/c \ge N^{1/(2r)}$. Dropping the other product restrictions gives

$$\mathcal{M} \le \sqrt{N} + O\left(\frac{rN}{\log N}\frac{H_N^{r-1}}{(r-1)!}\right).$$

The factor $1/(r-1)!$ comes from choosing a position in an ordered r-tuple of weight $1/r!$; repeats remain included.

To compare with $a_X = |\mathcal{D}^{\mathrm{rough}}(X; y)|$, put $Q = X^{1/(2r)}$. Here $y < Q$ eventually, since $\log Q = e^L/(2r) \gg L^B$. Choose $r-1$ distinct primes $q \equiv 1 \pmod 4$ in $(y, Q]$, and let a be their product. A final prime 1 mod 4 in $(X/(2a), X/a]$ has $\gg X/(a\log X)$ choices. Since $a \le X^{(r-1)/(2r)}$, this last prime exceeds Q for large X, so the representation is unique. Uniformly in the stated range of y,

$$H := \sum_{y<q\le Q,\ q\equiv 1(4)} q^{-1} \ge \tfrac{1}{2}L - O_B(\log L).$$

The repeated-index fraction in H^{r-1} is at most $\binom{r-1}{2}H^{-2}\sum_{q>y} q^{-2} \ll r^2/(yH^2) = o(1)$. Dividing distinct tuples by $(r-1)!$ and using Stirling therefore gives

$$a_X \gg \frac{X}{\log X}\frac{H^{r-1}}{(r-1)!} \ge X\exp\{-O_{B,c,C}(L)\}, \qquad \frac{\mathcal{M}}{a_X} \ll L^{C_0}. \tag{10}$$

For the last bound use $H = H_Q \asymp L$, with $H_Q \ge L/2 - O_B(\log L)$, and $H_N - H_Q = O(\log L)$, so $(H_N/H_Q)^{r-1} \le L^{O(1)}$; the fixed exponent C_0 may depend on B, c, C, and the $\sqrt{N}$ term is negligible. In fact the upper and lower bounds, with Stirling's formula, give

$$\log(a_X/X) = -L + (r-1)\left(1 + \log\frac{L}{2(r-1)}\right) + O_{B,c,C}(\log L).$$

This controls population, with no factor equal to the number of cells or boxes.

Retain the cell labels when summing (9). Only input-box indices are sorted; both output marks and graph are relabelled. For each output mark the triangle inequality bounds the maximum atom discrepancy by the sum of its box contributions. Summing over output marks charges each box just once, since the relabelling weights sum to one. Thus

$$\Delta_{X,y,\ell} \ll \frac{\mathcal{M}}{a_X}\{(\log X)D_r^{(1)} + r^2 2^r/P_{\min} + N^{-1}\}. \tag{11}$$

This step establishes a joint estimate, not just its graph marginal.

Uniformly for $Y_r \le \log y \le L^B$, $r^2 2^r \eta \le \exp\{-L + O_{\ell,B,c,C}(\log L)\} = o(1)$, so $D_r^{(1)} \ll r^2 2^r \eta$. Combining (6), (9) and (10) with this assembly budget proves (2). The second term includes the artificial-diagonal operator cost and its Perron factor as well as the smaller injectivity-repair cost. The structural cutoff makes the first exponent $-(\log y - Y_r)/12 + O_{\ell,B,c,C}(\log L)$ tend to minus infinity faster than any fixed multiple of $\log L$; the second term and a_X^{-1} are smaller. This

proves $\Delta_{X,y,\ell} L^J \to 0$ for every fixed J. If $r \le L$ and $\log y \ge 64L$, the first exponent is at most $-(11/4 - \log 2)L + O_{\ell,B,c,C}(\log L)$; its coefficient $11/4 - \log 2 = 2.056852\ldots > 2$. The other terms are smaller. This gives the stated stronger rate and the special case $y = \exp(L^2)$, also with Δ in place of ϵ.

For a graph event C of reference mass $p > 0$, the relative density bound gives $|\widetilde{\mu}_{X,y}(C) - p| \le \epsilon p$ and localized variation at most ϵp. Normalizing as in Proposition M4.1 gives the conditional bound without reciprocal mass loss. For the marked assertion put $\delta_{\mathbf{j}} = \max_\tau |2^e \mathsf{P}(\mathbf{j}, \tau) - \rho(\mathbf{j})|$. The unnormalized L^1 discrepancy on C is at most $p \sum_{\mathbf{j}} \delta_{\mathbf{j}} = p\Delta$, and its arithmetic normalizer is at least $p(1 - \Delta)$. The same normalization proves the marked bound. □

S4 Blocks outside prescribed vertices

The relative estimate and the main article's partial-observation comparison concern different steps. The former reaches rare arithmetic conditions; the latter determines which observed spectra retain GOE limits after those conditions.

Corollary S4.1 (Exterior block laws under internal constraints)**.** *In Theorem S1.1, fix an integer-valued corank $k = k(X)$ and put $s = r - k - 1$. Let $A \subseteq [r]$ be deterministic, and let $\mathcal{C}$ be any event determined by the internal off-diagonal signs on A, with $U_r(\kappa = k, \mathcal{C}) > 0$. For a fixed number of deterministic sets $I_1, \ldots, I_b \subseteq [r] \setminus A$, suppose $|I_i \cap I_j| \le 1$ for $i \ne j$, $m_j = |I_j| \to \infty$, and*

$$\sum_{j=1}^{b} m_j^{1/3} (\log m_j)^{3/2} = o(s).$$

Conditioned on $\kappa(K_d) = k$ and $\mathcal{C}$, the $2b$ processes

$$\sum_{\xi \in \mathrm{Spec}(K_d)_{I_j, I_j}} \delta_{m_j^{1/6}(\pm\xi - 2\sqrt{m_j})}$$

converge jointly to independent GOE edge processes. Their largest points converge to independent TW_1 variables. The result is uniform over all feasible internal conditions, including cyclic or dense patterns, with no lower bound on their probability.

Proof. Apply Theorem S1.1 to the single graph event $C = \{\kappa = k\} \cap \mathcal{C}$. Its conditional labelled law differs from $U_r(\cdot \mid \kappa = k, \mathcal{C})$ by at most $\epsilon_{X,y}/(1 - \epsilon_{X,y}) = o(1)$. The latter law has the asserted limits by Theorem M3.11. Observation cannot increase total variation, so the same limits hold arithmetically. □

For example, at fixed k, an arbitrary feasible internal pattern on $\lfloor \gamma r \rfloor$ vertices, $0 < \gamma < 1$, leaves the entire complement with GOE edge laws. This is an assertion about the complement, not about the spectrum of the full conditioned matrix. All sets and conditions refer to the independent uniform prime labelling above; prime-size-selected blocks, cross-sign conditions and mark-adaptive constraints are not covered. No extension to the main article's global CKMP family follows from this relative estimate for a different family.

S5 Proof of the joint edge laws

We prove Theorem M3.11 of the main article. Use its conditioned binary model and quotient labels. Here W_r is the symmetric sign matrix with independent uniform off-diagonal signs and zero diagonal, and $\mathfrak{c}(m) = m^{1/3}(\log m)^{3/2}$.

We first prove the unconditioned whole-matrix case $A = \varnothing$, $b = 1$, $I_1 = [r]$, under

$$s_r/[r^{1/3}(\log r)^{3/2}] \longrightarrow \infty, \qquad s_r = r - k(r) - 1. \tag{12}$$

The block argument is given at the end of the proof. In the base case put

$$A_r = K/(2\sqrt{r}), \quad A_r^{\mathrm{Wig}} = W_r/(2\sqrt{r}), \quad d_r = r^{1/3}\sqrt{\log r}, \quad \rho_r = r^{-2/3}\log r, \quad R_r = 1 + \rho_r,$$

and write $\Xi_r^{\pm} = \sum_i \delta_{r^{1/6}(\pm\lambda_i(K) - 2\sqrt{r})}$. Let $\Xi_{\mathrm{GOE}}^{\pm}$ be independent GOE edge processes.

Compressed traces. Let S_p be a sum of p independent uniform signs and T_j the first-kind Chebyshev polynomial, with $T_{-j} = T_j$. For fixed $t > 0$ and $\epsilon \in \{0, 1\}$ define

$$p = 2\lfloor tr^{2/3} \rfloor + \epsilon, \quad m = \lceil 4(t+1)d_r \rceil, \qquad P_{r,t,\epsilon}(x) = R_r^p\, \mathbb{E}\big[T_{S_p}(x/R_r)\mathbf{1}_{\{|S_p| \le m\}}\big].$$

This has degree at most m and parity ϵ; the underlying identity $x^p = \mathbb{E}T_{S_p}(x)$ is classical [5, Section 3].

Lemma S5.1 (Approximation and exterior control)**.** *For every fixed $t > 0$ and $\epsilon \in \{0, 1\}$, eventually*

$$\sup_{|x| \le R_r} |P_{r,t,\epsilon}(x) - x^p| \le r^{-8}. \tag{13}$$

The absolute monomial coefficient sum is at most $R_r^p 3^m$. For every fixed positive integer h,

$$\mathbb{E}_{\mathrm{Wig}} \left| \operatorname{tr} P_{r,t,\epsilon}(A_r^{\mathrm{Wig}}) - \operatorname{tr}(A_r^{\mathrm{Wig}})^p \right|^h \longrightarrow 0. \tag{14}$$

Proof. The cosine identity gives $x^p = \mathbb{E}T_{S_p}(x)$ on $[-1, 1]$, hence everywhere by polynomial continuation. Hoeffding's bound gives

$$\sup_{|x| \le R_r} |P_{r,t,\epsilon}(x) - x^p| \le 2R_r^p e^{-m^2/(2p)}, \qquad p\log R_r - \frac{m^2}{2p} = \big(2t - 4(t+1)^2/t + o(1)\big)\log r < -12\log r.$$

The recurrence $T_{j+1} = 2xT_j - T_{j-1}$ gives coefficient sum at most 3^j, proving the remaining coefficient assertion.

For $|x| = R_r \cosh u$, $u \ge 0$, we have $|P_{r,t,\epsilon}(x)| \le R_r^p e^{mu}$. If L is even and $m/L \le 1/2$, maximizing at $u = \operatorname{arctanh}(m/L)$ and using $\operatorname{arctanh} v \le 2v$ give

$$\sup_{u \ge 0}(mu - L\log\cosh u) \le m^2/L, \qquad |P_{r,t,\epsilon}(x)| \le e^{-(L-p)\log R_r + m^2/L}|x|^L \quad (|x| \ge R_r).$$

Take $L = 2\lfloor 8(t+1)^2 r^{2/3} \rfloor$. The exponent divided by $\log r$ tends to $-16t^2 - 30t - 15 < -8$; $|x|^p$ satisfies the same bound without m^2/L. Hence for every real symmetric A,

$$|\operatorname{tr} P_{r,t,\epsilon}(A) - \operatorname{tr} A^p| \le r^{-7} + 2r^{-8}\operatorname{tr} A^L.$$

Soshnikov's mixed high-trace bounds [6, Theorem 3] bound every fixed moment of the last trace for $A = A_r^{\mathrm{Wig}}$, proving (14). □

Proposition S5.2 (Compressed trace comparison)**.** *Under* (12)*, for fixed ℓ, $t_a > 0$ and $\epsilon_a \in \{0, 1\}$,*

$$\mathbb{E}\left[\prod_{a=1}^{\ell} \operatorname{tr} P_{r,t_a,\epsilon_a}(A_r) \;\middle|\; \kappa(R) = k(r)\right] - \mathbb{E}_{\mathrm{Wig}}\left[\prod_{a=1}^{\ell} \operatorname{tr}(A_r^{\mathrm{Wig}})^{2\lfloor t_a r^{2/3} \rfloor + \epsilon_a}\right] \longrightarrow 0.$$

Each compressed trace has bounded moments of every fixed even order.

Proof. Put $M = \sum_a \lceil 4(t_a+1)d_r \rceil$. The product is an edge-sign polynomial of degree at most M, with absolute coefficient sum

$$r^{\ell} R_r^{\sum_a p_a} (3\sqrt{r})^M = \exp\{O(d_r \log r)\}.$$

The odd-multiplicity edges of its closed walks form Eulerian graphs, so each nonconstant monomial uses at most $M < r$ vertices. Its bias is at most $11\, 2^{M-s_r}$ by Proposition M2.2 (use the global spanning bound for $s_r \le (r-1)/2$, the local bound otherwise). The expectation error is thus $\exp\{-s_r \log 2 + O(d_r \log r)\} = o(1)$. Lemma S5.1, Hölder and the Wigner mixed-trace bounds replace the Wigner polynomials by powers. Repeating with a fixed even number of copies proves the moment bounds; absolute odd-power moments are bounded by even powers at $t/2$ and $2t$. □

Lemma S5.3 (Spectral confinement)**.** *Under* (12)*, for every fixed $B > 0$,* $\Pr(\|A_r\| > R_r \mid \kappa(R) = k(r)) \le C_B r^{-B}$*. Outside this event both rescaled edges have no points above $w_r = 2\log r$.*

Proof. For fixed $D > 0$ put $R_{0,r} = 1 + \rho_r/2$, $d = \lceil D d_r \rceil$ and $Q_r(x) = T_d(x/R_{0,r})^2$. This is nonnegative everywhere and at most one on $[-R_{0,r}, R_{0,r}]$. The preceding exterior argument, with $2d$ in place of m, gives

$$Q_r(x) \le e^{-L \log R_{0,r} + 4d^2/L} |x|^L \qquad (|x| \ge R_{0,r}).$$

For $L = 2\lfloor u_D r^{2/3} \rfloor$, $u_D = 4(D+1)^2$, the exponent is $(-u_D + 2D^2/u_D + o(1)) \log r$, with coefficient below $-7/2$. The Wigner even-trace bound gives $\mathbb{E}_{\mathrm{Wig}} \operatorname{tr} Q_r(A_r^{\mathrm{Wig}}) \le r + O(r^{-2})$. Its degree is $2d = O(d_r)$ and its edge-sign coefficient sum is at most $r(3\sqrt{r})^{2d}$, so the same character comparison gives $\mathbb{E} \operatorname{tr} Q_r(A_r) \le r + o(1)$. For $|x| \ge R_r$,

$$Q_r(x) \ge \tfrac{1}{4} e^{2d \operatorname{arcosh}(R_r/R_{0,r})} = \tfrac{1}{4} r^{2D + o(1)},$$

since $\operatorname{arcosh}(R_r/R_{0,r}) \sim \sqrt{\rho_r}$. Positivity and Markov give $O(r^{1-2D+o(1)})$; choose $D > B + 2$. The edge cutoff follows from $2r^{2/3}(R_r - 1) = 2\log r$. □

All times, moment orders and confidence exponents are fixed before $r \to \infty$; (12) pays every fixed multiple of the compressed degree.

Recovering the edge processes. Retain in Ξ_r^+ the points at most w_r that come from positive eigenvalues, and call the result $\widetilde{\Xi}_r^+$. Define $\widetilde{\Xi}_r^-$ analogously from negative eigenvalues, with the same inclusive cutoff. These measures agree with $\Xi_r^{\pm}$ on every fixed compact set eventually.

Lemma S5.4 (Local factorial-Laplace transfer)**.** *The joint factorial moment measures of $(\widetilde{\Xi}_r^+, \widetilde{\Xi}_r^-)$ converge vaguely in every fixed order to those of $(\Xi_{\mathrm{GOE}}^+, \Xi_{\mathrm{GOE}}^-)$. They identify the joint point-process limit.*

Proof. Trace replacement. Fix $t > 0$, let $m(u) = \lfloor u r^{2/3} \rfloor$, $E_r = \{\|A_r\| > R_r\}$ and put

$$L_r^{\pm}(t) = \int e^{tx}\, d\widetilde{\Xi}_r^{\pm}(x), \quad H_r^{\pm}(t) = \tfrac{1}{2} \operatorname{tr}(P_{r,t,0} \pm P_{r,t,1})(A_r), \quad G_r^{\pm}(t) = \tfrac{1}{2} \operatorname{tr} A_r^{2m(t)}(1 \pm A_r).$$

Set $Z_r(t) = 1 + \operatorname{tr} A_r^{2m(t/2)} + \operatorname{tr} A_r^{2m(4t)}$, and obtain $\widehat{Z}_r(t)$ by replacing its powers by compressed polynomials. On E_r^c their trace errors, and $H_r^{\pm} - G_r^{\pm}$, are $O(r^{-7})$. Proposition S5.2 gives, for fixed h,

$$\mathbb{E}|\widehat{Z}_r(t)|^{2h} = O_{h,t}(1), \qquad \mathbb{E}[Z_r(t)^{2h} \mathbf{1}_{E_r^c}] = O_{h,t}(1). \tag{15}$$

For $|x| \le w_r$ and $y = 1 + x/(2r^{2/3})$, Taylor's formula gives $y^{2m(t)} = e^{tx}\{1 + O_t(r^{-2/3}(|x| + x^2))\}$. The smaller and larger powers in Z_r control the two signs of x; the parity projector's error at either edge is $O(r^{-2/3}w_r)$. In the inner region $1/2 \le y \le 1 - w_r/(2r^{2/3})$,

$$y^{2m(t)} + e^{2tr^{2/3}(y-1)} \le C_t e^{-c_t w_r} y^{2m(t/2)}.$$

This follows by subtracting powers and using $2(1-y) + \log y \ge (1-y)/2$ on $[1/2, 1]$. The region $|y| < 1/2$ contributes $O_t(re^{-c_t r^{2/3}})$. Applying these bounds at both edges yields, on E_r^c,

$$|G_r^{\pm}(t) - L_r^{\pm}(t)| \le C_t\{r^{-2/3}(\log r)^2 + r^{-c_t}\}Z_r(t) + C_t re^{-c_t r^{2/3}}.$$

This tends to zero in every fixed moment by (15). On E_r, $H_r^{\pm}$ has bounded $2h$th moment and $L_r^{\pm}(t) \le r^{1+2t}$ deterministically. Cauchy–Schwarz and Lemma S5.3 bound its contribution by $O_{h,t}(\Pr(E_r)^{1/2}) + r^{h(1+2t)}\Pr(E_r) = o(1)$, taking $B > h(1+2t) + 1$. Thus

$$L_r^{\pm}(t) - H_r^{\pm}(t) \longrightarrow 0 \quad \text{in every fixed } L^h. \tag{16}$$

Factorial measures. Fix q and an open rectangle

$$U_q \Subset (0, \infty)^q. \tag{17}$$

For one colour the distinct-index sum satisfies

$$\sum_{i_1,\dots,i_q \text{ distinct}} e^{\sum_j t_j x_{i_j}} = \sum_{\pi \in \Pi_q} \prod_{B \in \pi} (-1)^{|B|-1}(|B|-1)!\, L_r\left(\sum_{j \in B} t_j\right). \tag{18}$$

Apply this separately to the two colours, whose underlying eigenvalue index sets are disjoint. Equations (16)–(18), Proposition S5.2 and [6, Corollary 6] give convergence of every joint factorial Laplace transform on U_q to the independent GOE transforms. The same comparison at doubled orders bounds the required second moments.

Write $\alpha_r^{a,b}$ for the order-(a,b) factorial measure, $a + b = q$. Choose $\mathbf{t}^0 + [-2\delta, 2\delta]^q \subset U_q$ and set $d\nu_r(\mathbf{x}) = e^{\langle \mathbf{t}^0, \mathbf{x}\rangle} d\alpha_r^{a,b}(\mathbf{x})$. The convergent transforms at the cube corners imply

$$\sup_r \int e^{2\delta\|\mathbf{x}\|_1}\, d\nu_r < \infty, \qquad e^{2\delta\|\mathbf{x}\|_1} \le \sum_{\boldsymbol{\sigma} \in \{\pm 1\}^q} e^{\langle 2\delta\boldsymbol{\sigma}, \mathbf{x}\rangle}.$$

Thus these finite measures are tight, and their transform tails are $O(e^{-\delta R})$ on $\|\mathbf{x}\|_1 > R$ when $\|\mathbf{z}\|_\infty \le \delta$. Every subsequential limit has the prescribed transform near zero; uniqueness of the multivariate Laplace transform identifies it with the tilted GOE factorial measure. Untilting on compact sets proves the asserted vague convergence.

Point-process identification. On a bounded interval I, let M_I bound the entries of the 2×2 GOE kernel [7, Section IV]. Its Pfaffian expansion gives

$$0 \le \rho_j(\mathbf{x}) \le (2j-1)!!M_I^j \le j!(2M_I)^j, \qquad \mathbb{E}(N_I)_j \le j!(2M_I|I|)^j.$$

Hence $\mathbb{E}(1+u)^{N_I} < \infty$ for small $u > 0$; the independent colours also have an exponential moment for their total count. Their joint occupation laws on finitely many bounded intervals are therefore moment determinate.

First factorial measures give local tightness. Any subsequential limit has no atom at a fixed endpoint: test a shrinking continuous bump there and use the continuous GOE first-measure density.

Counts on finitely many bounded intervals consequently converge along that subsequence. Their joint factorial moments converge, and bounds at twice each order give uniform integrability of all count products. The limiting counts thus have the GOE moments; determinacy identifies their joint law and hence the two point processes. Since $w_r \to \infty$, the same local identification holds for the untruncated measures. □

Proof of Theorem M3.11. First take $b = 1$ and $I_1 = [r]$. Lemma S5.4 identifies the two processes. Moreover its proof gives $\sup_r \mathbb{E}L_r^{\pm}(t) < \infty$ for fixed $t > 0$. Markov and Lemma S5.3 therefore give

$$\limsup_{r\to\infty} \Pr(\Xi_r^{\pm}((L,\infty)) > 0) \ll_t e^{-tL}.$$

For fixed a, b and $T > \max(a, b)$, the joint maximum distribution at (a, b) differs from the probability

$$\Pr\{\Xi_r^{+}((a,T]) = 0, \Xi_r^{-}((b,T]) = 0\}$$

by at most the probability of a point above T. The zero-count probability converges by point-process convergence and the absence of GOE endpoint atoms. The same exponential tail bound holds for the limiting processes. Letting first $r \to \infty$, then $T \to \infty$, proves the joint extreme-value law. *Observed blocks after a partial observation.* Put $M = r - 1 - |A|$ and $S = \sum_j \mathfrak{c}(m_j)$. In block j, repeat the compression at size m_j, with $A_j = K_{I_j,I_j}/(2\sqrt{m_j})$. For fixed moment orders, the total degree is $L = O(\sum_j m_j^{1/3}\sqrt{\log m_j})$, and the absolute coefficient sum is $\exp\{O(S)\}$. If $m_* = \max_j m_j$, then $M \geq m_* - 1$ and $S \leq b\mathfrak{c}(m_*) = o(M)$. Hence $|A| + L < r$. Proposition M2.2 now gives, uniformly after $\mathcal{C}$, the mixed-trace error

$$\exp\{-s\log 2 + O(S)\} + \exp\{-\tfrac{1}{2}M\log 2 + O(S)\} = o(1).$$

The first term uses the stated rank condition; the second is paid automatically by the observed sizes. Thus fixing A introduces no extra restriction on the rank and no inverse probability cost. Constant characters agree exactly. The reference Wigner blocks are independent because their off-diagonal edge sets are disjoint, even when they share one vertex. The same comparison bounds all fixed moments of the compressed statistics.

The positive detector in block j gives $\Pr(\|A_j\| > 1 + m_j^{-2/3}\log m_j \mid \mathcal{C}) \leq C_D m_j^{-D}$ for each fixed D; its comparison error is smaller than any fixed power of m_j. The trace replacement in Lemma S5.4 consequently holds in every fixed L^p at each block's own size. Telescope mixed products and apply Hölder to these individual errors and moment bounds. This avoids charging a large block to a smaller block's exceptional probability. The factorial-transform and count-determinacy argument then applies with $2b$ fixed colours; its individual tail bounds also give convergence of their largest points. Without $\mathcal{C}$, the global approximations adjoin because their metric errors tend to zero in probability, by Slutsky's theorem on subsequences with limiting ratios. Every estimate is uniform in B, proving the theorem before averaging over the quotient form. □

Remark S5.5. *The additional condition may inspect only B and the exposed labels. The complete matrix can have different extreme-eigenvalue scales when many exposed labels coincide, as in Example M6.5. Corollary S4.1 uses the relative comparison to transfer the joint block law to the rough-prime family.*

S6 Spectral limits with prescribed small-prime signs

We prove the model estimates used in Theorems M6.3 and M6.4. The first $g = O(\log r)$ labels, on a set A, give the prescribed small-prime block. The remaining r labels give the large-prime block. Put

$n = r + g$, $N = 2^s$ and $\lambda = r/N$. We use the normalization b_N from Theorem M3.7. All estimates are uniform in the prescribed pattern and the nondegenerate symmetric form B.

The large-prime submatrix. After exposing B and the labels on A, use Lemma M6.2 for the remaining r labels. For the sparse bulk and sparse edge locations, use the conditional character comparison in Proposition M2.2 with $M = r - 1$. For t traces of total walk length J, the additional spanning error after normalization is at most $2\sqrt{5}\, r^{J/2-t} 2^{(J-r+1)/2}$. It is negligible for the fixed walk lengths for the bulk and the logarithmic lengths in the edge-location argument. The quadratic character bound is unchanged. When $s > r - 1$ this range gives the semicircle limit, since $r/2^s \to 0$.

In the other bulk, occupancy and dense ranges one has $s = O(\log r)$. Given the exposed labels, the probability of a failure to span among the $r - 1$ independent labels is at most 2^{s-r+1}. Changing the last label changes the empirical distribution by at most $2/r$ in distribution-function distance and each occupancy by at most one. The replacement affects one row and column of the sign matrix, so its operator cost is $O(\sqrt{r})$, smaller than the occupancy edge scale $\sqrt{r b_N}$ since $b_N \to \infty$. This gives the bulk and occupancy conclusions. In the dense range, every class is occupied with high probability, and the occupancy comparison gives the count of the large eigenvalues. For the growing-outlier estimate, the last label changes the relative occupancy deviations by $O(N/r)$ in norm, or $O(\sqrt{N})$ on the eigenvalue scale. This is $o(\sqrt{r \log N})$ in the dense range. The block edge law follows from Theorem M3.11 with $b = 1$ and $I_1 = A^c$, and the other estimates hold uniformly in the exposed labels and B.

Products of signs for the complete matrix. Suppose $s/(\log r \log\log r) \to \infty$. Before conditioning on span, a tuple of independent labels on A with the prescribed signs can be chosen in at least $\prod_{i=0}^{g-1}(2^{s-i} - 2^i)$ ways. At step i, the i linear conditions are independent and at most 2^i solutions are in the previous span. A tuple of rank $t < g$ can be chosen in at most

$$\binom{g}{t} 2^{ts - t(t-1)/2} 2^{t(g-t)}$$

ways, by choosing an independent subtuple and then choosing the other labels in its span. The ratio to the lower bound is at most $(1 + o(1))\binom{g}{t} 2^{-(s-2g)(g-t)}$. Conditioning on full span costs a factor of at most four in this ratio, since an independent exposed tuple has $s - g \le r - 1$ and extends to full span with probability greater than $1/4$. The probability that the exposed labels are dependent is thus $O(g 2^{2g-s}) = o(1)$. We may discard this event before taking moments.

Let W^A be a sign Wigner matrix whose internal block on A has been set equal to the prescribed block. For a product of signs, cancel even powers of each entry. Let u be the number of unexposed vertices still used. If $g + u < n$, before the spanning condition these u labels are independent uniforms, even if one is the sum of the others. The change in expectation under the spanning condition is at most $2\sqrt{5}\, 2^{(u-r+1)/2}$, by the proof of Proposition M2.2 applied to these labels. If any edge between two unexposed vertices remains, the quadratic part of the exponent has polar matrix of rank at least $2s$. The cross edges add only linear terms, so the expectation has absolute value at most 2^{-s}. If only cross and internal edges remain and there is a cross edge, the linear phase is nontrivial by independence of the exposed labels. Its average is zero. A product entirely inside A agrees exactly with that for W^A. Thus an arbitrary polynomial in the signs, using at most j unexposed vertices per monomial with $g + j < n$, has comparison error at most

$$\left(2^{-s} + 2\sqrt{5}\, 2^{(j-r+1)/2}\right) \sum_F |\widehat{P}(F)|,$$

where $\widehat{P}(F)$ are the coefficients after cancelling even powers.

Sparse edge locations. Let $H = K_b/\sqrt{r}$ and let C be the $g \times r$ cross block. For any fixed polynomial P the comparison just proved, applied to the squared Frobenius norm, gives

$$\left\| r^{-1} C P(H) C^{\mathsf{T}} - r^{-1} \operatorname{tr}(P(H)) I_g \right\| = o_{\mathrm{Pr}}(1).$$

Indeed the scalar polynomial for the squared Frobenius norm has coefficient sum $r^{O_P(1)}$, which the comparison error absorbs. In the Wigner reference, condition on the outside block H. The cross rows are independent sign vectors. Each centred entry of the displayed matrix has conditional second moment at most $2r^{-2}\operatorname{tr}(P(H)^2)$. The expected trace on the right is $O_P(r)$ by the fixed Wigner moment bounds, and $g^2/r = o(1)$.

In particular $CC^{\mathsf{T}}/r = I_g + o_{\mathrm{Pr}}(1)$ in operator norm. The large-prime block has semicircle law and both edges tending to ± 2 after normalization. For fixed $z > 2$, approximate $(z-x)^{-1}$ by a polynomial on $[-2-\varepsilon, 2+\varepsilon]$ with $0 < \varepsilon < z-2$. The bound for CC^{T}/r controls the error of this approximation after multiplication by C and C^{T}. The displayed comparison shows that $r^{-1}C(z-H)^{-1}C^{\mathsf{T}}$ converges in operator norm to $I_g \int (z-x)^{-1}\, d\nu_0(x)$. This integral is at most 1. Since $\|K_a\|/\sqrt{r} \le g/\sqrt{r} = o(1)$, the Schur complement of $zI - H$ in $zI - K/\sqrt{r}$ is positive with probability tending to one. Apply the same argument to $-K$ for the other bound. Interlacing with K_b then gives both locations.

The two edge processes. Suppose $s/\{r^{1/3}(\log r)^{3/2}\} \to \infty$. We first compare $W^A/\sqrt{n}$ with an ordinary sign Wigner matrix $V = W_n/\sqrt{n}$, using the same signs outside A. Their difference D_* is supported on $A \times A$ and has norm at most $2g/\sqrt{n}$. The local semicircle law [1, Theorem 2.1] at imaginary part $\eta = n^{-1/3}$ gives

$$\|V\| \le 3, \qquad \max_{i,j} \sup_{|E| \le 4} |(V - E - i\eta)^{-1}_{ij}| = O(1)$$

except on a set of probability $O(\exp(-(\log n)^3))$. To apply it, normalize the row variance to one with $\sqrt{n-1}$. The variance matrix has eigenvalues 1 and $-1/(n-1)$ and all entries at most $1/(n-1)$. The bounded entry tails permit the stated probability, by choosing L a sufficiently large multiple of $\log\log n$ in that theorem. Changing $\sqrt{n-1}$ back to $\sqrt{n}$ preserves the estimates.

Let G_A denote the restriction of the resolvent to $A \times A$. It has norm $O(g)$. The restriction for $V + tD_*$ is $(I + tG_A(D_*)_{AA})^{-1}G_A$, of norm $O(g)$ uniformly in $0 \le t \le 1$, since $g^2/\sqrt{n} = o(1)$. For any unit eigenvector v of $V + tD_*$ take E to be its eigenvalue. The spectral expansion of the imaginary part bounds $\eta^{-1}(v|_A)(v|_A)^{\mathsf{T}}$ by that imaginary part in positive semidefinite order. Hence $\|v|_A\|^2 \ll g\eta$. Each ordered eigenvalue is Lipschitz in t. At differentiability points its derivative is bounded by the maximum of $|v^{\mathsf{T}} D_* v|$ on its eigenspace, by the min–max principle. It follows that on this event all shifts are $O(g^2 n^{-5/6}) = o(n^{-2/3})$.

For powers $p \asymp n^{2/3}$ the mean-value theorem now makes the difference of the traces of $(V/2)^p$ and $(W^A/(2\sqrt{n}))^p$ tend to zero in every fixed moment. On the good event, for $|x| \ge 1/2$ the error in x^p is at most $O(pd)e^{2pd}|x|^p$, where $d = O(g^2 n^{-5/6})$. The terms with $|x| < 1/2$ are exponentially small. The absolute power sums have bounded fixed moments by the Wigner high-moment bounds in Section S5, using adjacent even powers for odd p. On the exceptional event, the law of W^A is the Wigner law conditioned on an event of probability $2^{-\binom{g}{2}}$. Allowing that factor in a higher moment and using Cauchy–Schwarz with the exceptional probability still gives $o(1)$.

Use the polynomials of Section S5 with size n. A fixed product of compressed traces, or one of the confinement polynomials, has degree $j = O(n^{1/3}\sqrt{\log n})$ and coefficient sum at most $\exp(O(j \log n))$. The sign comparison above therefore replaces the expectation under the quotient law by that for W^A, with error $o(1)$. The exterior polynomial bounds of Lemma S5.1 and Lemma S5.3 use upper

bounds on even high moments. These hold for W^A by the preceding comparison with V. The compressed mixed moments consequently have the same limits as in Section S5, and the same confinement bounds hold. The transfer of factorial Laplace transforms in Lemma S5.4 gives the two independent GOE processes. The probability $o(1)$ of dependent exposed labels does not affect convergence in distribution.

The edges when labels repeat. Here $N \to \infty$ and $\lambda \geq (\log N)^{-\beta}$, with $0 < \beta < 1$. First take r independent free labels and add the exposed counts c_x, supported on a set S of at most g labels. Write ℓ for the largest combined count. The free maximum is asymptotic to b_N by the independent occupancy calculation in the main article. The binomial tail bound, summed over S, gives free counts there equal to

$$\lambda + O_{\Pr}(\sqrt{\lambda \log\log N} + \log\log N) = \lambda + o_{\Pr}(b_N) \qquad (\lambda \leq (\log N)^2).$$

Here $g = O(\log N)$ on this range, and b_N is at least λ and a constant times $\log N/\log\log N$. Thus $\ell = (1 + o_{\Pr}(1)) \max\{b_N, \lambda + \max_x c_x\}$. For $\lambda \geq (\log N)^2$, all free counts are asymptotic to λ uniformly, and $g = O(\log N + \log\lambda) = o(\lambda)$.

For the spectral estimate when $\lambda \leq (\log N)^2$, Poissonize the free counts as in the proof of the occupancy theorem in the main article. If H_0 is the set of exceptional Poisson counts in that proof, enlarge it to $H = H_0 \cup S$. The same size bound $|H|/N \ll (\log N)^{-8}$ holds with the same probability. The flat entries of U give

$$\|U_{HH}\| \leq \|U_{H_0H_0}\| + 2\sqrt{g|H_0|/N} + g/\sqrt{N} = O((\log N)^{-7/2}).$$

The counts on H^c have no exposed contribution and retain the independent truncated laws of the original proof, conditionally on H_0. For $\lambda \leq \sqrt{\log N}$ the Bernstein estimate for $P_H U D_{H^c} U P_H$ therefore gives $\lambda(I_H + o(1))$. The coupling block has norm $(1 + o(1))\sqrt{\lambda\ell}$. The largest count still lies in H, and the diagonal block errors are $o(\sqrt{\lambda\ell})$: the high block is bounded by $\ell\|U_{HH}\|$ and the low block has the original truncation bound. In particular $\ell = O(\log N + \lambda)$ on the operator event. For $\sqrt{\log N} \leq \lambda \leq (\log N)^2$, replace the low counts by λ as in the main article. The same involution decomposition gives both extrema asymptotic in magnitude to $\sqrt{\lambda\ell}$, because the error $O(\ell\|U_{HH}\|)$ is $o(\sqrt{\lambda\ell})$. Each operator estimate fails with probability $O(N^{-3})$ before conditioning the total free count to be r, so it holds with probability tending to one after conditioning. For $\lambda \geq (\log N)^2$, near-uniform counts give the estimate directly by $D^{1/2} U D^{1/2} = \lambda U + o_{\Pr}(\lambda)$.

Multiplying the quotient eigenvalues by $\sqrt{N}$ now gives the two extreme magnitudes $(1 + o_{\Pr}(1))\sqrt{r\ell}$. Replacing one free row by the required sum and removing the sign diagonal costs $O(\sqrt{n})$ in operator norm, below this scale since $b_N \to \infty$. The first $r - 1$ free labels fail to span with probability at most $N2^{-r+1} = o(1)$, so conditioning on full span changes the coupled law by $o(1)$ in total variation.

Declarations

No external funding was received. The author declares no competing interests. OpenAI Codex assisted with ideas, checking arguments and references, computations, and drafting. Computational checks do not replace the proofs; the author takes responsibility for the article. Anthropic Claude also assisted with project-internal reviews and numerical checks. These checks are not external peer review or formal verification.

References for the supplementary proofs

[1] L. Erdős, H.-T. Yau and J. Yin, *Rigidity of eigenvalues of generalized Wigner matrices*, Adv. Math. **229** (2012), 1435–1515. arXiv:1007.4652.

[2] J. Friedlander and H. Iwaniec, *Ternary quadratic forms with rational zeros*, J. Théor. Nombres Bordeaux **22** (2010), 97–113. doi:10.5802/jtnb.706.

[3] Z. Gao, *Four-ranks of class groups and the spectrum of Legendre symbols*, accompanying main article (2026).

[4] M. Krivelevich and B. Sudakov, *Pseudo-random graphs*, in *More Sets, Graphs and Numbers*, Bolyai Soc. Math. Stud. **15**, Springer, 2006, 199–262. arXiv:math/0503745.

[5] S. Sachdeva and N. K. Vishnoi, *Faster Algorithms via Approximation Theory*, Found. Trends Theor. Comput. Sci. **9** (2013), 125–210. DOI: 10.1561/0400000065. Section 3.

[6] A. Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys. **207** (1999), 697–733. DOI: 10.1007/s002200050743.

[7] C. A. Tracy and H. Widom, *Matrix kernels for the Gaussian orthogonal and symplectic ensembles*, Ann. Inst. Fourier **55** (2005), 2197–2207. Published article.