COMPARISON THEOREM FOR SOME EXTREMAL EIGENVALUE STATISTICS

BY BENJAMIN LANDON[1], PATRICK LOPATTO[2,*] AND JAKE MARCINEK[2,†]

[1]*Department of Mathematics, Massachusetts Institute of Technology, blandon@mit.edu*
[2]*Department of Mathematics, Harvard University, *lopatto@math.harvard.edu; †marcinek@math.harvard.edu*

We introduce a method for the comparison of some extremal eigenvalue statistics of random matrices. For example, it allows one to compare the maximal eigenvalue gap in the bulk of two generalized Wigner ensembles, provided that the first four moments of their matrix entries match. As an application, we extend results of Ben Arous–Bourgade and Feng–Wei that identify the limit of the maximal eigenvalue gap in the bulk of the GUE to all complex Hermitian generalized Wigner matrices.

There has been significant progress in understanding local eigenvalue statistics of random matrices over the past decade. A fundamental example is the proof of bulk universality for Wigner matrices established in the series of works [9, 11, 12, 14, 16, 19]. It states that in the high dimensional limit, the local eigenvalue fluctuations are universal and depend only on the symmetry class of the random matrix ensemble. Parallel results were established in certain cases in [33, 34], including a "four moment theorem," which shows that the local statistics are determined by the first four moments of the entries. By local eigenvalue statistics, we refer to the fluctuations of the eigenvalues in the bulk (the interior of the spectrum) on the interparticle scale N^{-1}.

The *generalized Wigner ensembles* constitute a fundamental class of random matrix ensembles. These are self-adjoint $N \times N$ random matrices whose entries above the diagonal are independent centered random variables. The variances of the entries are taken to be order N^{-1}, with the constraint that the variances of each row sum to 1. Examples include the Gaussian Orthogonal Ensemble and Gaussian Unitary Ensemble (GOE/GUE), whose entries are real and complex Gaussians, respectively. In these cases, arguments based on orthogonal polynomials show that the asymptotic local statistics are described by explicit formulas. Universality implies that these formulas describe the asymptotic local statistics all generalized Wigner ensembles.

Examples of local statistics typically studied are the distribution of a single gap between consecutive eigenvalues, the local correlation functions at an energy $E \in (-2, 2)$ in the bulk, and averaged versions of these quantities. In this work, we consider certain extremal eigenvalue statistics. In particular, we study the maximal gap between consecutive eigenvalues in the bulk. To place our main technical contribution in context, we briefly review the steps for proving bulk universality for generalized Wigner matrices. We refer to [15, 17] for a comprehensive review and bibliography.

The first step is dynamical and originates in the work of Erdős, Schlein and Yau [14]. It is based around analyzing the local time to equilibrium of the matrix eigenvalues under a stochastic matrix dynamics known as *Dyson Brownian motion* (DBM). The equilibrium measure of this dynamics is the GOE or GUE, depending on the symmetry class under consideration. It is now known that local observables reach their equilibrium distribution under

the DBM flow after only a short time. At the level of matrices, this result proves that the local statistics of matrix ensembles with small Gaussian perturbations coincide with those of the Gaussian ensembles. It therefore establishes that bulk universality holds for the former matrices, which are known as *Gaussian divisible ensembles*.

The second step, extending universality from Gaussian divisible ensembles to all generalized Wigner matrices, is a density argument. One approach is the four moment method of Tao and Vu [33, 34]. Their work implies that the local eigenvalue statistics of two matrix ensembles with four matching moments coincide as $N \to \infty$, which yields universality for ensembles whose moments match those of the GOE or GUE. In general, it is possible to construct Gaussian divisible ensembles matching four moments of a given generalized Wigner ensemble. Combining this fact and the four moment method with the first step, we obtain universality for all generalized Wigner matrices. Other approaches to this step include the Green's function comparison theorem [18], the matrix continuity estimate of [5] and the reverse heat flow of [12].

We refer to the two steps of this approach as the dynamical step and the comparison step, and now consider how they may be applied to prove universality of the maximal eigenvalue gap. For the dynamical part, the works [4, 27] employ a coupling of Dyson Brownian motion to an auxiliary process which is already at equilibrium. They show that the ith eigenvalue gap of the DBM is the same as the ith gap of the equilibrium process, down to the scale N^{-1-c}. The work [27] proved further that this estimate holds with a probability large enough to permit a union bound over all N eigenvalues. This is sufficient to obtain universality of the maximal gap distribution for Gaussian divisible ensembles. We remark that the polynomial error rate $c > 0$ obtained in this work is not optimal, and in particular, it is not sufficient to treat the minimal eigenvalue gap.

For the comparison step, the reverse heat flow provides strong estimates, but requires that the matrix entry distributions are smooth, an assumption we would like to avoid. The other approaches are based on the Lindeberg moment matching strategy and apply either to individual eigenvalues or resolvent entries. They are local in that they work only near a spectral energy E or for functions of a few eigenvalues. However, the maximal gap is a statistic involving a macroscopic number of eigenvalues distributed throughout the spectrum. Therefore, a new approach is needed in order to handle the maximal gap and other extremal spectral statistics for ensembles with discrete distributions.

Our contribution is to provide comparison theorems for the maximal gap between consecutive eigenvalues in the bulk (and certain other extremal statistics). These theorems extend universality for the largest eigenvalue gaps from the Gaussian divisible ensembles to general matrices. As a consequence, we show that the results of Ben Arous–Bourgade [1] and Feng–Wei [21] on the distribution of the maximal gaps of the GUE in fact hold for all Hermitian generalized Wigner matrices; we describe these more below. Our results can also be viewed as providing a four-moment theorem for extremal spectral statistics in the spirit of Tao and Vu [33, 34].

At a technical level, we construct a regularization of the maximal eigenvalue gap that is amenable to comparison. The four-moment approach is based around the Lindeberg moment matching strategy, which requires estimates on the derivatives of the regularization with respect to the matrix entries. The application of the Lindeberg strategy to random matrices originates in the work of Chatterjee [6]. Our derivative estimates also allow us to extend the matrix continuity approach of [5] to maximal eigenvalue gaps.

The regularization we use for the maximum is

$$(0.1) \qquad \max_i x_i \sim \frac{1}{\beta} \log\left(\sum_i e^{\beta x_i}\right)$$

for appropriately chosen $\beta \to \infty$ as $N \to \infty$. This regularization was used previously in random matrix theory in conjunction with the Lindeberg strategy by Korada and Montanari [26] to show universality of the minima of loss functions associated to certain statistical algorithms; this is otherwise unrelated to our work. Our inspiration for this regularization comes from statistical mechanics, where it represents the "zero temperature limit" of a Gibbs measure. (In our notation, β is the inverse temperature.) This interpretation plays a key role in many areas of mathematics inspired by statistical mechanics, for example, in the study of random constraint satisfaction problems through the theory of mean field spin glasses.

We do not apply the regularization (0.1) directly to the eigenvalue gaps. Derivatives of eigenvalues with respect to matrix entries are singular as they involve in the denominator the differences of nearby eigenvalues. Instead, we first construct a regularized version of the eigenvalues using the matrix resolvent in combination with the Helffer–Sjöstrand formula. Part of this construction is essentially implicit in the work of Knowles and Yin [25]. Our use of the Helffer–Sjöstrand formula simplifies many of the estimates, and we feel there is value in isolating the relevant arguments and presenting the construction as its own lemma. More-over, we give a precise estimate on the relation between the accuracy of the regularization and the growth of its derivatives with respect to matrix entries. Roughly, if the regularization is chosen so that the error between it and the eigenvalue is less than $N^{-1-\delta}$, then the kth derivative is no larger than $N^{-1+(k-1)\delta}$. Tracking this dependence may be useful for future works probing the eigenvalue behavior below the interparticle scale N^{-1}.

We now turn to related results in the literature. The problem of determining the distribution of the largest gap between eigenvalues was raised for random unitary matrices by Diaconis [8]. As motivation, he gives a striking conjecture relating the distribution of the largest gaps between the unitary eigenvalues to extremal gaps between zeros of the Riemann zeta function [32]. For more on this connection, we refer to the reader to [1], Section 1.3.

A heuristic approach to studying the maximal gap of random unitary matrices was given by Vinson, who obtained the correct order of magnitude [35]. In [1], Ben Arous and Bour-gade proved for both unitary matrices and the GUE that the maximal gap, normalized by $\sqrt{\log(N)}/N$, converges to an explicit constant (which depends on the ensemble). (For the GUE, they restrict to gaps in the bulk of the spectrum.) Feng and Wei [21] found that, up to a deterministic recentering, the maximal gap of both ensembles fluctuates on the scale $(N\sqrt{\log(N)})^{-1}$ around its limit and identified the limiting distributions. We determine the maximal gap down to the scale N^{-1-c} for some $c > 0$, and so our work extends the GUE results to all Hermitian generalized Wigner ensembles. In another direction, we note Figalli and Guionnet extended the results of [1] to β-ensembles with $\beta = 2$ [23].

Our paper is related to recent work of Bourgade [3]. He gives a new approach to the analysis of Dyson Brownian motion and obtains strong estimates enabling access not only to the maximal gap, but also the *minimal* gap, a much more singular quantity which is beyond the methods of our work. The work [3] relies on the reverse heat flow for the comparison step. Our works are therefore complementary, as we focus on proving comparison theorems and do not study the dynamical side. In particular, our comparison theorems apply to ensembles whose matrix entries are discrete random variables, which are outside the scope of the reverse heat flow technique.

We note that other previous works have also studied the minimal eigenvalue gap. Vinson obtained the limit of the smallest eigenvalue gap for both the Circular Unitary Ensemble and GUE in his Ph.D. thesis [35]; multiple smallest gaps of these ensembles were then considered in [1]. Feng, Tian and Wei have established the behavior of the smallest gaps of the Gaussian Orthogonal Ensemble; this work builds on earlier work of Feng and Wei on the circular β-ensembles [20, 22]. Other works have studied Wigner [13, 31] and sparse [29, 30] matrices, obtaining tail bounds for the size of the gap. Showing universality for the minimal gap of

Hermitian generalized Wigner matrices without imposing a smoothness assumption on the entries remains an intriguing open problem.

Another interesting problem is to compute the normalized limit of the maximal bulk eigenvalue gap for the GOE. Unfortunately, the argument of [1] for the GUE relies on its determinantal structure, which is not present for the GOE. If the normalized limit of the maximal bulk gap could be established for the GOE, our arguments would immediately show it is the same for all real symmetric generalized Wigner matrices. One could also ask analogous questions about extremal gaps for β-ensembles with arbitrary β and potential.

The remainder of the paper is organized as follows. In Section 1, we define the eigenvalue statistics we consider and state our main comparison results and universality corollaries. In Section 2, we prove some of our comparison theorems. Section 3 contains an eigenvalue regularization lemma based on the method of Knowles and Yin [25]. In Section 4, we prove the remainder of our comparison theorems. Section 5 and the Appendix are devoted to deducing the universality corollaries from existing results in the literature and our comparison results.

1. Main result.

1.1. *Definition of model.* We begin by defining generalized Wigner matrices.

DEFINITION 1.1. A generalized Wigner matrix H_N is a real symmetric or complex Hermitian $N \times N$ matrix whose upper triangular elements $\{h_{ij}\}_{i \le j}$ are independent random variables with mean zero and variances $\sigma_{ij}^2 = \mathbb{E}(|h_{ij}|^2)$ that satisfy

$$(1.1) \qquad \sum_{i=1}^{N} \sigma_{ij}^2 = 1 \quad \text{for all } j \in [\![1, N]\!]$$

and

$$(1.2) \qquad \frac{c}{N} \le \sigma_{ij}^2 \le \frac{C}{N} \quad \text{for all } i, j \in [\![1, N]\!]$$

for some constants $c, C > 0$.

When H_N is Hermitian, we further assume

$$(1.3) \qquad c \operatorname{Var} \operatorname{Re} h_{ij} \le \operatorname{Var} \operatorname{Im} h_{ij} \le C \operatorname{Var} \operatorname{Re} h_{ij}, \quad i \ne j$$

and that $\operatorname{Re} h_{ij}$ and $\operatorname{Im} h_{ij}$ are independent.

Finally, we suppose that the normalized entries have finite moments, uniformly in N, i, and j, in the sense that for all $p \in \mathbb{N}$ there exists a constant μ_p such that

$$(1.4) \qquad \mathbb{E} \left| \frac{h_{ij}}{\sigma_{ij}} \right|^p \le \mu_p$$

for all N, i and j.

Throughout this work, we suppress the dependence of various constants in our results on the constants in this definition. This dependence does not affect our arguments in any substantial way.

We remark that the moment condition on the entries is technical and used only to get convergence of the largest gaps in the topology of L^p for every p in (1.12) below. If one desires only convergence in some L^p for fixed p, then the requirement can be relaxed to requiring that the matrix entries have a certain large moment.

We also use the following notion of *overwhelming probability*.

DEFINITION 1.2.   We say that a set of events $\{\mathcal{A}(u)\}_{u \in U^{(N)}}$, where $U^{(N)}$ is a parameter set which may depend on N, holds with *overwhelming probability* if, for any $D > 0$, there exists $N(D, U^{(N)})$ such that for $N \geq N(D, U^{(N)})$,

$$(1.5) \qquad \inf_{u \in U^{(N)}} \mathbb{P}(\mathcal{A}(u)) \geq 1 - N^{-D}.$$

1.2. *Comparison results.*   We first present a theorem on the comparison of the largest eigenvalue gaps of generalized Wigner matrices with matching moments.

We denote eigenvalues of matrices $\lambda_1 \leq \cdots \leq \lambda_N$ and use $\boldsymbol{\lambda}$ to denote the vector of N eigenvalues. The statistics we consider involve the maxima of eigenvalue gaps away from the spectral edges. Taking such a maximum requires the specification of which eigenvalues we take the maximum over. We do this in two different ways. The first statistic, which we denote by $\mathcal{T}_{\ell, J}$, considers gaps $\lambda_{i+1} - \lambda_i$ with $i \in J$ where J is a fixed index set. The second, which we denote by $\widehat{\mathcal{T}}_{\ell, I}$, considers gaps with $\lambda_i \in I$ where I is an interval. We first consider $\mathcal{T}$ before considering $\widehat{\mathcal{T}}$ later. These statistics are closely related, but dealing with $\widehat{\mathcal{T}}$ is more difficult than $\mathcal{T}$ due to the fact that the eigenvalue fluctuations naturally cause the set of indices over which the maximum is taken to itself be random.

For any index set $J \subseteq [\![1, N-1]\!]$ we denote by $\mathcal{T}_{\ell, J}(\boldsymbol{\lambda})$ the ℓth largest eigenvalue gap in the index set J, formally defined by

$$\sup_{j \in J} \{ \lambda_{j+1} - \lambda_j : \exists j_1 \neq \cdots \neq j_{l-1} \in J \text{ and } \lambda_{j_1+1} - \lambda_{j_1} \geq \cdots \geq \lambda_{j_{l-1}+1} - \lambda_{j_{l-1}} \geq \lambda_{j+1} - \lambda_j \}.$$

The proofs of the next two theorems appear in Section 2.

THEOREM 1.3.   *Let $H^{(v)}$ and $H^{(w)}$ be two generalized Wigner matrices with matrix elements denoted by v_{ij} and w_{ij} respectively. Suppose that there exists a constant $c > 0$ such that*

$$(1.6) \qquad |\mathbb{E}[v_{ij}^a \bar{v}_{ij}^b] - \mathbb{E}[w_{ij}^a \bar{w}_{ij}^b]| \leq N^{-2-c}$$

for all nonnegative integers a, b such that $a + b \leq 4$. Fix $\alpha \in (0, 1/2)$ and choose some index set $J \subset [\![\alpha N, (1-\alpha)N]\!]$. Let $S \in C^{\infty}(\mathbb{R})$ be a test function, and let $\nu = N/\sqrt{\log(N)}$. Suppose $\ell = \ell(N)$ satisfies $\ell = N^{a_N}$ for some sequence $a_N \geq 0$. Then there exists constants $c_1 = c_1(c) > 0$ and $C = C(c, \alpha) > 0$ such that if $a_N \leq c_1$, then[1]

$$(1.7) \qquad |\mathbb{E}_{H^{(v)}} S(\nu \mathcal{T}_{\ell, J}(\boldsymbol{\lambda})) - \mathbb{E}_{H^{(w)}} S(\nu \mathcal{T}_{\ell, J}(\boldsymbol{\lambda}))| \leq C \Big(\sup_{0 \leq d \leq 5} \|S^{(d)}\|_{\infty} \Big) N^{-c_1}.$$

The following interpolation, called Dyson Brownian motion, is often used in random matrix theory. We define for any real symmetric generalized Wigner matrix X the matrix Ornstein–Uhlenbeck process $X_t = \{x_{ij}(t)\}_{i,j=1}^N$ given by

$$(1.8) \qquad dx_{ij}(t) = \frac{dB_{ij}(t)}{\sqrt{N}} - \frac{1}{2Ns_{ij}} x_{ij}(t)\, dt,$$

where the initial data are the entries $\{x_{ij}\}$ of X and $s_{ij} = \mathbb{E}[x_{ij}^2]$. We let $\boldsymbol{\lambda}(t)$ denote the vector of eigenvalues of X_t.

For a Hermitian generalized Wigner matrix, we define Dyson Brownian motion by applying (1.8) separately to its real and imaginary parts, with corresponding rescaling parameters $s_{ij}^{\mathbb{R}} = \mathbb{E}[(\mathrm{Re}\, x_{ij})^2]$ and $s_{ij}^{\mathbb{C}} = \mathbb{E}[(\mathrm{Im}\, x_{ij})^2]$ for the dt terms.

---

[1]We did not work out the optimal dependence of c_1 on c, but as indicated in the proof below, it suffices to take $c_1 \leq c/100$.

THEOREM 1.4. *Consider the SDE* (1.8) *for a matrix X_t with initial data given by some generalized Wigner matrix X. Suppose $t \in (N^{-1+\delta}, N^{-1/2-\delta})$ for some $\delta > 0$. Fix $\alpha \in (0, 1/2)$ and choose some index set $J \subset [\![\alpha N, (1-\alpha)N]\!]$. Let $S \in C^\infty(\mathbb{R})$ and set $v = N/\sqrt{\log N}$. Suppose $\ell = \ell(N)$ satisfies $\ell = N^{a_N}$ for some sequence $a_N \geq 0$. Then there exist constants $c = c(\delta) > 0$ and $C = C(\delta, \alpha) > 0$ such that if $a_N < c$, then*

$$(1.9) \qquad \left| \mathbb{E}\big[S(v\mathcal{T}_{\ell,J}(\boldsymbol{\lambda}(t)))\big] - \mathbb{E}\big[S(v\mathcal{T}_{\ell,J}(\boldsymbol{\lambda}(0)))\big] \right| \leq C \Big(\sup_{0 \leq d \leq 3} \|S^{(d)}\|_\infty \Big) N^{-c}.$$

With more work, we can extend the above comparison theorems to the following related extremal statistic. Let $I \subseteq (-2+\kappa, 2-\kappa)$ be an interval with $\kappa > 0$. For $v_1 \leq v_2 \cdots \leq v_N$ with $\boldsymbol{v} \in \mathbb{R}^N$, let

$$(1.10) \qquad \widehat{\mathcal{T}}_{1,I}(\boldsymbol{v}) = \max_{i:v_i \in I} (v_{i+1} - v_i),$$

and define similarly $\widehat{\mathcal{T}}_{\ell,I}(\boldsymbol{v})$ to be the ℓth largest gap $v_{i+1} - v_i$ such that $v_i \in I$. Then we have the following analog of the above theorems. Its proof is deferred until Section 4.

THEOREM 1.5. *First, suppose that $H^{(v)}$ and $H^{(w)}$ are two generalized Wigner matrices such that the first four moments of the matrix entries match in the sense of* (1.6). *Then the estimate* (1.7) *holds for some constants $C, c_1 > 0$ with $\widehat{\mathcal{T}}_{\ell,I}$ in place of $\mathcal{T}_{\ell,J}$ (which depend only on c from* (1.6) *and κ).*

Second, let X be a generalized Wigner matrix and X_t be the process in (1.8). *Let $t \in (N^{-1+\delta}, N^{1/2-\delta})$ for some $\delta > 0$. Then the estimate* (1.9) *holds with $\widehat{\mathcal{T}}_{\ell,I}$ in place of $\mathcal{T}_{\ell,J}$ for some constants $C, c > 0$ (which depend only on δ and κ).*

In particular, for a fixed κ, these constants are uniform in the choice of $I \subseteq (-2+\kappa, 2-\kappa)$.

We remark that the method we introduce is fairly robust and can handle a variety of extremal spectral statistics. For example, in light of the result [1], Theorem 1.7, one may wish to scale each gap by the semicircle density near λ_i, in addition to the factor v. Our proof applies without change to the maxima over quantities such as $\{v\alpha_i(\lambda_{i+1} - \lambda_i)\}_i$ for deterministic positive α_i satisfying $\alpha_i \leq N^c$ for a sufficiently small $c > 0$. Additionally, one may also consider the deviations $\{v\alpha_i(\lambda_i - \beta_i)\}_i$ for similar α_i and arbitrary constants β_i (or their absolute values).

In certain cases, it may be necessary to use the four-moment approach instead of the matrix continuity estimate, if one can prove universality only for Gaussian divisible ensembles with somewhat larger Gaussian components. This means using Theorem 1.3 instead of Theorem 1.4 (or the first claim of Theorem 1.5 instead of the second); compare the restriction $t \ll N^{-1/2}$ with the weaker (1.6).

1.3. *Corollaries for the universality of extremal gaps.* The following is a corollary of Theorem 1.4 and the homogenization result [27], Theorem 3.1.

COROLLARY 1.6. *Let H be a generalized Wigner matrix and G the GOE or GUE matrix of the same symmetry class. There are constants $C, c > 0$ such that*

$$(1.11) \qquad \left| \mathbb{E}_H\big[S(v\mathcal{T}_{\ell,J}(\boldsymbol{\lambda}))\big] - \mathbb{E}_G\big[S(v\mathcal{T}_{\ell,J}(\boldsymbol{\lambda})\big] \right| \leq CN^{-c} \sup_{0 \leq d \leq 3} \|S^{(d)}\|_\infty.$$

A similar estimate holds for $\widehat{\mathcal{T}}_{\ell,I}$.

Together with the results of Ben Arous–Bourgade ([1], Theorem 1.7) and Feng–Wei ([21], Theorem 2), it implies the following corollary.

COROLLARY 1.7. *Suppose H is a Hermitian generalized Wigner matrix. Let $I \subset (-2, 2)$ be a compact interval, and set $M = M(I) = \inf_I \sqrt{4 - x^2}$.*

1. *Let $\ell_N = N^{k_N}$ be a sequence of positive integers with $\lim_{N \to \infty} k_N = 0$. Then for any $p > 0$,*

$$(1.12) \qquad \frac{MN}{\sqrt{32 \log N}} \widehat{\mathcal{T}}_{\ell_N, I} \xrightarrow{L^p} 1.$$

2. *Fix an index k. Let*

$$(1.13) \qquad \tau_k^* = \frac{1}{4} (2 \log(N))^{\frac{1}{2}} (MN \widehat{\mathcal{T}}_{k,I} - \sqrt{32 \log(N)}) + \frac{5}{8} \log(2 \log(N)).$$

Then for any bounded interval $I_1 \subseteq \mathbb{R}_+$ we have

$$(1.14) \qquad \mathbb{P}(\tau_k^* \in I_1) \to \int_{I_1} \frac{e^{k(c_2 - x)}}{(k - 1)!} e^{-e^{c_2 - x}} \, dx,$$

where c_2 is an explicit constant depending on the interval. It is the same as in [21], Theorem 2.

The proofs of these corollaries appear in Section 5.

1.4. *Preliminaries.* In this section, we recall some standard facts about generalized Wigner matrices. The Stieltjes transform of H is defined by

$$(1.15) \qquad m_N(z) = \frac{1}{N} \sum_i \frac{1}{\lambda_i - z}$$

and the semicircle law and corresponding Stieltjes transform are

$$(1.16) \qquad \rho_{\mathrm{sc}}(E) = \frac{\sqrt{(4 - E^2)_+}}{2\pi}, \qquad m_{\mathrm{sc}}(z) = \int \frac{\rho_{\mathrm{sc}}(x) \, dx}{x - z}.$$

The Green's function of H is

$$(1.17) \qquad G(z) = \frac{1}{H - z},$$

and its matrix elements are denoted $G_{ij}(z)$.

We also define the spectral domain

$$(1.18) \qquad \mathcal{D} = \left\{ z = E + i\eta \in \mathbb{C} : |E| \leq 10, \frac{N^\delta}{N} \leq \eta \leq 10 \right\}.$$

We have the following from [10].

THEOREM 1.8 ([10], Theorem 2.3). *Fix $\varepsilon > 0$ and $\delta > 0$. Then, with overwhelming probability, we have*

$$(1.19) \qquad \sup_{z \in \mathcal{D}} |m_N(z) - m_{\mathrm{sc}}(z)| \leq \frac{N^\varepsilon}{N\eta}.$$

For the individual Green's function elements, with overwhelming probability we have

$$(1.20) \qquad \sup_{z \in \mathcal{D}} |G_{ij}(z) - \delta_{ij} m_{\mathrm{sc}}(z)| \leq N^\varepsilon \left(\sqrt{\frac{\operatorname{Im} m_{\mathrm{sc}}(z)}{N\eta}} + \frac{1}{N\eta} \right).$$

The classical eigenvalue locations of the semicircle law are denoted by γ_i and defined by

$$(1.21) \qquad \frac{i}{N} = \int_{-2}^{\gamma_i} \rho_{sc}(x)\,dx.$$

We next state eigenvalue rigidity and complete eigenvector delocalization estimates. The former is [10], Theorem 7.6, while the latter follows by a standard argument from the Green's function estimates in Theorem 1.8 (see, e.g., the proof of [2], Theorem 2.10).

THEOREM 1.9 ([10], Theorem 7.6). *The following estimates hold for a generalized Wigner matrix H, simultaneously for all $i \in [\![1, N]\!]$, with overwhelming probability. For any $\varepsilon > 0$,*

$$(1.22) \qquad |\lambda_i - \gamma_i| \leq \frac{N^\varepsilon}{N^{2/3}(\min\{i^{1/3}, (N-i+1)^{1/3}\})},$$

and for the eigenvector u_i of H corresponding to λ_i,

$$(1.23) \qquad \|u_i\|_\infty \leq \frac{N^\varepsilon}{\sqrt{N}}.$$

2. Maximal gap over a set of indices.

In this section, we prove the comparison results, Theorems 1.3 and 1.4. Both are based on defining a suitable regularization of the kth largest gaps and proving estimates on the partial derivatives of this regularization with respect to matrix entries. Theorem 1.3 is then based on the Lindeberg strategy, while Theorem 1.4 is based on the matrix continuity estimate of [5], Lemma A.1. The real symmetric and complex Hermitian cases are nearly identical, and we give full details in only the symmetric case as it is notationally simpler.

Throughout this section, we fix an $\alpha \in (0, 1/2)$ and only consider index sets $J \subset [\![\alpha N, (1-\alpha)N]\!]$. We first begin by introducing our regularization of the L^∞ norm. Given a vector $v \in \mathbb{R}^N$, define the associated largest bulk gap by

$$(2.1) \qquad \mathcal{T}_1 = \mathcal{T}_{1,J}(v) = \sup_{i \,:\, i \in J} v_{i+1} - v_i.$$

For general ℓ, we recall that $\mathcal{T}_{\ell,J}(v)$ is defined as the ℓth largest gap $v_{i+1} - v_i$ in consecutive elements of v with $i \in J$ (or zero if ℓ is greater than the number of i with $i \in J$). It is convenient to write this quantity as

$$(2.2) \quad \mathcal{T}_\ell(v) = \mathcal{T}_{\ell,J}(v) = \sup_{i_1 < \cdots < i_\ell \,:\, i_k \in J} \sum_{k=1}^{\ell}(v_{i_k+1} - v_{i_k}) - \sup_{i_1 < \cdots < i_{\ell-1} \,:\, i_k \in J} \sum_{k=1}^{\ell-1}(v_{i_k+1} - v_{i_k}).$$

Set $\nu = N/\sqrt{\log N}$, which represents the scale of the largest gap. For $\beta > 0$, we introduce

$$(2.3) \qquad \begin{aligned} Z = Z_{\ell,\beta,J}(v) &= \sum_{i_1 < \cdots < i_\ell \,:\, i_k \in J} \exp\left(\beta\nu \sum_{k=1}^{\ell} v_{i_k+1} - v_{i_k}\right), \\ G = G_{\ell,\beta,J}(v) &= \frac{1}{\beta}\log Z, \end{aligned}$$

where we use the convention $G_{0,\beta,J}(v) = 0$. We also set

$$(2.4) \qquad F_{\ell,\beta,J}(v) = G_{\ell,\beta,J}(v) - G_{\ell-1,\beta,J}(v).$$

The quantity $F_{\ell,\beta,J}(v)$ is our regularization of the ℓth largest gap of the vector v. The following lemma is elementary and its proof is omitted.

LEMMA 2.1.   *For any $v \in \mathbb{R}^n$ and index set J, we have*

$$(2.5) \qquad\qquad \left| v\mathcal{T}_\ell(v) - F_{\ell,\beta}(v) \right| < \frac{2\ell \log N}{\beta}.$$

Instead of directly taking v to be the eigenvalues of a random matrix, we are going to work with a regularized version, which we denote by the vector $\tilde{\lambda} = (\tilde{\lambda}_i)$. The construction of this regularization is somewhat involved, but we will see that $\tilde{\lambda}_i$ is primarily an integral of the empirical Stieltjes transform $m_N(z)$ over a certain domain in the complex plane. We state the following lemma which asserts the existence of $\tilde{\lambda}$ and gives the key estimates on its derivatives that are used in the present section; its proof is deferred until Section 3.

Given a matrix H, $0 \le \theta \le 1$, and indices c, d, we denote by $\theta^{cd} H$ the matrix with entries

$$(2.6) \qquad\qquad (\theta^{cd} H)_{ij} = \begin{cases} \theta H_{ij} & \text{if } (i, j) = (c, d) \text{ or } (j, i) = (c, d), \\ H_{ij} & \text{otherwise.} \end{cases}$$

This object naturally appears as the error in Taylor expansions with respect to matrix entries.

LEMMA 2.2.   *Fix $\delta, \varepsilon > 0$. There are smooth functions $\tilde{\lambda}_i(X)$ (depending on ε, δ) on the space of $N \times N$ symmetric matrices with the following properties. Suppose that H is a real symmetric generalized Wigner matrix. With overwhelming probability, uniformly for all $i \in [\![\alpha N, (1-\alpha)N]\!]$, all integers $1 \le k \le 5$, and all choices of indices $1 \le a, b, c, d \le N$, we have*

$$(2.7) \qquad \left| \tilde{\lambda}_i(H) - \lambda_i(H) \right| \le C \frac{N^\varepsilon}{N^{1+\delta}}, \quad \sup_{0 \le \theta \le 1} \left| \partial_{ab}^k \tilde{\lambda}_i(\theta^{cd} H) \right| \le C \frac{N^{\varepsilon + (k-1)\delta}}{N},$$

where $\partial_{ab} = \partial_{X_{ab}}$ denotes the partial derivative with respect to the matrix entry with index (a, b) and $C = C(\varepsilon, \delta) > 0$ is a constant. The matrix $\theta^{cd} H$ is defined in (2.6).

Further, we have

$$(2.8) \qquad\qquad \sup_{0 \le \theta \le 1} \left| \partial_{ab}^k \tilde{\lambda}_i(\theta^{cd} H) \right| \le C N^{Ck}$$

almost surely for all integers $1 \le k \le 5$.

In the complex Hermitian case, the same estimates hold where instead ∂_{ab}^k is replaced by $\partial_{\text{Re}[X_{ab}]}^i \partial_{\text{Im}[X_{ab}]}^j$ for $i + j = k$, and H is a complex Hermitian generalized matrix.

The following is the main technical proposition of this work. It provides a regularization of the largest eigenvalue gaps and estimates on the derivatives of this regularization with respect to the matrix entries.

PROPOSITION 2.3.   *Let X be a real symmetric generalized Wigner matrix, and let $J \subseteq [\![\alpha N, (1-\alpha)N]\!]$. Let $\beta = N^\gamma$ and suppose that $\ell \le N^{\mathfrak{a}}$ for some $\gamma, \mathfrak{a} > 0$. Let $\tilde{\lambda}_i$ and $F_{\ell,\beta,J}$ be as above, and fix $\delta, \varepsilon > 0$. Then the following statements hold for all choices of indices $1 \le a, b \le N$. First, with overwhelming probability,*

$$(2.9) \qquad \left| v\mathcal{T}_{\ell,J}(\lambda) - F_{\ell,\beta,J}(\tilde{\lambda}) \right| \le Cv \frac{N^\varepsilon}{N^{1+\delta}} + 2 \frac{N^{\mathfrak{a}} \log(N)}{N^\gamma}.$$

Second, with overwhelming probability for all $1 \le k \le 5$,

$$(2.10) \qquad \sup_\theta \left| \partial_{ij}^k F_{\ell,\beta,J} [\tilde{\lambda}(\theta_{ab} H)] \right| \le \frac{C N^{k\delta}}{\beta} (1 + N^{k(\varepsilon + \gamma + \mathfrak{a} - \delta)}).$$

Third, we have almost surely

$$\sup_\theta \left| \partial_{ij}^k F_{\ell,\beta,J} \big[\tilde{\lambda}(\theta_{ab} H) \big] \right| \leq C N^{Ck}. \tag{2.11}$$

Here, $C = C(\delta, \varepsilon, \gamma, \mathfrak{a}, \alpha) > 0$ is a constant. The statements in the complex Hermitian case are adjusted as indicated in Lemma 2.2.

Before proving the above proposition, we derive the following elementary estimate.

LEMMA 2.4. *The partial derivatives of $F_{\ell,\beta}(v)$ with respect to the entries of the vector v satisfy*

$$\sum_{\underline{j}} \left| \frac{\partial^d F(v)}{\partial_{j_1} \cdots \partial_{j_d}} \right| \leq C_d \beta^{d-1} v^d \ell^d, \tag{2.12}$$

where the sum runs over all multiindices $\underline{j} = (j_1, \ldots, j_d)$ with values in J^d, and $\partial_j = \partial_{v_j}$.

PROOF. We have

$$\partial_j Z(v) = \beta v \sum_{\underline{i}} \varepsilon_{\underline{i}}(j) \exp\left(\beta v \sum_{k=1}^{\ell} v_{i_k+1} - v_{i_k} \right), \tag{2.13}$$

where $\varepsilon_{\underline{i}}(j)$ is defined as

$$\varepsilon_{\underline{i}}(j) = \begin{cases} 0 & \text{if } j = i_k \text{ and } i_k = i_{k-1} + 1 \text{ for some } k \in [\![1, \ell]\!], \\ -1 & \text{if } j = i_k \text{ for some } k \in [\![1, \ell]\!], \text{ but } i_k \neq i_{k-1} + 1, \\ 1 & \text{if } j = i_k + 1 \text{ for some } k \in [\![1, \ell]\!], \text{ but } i_{k+1} \neq i_k + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.14}$$

Higher derivatives are similar, yielding

$$\frac{\partial^d Z(v)}{\partial_{j_1} \cdots \partial_{j_d}} = (\beta v)^d \sum_{\underline{i}} \varepsilon_{\underline{i}}(\underline{j}) \exp\left(\beta v \sum_{k=1}^{\ell} v_{i_k+1} - v_{i_k} \right), \tag{2.15}$$

with $\varepsilon_{\underline{i}}(\underline{j}) \in \{-1, 0, 1\}$ satisfying $\varepsilon_{\underline{i}}(\underline{j}) = 0$ if $\underline{j} \not\subset \underline{i} \cup \underline{i} + 1$ (where we abuse notation and consider these tuples as sets). Hence, for all d,

$$\sum_{\underline{j}} \left| \frac{\partial^d Z(v)}{\partial_{j_1} \cdots \partial_{j_d}} \right| \leq (\beta v)^d \sum_{\underline{i}} \exp\left(\beta v \sum_{k=1}^{\ell} v_{i_k+1} - v_{i_k} \right) \sum_{\underline{j}} |\varepsilon_{\underline{i}}(\underline{j})| \tag{2.16}$$

$$\leq (\beta v)^d Z(v) (2\ell)^d.$$

By the chain rule and the definition of $F_{\ell,\beta,J}(v)$ in terms of $\log(Z)$, we then find

$$\sum_{\underline{j}} \left| \frac{\partial^d F(v)}{\partial_{j_1} \cdots \partial_{j_d}} \right| \leq \frac{1}{\beta} C_d \ell^d (\beta v)^d \tag{2.17}$$

for some combinatorial factor C_d. $\quad\square$

We are now prepared to prove Proposition 2.3, Theorem 1.3 and Theorem 1.4.

PROOF OF PROPOSITION 2.3. The estimate (2.9) follows from the first estimate of (2.7) and Lemma 2.1. For the second estimate, fix integers $1 \leq d \leq k \leq 5$ and let $\{s_i\}_{i=1}^d$ be positive

integers such that $s_1 + s_2 + \cdots s_d \leq k$. We estimate the following function evaluated at the matrix $\theta_{ab} H$:

$$
\left| \sum_{j_1, \ldots, j_d} \frac{\partial^d F}{\partial j_1 \cdots \partial j_d} (\tilde{\boldsymbol{\lambda}}) \frac{\partial^{s_1} \tilde{\lambda}_{j_1}}{\partial_{ij}^{s_1}} \cdots \frac{\partial^{s_d} \tilde{\lambda}_{j_d}}{\partial_{ij}^{s_d}} \right|
$$

$$
(2.18) \qquad \leq \sum_{j_1, \ldots, j_d} \left| \frac{\partial^d F}{\partial j_1 \cdots \partial j_d} (\tilde{\boldsymbol{\lambda}}) \right| N^{d\varepsilon + k\delta - d\delta} \frac{1}{N^d}
$$

$$
\leq \frac{C_k N^{k\delta}}{\beta} \left(\frac{N^{\varepsilon + \mathfrak{a} + \gamma} \nu}{N^{1+\delta}} \right)^d \leq \frac{C_k N^{k\delta}}{\beta} N^{d(\varepsilon + \gamma + \mathfrak{a} - \delta)}.
$$

In the final inequality, we used $\nu \leq N$. The first inequality is by Lemma 2.2 and the second inequality is Lemma 2.4. By the chain rule, the kth partial derivative of F with respect to the (i, j) matrix element is a linear combination of such terms, and the claim follows. The third estimate is similar, using instead the last estimate of Lemma 2.2. $\quad\square$

PROOF OF THEOREM 1.3. By Proposition 2.3, it is enough to prove Theorem 1.3 for the function $S(F(\tilde{\boldsymbol{\lambda}}))$ instead of $\nu \mathcal{T}(\boldsymbol{\lambda})$, assuming the parameters $\delta, \varepsilon, \gamma, \mathfrak{a}$ are chosen such that $\gamma > \mathfrak{a}$ and $\delta > \varepsilon$. The comparison for $S(F(\tilde{\boldsymbol{\lambda}}))$ is a standard application of the Lindeberg four-moment matching method, which we now briefly recall. In this approach, one replaces the upper triangular matrix entries of $H^{(v)}$ by those of $H^{(w)}$ one by one. The difference (1.7) is then a telescoping sum with $\mathcal{O}(N^2)$ terms, each being the difference of $S(F(\tilde{\boldsymbol{\lambda}}))$ evaluated at two generalized Wigner matrices differing in only in the (a, b) and (a, b) entries. One then Taylor expands each term around $X_{ab} = 0$ to 5th order. The first four orders cancel to $\mathcal{O}(N^{-2-c'})$ thanks to the hypothesis (1.6), and the 5th order remainder term is $\mathcal{O}(N^{-2-c'})$ due (2.10) and (2.11). As there are $\mathcal{O}(N^2)$ terms in the telescoping sum, the total error is $\mathcal{O}(N^{-c'})$, for some $c' > 0$, and we deduce the result. For a detailed exposition of this method, we refer to Chapter 16 of [17].

For clarity, and to explain the origin of the restrictions on the constants, we compute the 4th order error term. Specifically, we fix a matrix entry (a, b) and bound the error that arises when replacing v_{ab} in the (a, b) and (b, a) entry with w_{ab} (assuming the remaining entries are some collection of v_{ij} and w_{ij} entries). Write $F = F(\tilde{\boldsymbol{\lambda}})$ and $S = S(F(\tilde{\boldsymbol{\lambda}}))$, let S_x denote S evaluated with the (a, b) entry equal to x, and denote $\partial = \partial_{ab}$. Then Taylor expanding about 0 in the (a, b) entry gives that the 4th order error term for $S_{v_{ab}} - S_0$ is

$$
(2.19) \qquad \frac{v_{ab}^4}{4!} \left(S^{(4)} (\partial F)^4 + 6 S^{(3)} (\partial F)^2 \partial^2 F + S^{(2)} (3(\partial^2 F)^2 + 4(\partial^3 F) \partial F) + S^{(1)} \partial^{(4)} F \right),
$$

where this expression is evaluated with the (a, b) entry set to zero. Taking expectation in (2.19) and subtracting the analogous expression for $S_{w_{ab}} - S_0$, we see the w_{ab}^4 and v_{ab}^4 factors are independent of the terms in the parentheses, and we make take the difference of their expectations to obtain a factor that is $O(N^{-2-c})$ by (1.6), for some given c depending on the matrix ensembles $H^{(v)}$ and $H^{(w)}$. It then suffices to show that the sum of the terms in parentheses in (2.19) is $O(N^{\mathfrak{b}})$ for some $\mathfrak{b} < c$.

We consider just the $S^{(1)} \partial^{(4)} F$ term, since the others are similar.[2] Because $S^{(1)}$ is bounded by $\|S^{(1)}\|_\infty$, it suffices to bound $\partial^{(4)} F$ in expectation. If we first work on the set of the overwhelming probability where (2.10) holds, then it suffices to choose $\delta, \varepsilon, \gamma, \mathfrak{a}$ such that $\gamma = 2\mathfrak{a}, \delta = 2\varepsilon$, and $c/100 > \delta, \varepsilon, \gamma, \mathfrak{a}$, for example.[3] On the set where this estimate does not

---

[2] The essential point is that the orders of the derivatives on F in each term sum to 4.
[3] This ensures that $\gamma > \mathfrak{a}$ and $\delta > \varepsilon$, as was required by the application of Proposition 2.3 in the beginning of this proof.

hold, we use (2.11). We may then take $c_1 = c/100$ in the statement of Theorem 1.3, and this completes the proof. $\square$

PROOF OF THEOREM 1.4. Theorem 1.4 is a consequence of Lemma A.1 of [5] and Proposition 2.3. $\square$

3. Eigenvalue regularization. This section is devoted to the proof of Lemma 2.2. We only treat the real symmetric case in detail, as the complex Hermitian case is analogous. We remark that the choice of $\tilde{\lambda}_i$ is the same in both cases, as it is just a function of the empirical Stieltjes transform $\text{tr}(X - z)^{-1}$. The proof of the derivative estimates is similar in both cases, with real symmetric being notationally simpler.

In the first subsection, we construct the regularized eigenvalues $\tilde{\lambda}_i$ and at the same time prove the first estimate of (2.7). In the remaining subsections we estimate derivatives.

3.1. *Construction of regularized eigenvalues.* Let $\{\lambda_i\}_{i=1}^N$ be the eigenvalues of a generalized Wigner matrix H and let $\alpha \in (0, 1/2)$ be fixed. Fix $i \in [\![\alpha N, (1 - \alpha)N]\!]$ in the bulk. Let $\varepsilon_1 > 0$ be a parameter. Let j and k be indices such that

$$(3.1) \qquad i - 2N^{\varepsilon_1} \le j \le i - N^{\varepsilon_1}, \qquad i + N^{\varepsilon_1} \le k \le i + 2N^{\varepsilon_1},$$

and set

$$(3.2) \qquad I = [\gamma_j, \gamma_k].$$

Let $N(E) = |\{j : \lambda_j < E\}|$ be the eigenvalue counting function for X.

Using rigidity for generalized Wigner matrices, we have with overwhelming probability that

$$(3.3) \qquad \lambda_i - \gamma_j = \int_{\gamma_j}^{\lambda_i} dE = \int_I \mathbb{1}_{\{\lambda_i \le E\}} \, dE = \int_I \mathbb{1}_{\{N(E) \ge i\}} \, dE.$$

Let now $r : \mathbb{R} \to [0, 1]$ be a smooth function such that $r(x) = 1$ for $x \ge i$ and $r(x) = 0$ for $x \le i - 1/2$. We take $|r'| + |r''| + |r'''| \le C$. Then since $N(E)$ is integer valued,

$$(3.4) \qquad \int_I \mathbb{1}_{\{N(E) \ge i\}} \, dE = \int_I r(N(E)) \, dE.$$

Let $\delta_1 > 0$ be a parameter and set

$$(3.5) \qquad \eta_1 = N^{-1-\delta_1}.$$

For each $E \in I$, we define the smoothed out eigenvalue counting function $f_E(x)$ on the scale η_1 as follows. We define $f_E(x) = 1$ for $-10 \le x \le E$ and $f_E(x) = 0$ for $x \ge E + \eta_1$ and $x \le -11$. We take $|f_E^{(k)}| \le C_k \eta_1^{-k}$ for $k = 1, 2, 3$ and $E \le x \le E + \eta_1$, and $|f_E^{(k)}(x)| \le C$ for $k = 1, 2, 3$ and $-11 \le E \le -10$. Now, we have

$$(3.6) \qquad |N(E) - \text{tr}(f_E)| \le |\{a : \lambda_a \in [E, E + \eta_1]\}|,$$

and so by rigidity we have with overwhelming probability

$$
\begin{aligned}
\left| \int_I r(N(E)) \, dE - \int_I r(\text{tr}(f_E)) \, dE \right| &\le C \int_I |N(E) - \text{tr}(f_E)| \, dE \\
&\le C \int_I |\{a : \lambda_a \in [E, E + \eta_1]\}| \, dE \\
(3.7) \qquad &= C \sum_a \int_I \mathbb{1}_{\{\lambda_a \in [E, E + \eta_1]\}} \, dE \\
&\le C \eta_1 |\{a : \lambda_a \in I + [0, \eta_1]\}| \\
&\le C \eta_1 N^{\varepsilon_1} = C \frac{N^{\varepsilon_1 - \delta_1}}{N}.
\end{aligned}
$$

Let $\eta_2 = N^{-\delta_2}/N$ for a parameter $\delta_2 \in (0, 1)$, and let χ be a smooth, symmetric cut-off function with $\chi(x) = 1$ for $|x| \leq 1$ and $\chi(x) = 0$ for $|x| > 2$. By the Helffer–Sjöstrand formula [17], we have

$$
\begin{aligned}
(3.8) \quad \operatorname{tr}(f_E) &= \frac{N}{2\pi} \int_{\mathbb{R}^2} (if_E(e)\chi'(\sigma) - \sigma f_E'(e)\chi'(\sigma))m_N(e+i\sigma)\,de\,d\sigma \\
&\quad - \frac{N}{2\pi} \int_{|\sigma|>\eta_2} \int_{\mathbb{R}} f_E''(e)\sigma\chi(\sigma)\operatorname{Im}[m_N(e+i\sigma)]\,de\,d\sigma \\
&\quad - \frac{N}{\pi} \int_0^{\eta_2} \int_{\mathbb{R}} f_E''(e)\sigma\operatorname{Im}[m_N(e+i\sigma)]\,de\,d\sigma.
\end{aligned}
$$

Since $m(z)$ is holomorphic, we have the Cauchy–Riemann equation $\partial_x \operatorname{Im} m(x + iy) = -\partial_y \operatorname{Re} m(x + iy)$. Using this, we integrate the second term by parts twice (first in e and then in σ) to obtain

$$
\begin{aligned}
(3.9) \quad \operatorname{tr}(f_E) &= \frac{N}{2\pi} \int_{\mathbb{R}^2} (if_E(e)\chi'(\sigma) - \sigma f_E'(e)\chi'(\sigma))m_N(e+i\sigma)\,de\,d\sigma \\
&\quad + \frac{N}{2\pi} \int_{|\sigma|>\eta_2} \int_{\mathbb{R}} f_E'(e)\partial_\sigma(\sigma\chi(\sigma))\operatorname{Re}[m_N(e+i\sigma)]\,de\,d\sigma \\
&\quad - \frac{N}{\pi} \int_0^{\eta_2} \int_{\mathbb{R}} f_E''(e)\sigma\operatorname{Im}[m_N(e+i\sigma)]\,de\,d\sigma \\
&\quad - \frac{N}{\pi} \int_{\mathbb{R}} f_E'(e)\eta_2\operatorname{Re}[m_N(e+i\eta_2)]\,de.
\end{aligned}
$$

Define

$$
\begin{aligned}
(3.10) \quad F_E &= \frac{N}{2\pi} \int_{\mathbb{R}^2} (if_E(e)\chi'(\sigma) - \sigma f_E'(e)\chi'(\sigma))m_N(e+i\sigma)\,de\,d\sigma \\
&\quad + \frac{N}{2\pi} \int_{|\sigma|>\eta_2} \int_{\mathbb{R}} f_E'(e)\partial_\sigma(\sigma\chi(\sigma))\operatorname{Re}[m_N(e+i\sigma)]\,de\,d\sigma.
\end{aligned}
$$

We estimate, using the definition of f_E,

$$
\begin{aligned}
(3.11) \quad &\left| \int_I r(\operatorname{tr} f_E)\,dE - \int_I r(F_E)\,dE \right| \\
&\leq C \int_I \int_{\mathbb{R}} \int_0^{\eta_2} N\sigma|f_E''(e)|\operatorname{Im}[m_N(e+i\sigma)]\,d\sigma\,de\,dE \\
&\quad + C \int_I \int_{\mathbb{R}} N\eta_2|f_E'(e)||m_N(e+i\eta_2)|\,de\,dE \\
&\leq \frac{C}{\eta_1^2} \int_I \int_0^{\eta_1} \int_0^{\eta_2} N\sigma\operatorname{Im}[m_N(E+e+i\sigma)]\,d\sigma\,de\,dE \\
&\quad + C\frac{\eta_2}{\eta_1} \int_I \int_0^{\eta_1} N|m_N(E+e+i\eta_2)|\,de\,dE \\
&\quad + CN^{\varepsilon_1}\eta_2 \\
(3.12) \quad &= \frac{C}{\eta_1^2} \int_0^{\eta_2} \int_0^{\eta_1} \left[\int_I N\sigma\operatorname{Im}[m_N(E+e+i\sigma)]\,dE \right] de\,d\sigma \\
(3.13) \quad &\quad + C\frac{\eta_2}{\eta_1} \int_0^{\eta_1} \left[\int_I N|m_N(E+e+i\eta_2)|\,dE \right] de \\
&\quad + CN^{\varepsilon_1}\eta_2.
\end{aligned}
$$

In the preceding equations, we used rigidity to bound the contributions for $e \in [-11, -10]$. We now estimate the inner integral of (3.12). Define $E' = E + e$. We have

$$(3.14) \quad \int_I \sigma N \operatorname{Im}[m_N(E' + i\sigma)] \, dE \leq \sum_{a: |i-a| \leq 3N^{\varepsilon_1}} \eta_2 \int_I \frac{\sigma}{(E' - \lambda_a)^2 + \sigma^2} \, dE$$

$$+ \sum_{a: |i-a| > 3N^{\varepsilon_1}} \eta_2 \int_I \frac{\sigma}{(E' - \lambda_a)^2 + \sigma^2} \, dE.$$

We estimate the first term by

$$(3.15) \quad \sum_{a: |i-a| \leq 3N^{\varepsilon_1}} \eta_2 \int_I \frac{\sigma}{(E' - \lambda_a)^2 + \sigma^2} \, dE$$

$$\leq \sum_{a: |i-a| \leq 3N^{\varepsilon_1}} \eta_2 \int_{\mathbb{R}} \frac{\sigma}{(E' - \lambda_a)^2 + \sigma^2} \, dE \leq C\eta_2 N^{\varepsilon_1}.$$

For the second term, we have with overwhelming probability for a with $|i - a| > 3N^{\varepsilon_1}$ that $|E' - \lambda_a| \geq cN^{\varepsilon_1}/N$ and so

$$(3.16) \quad \sum_{a: |i-a| > 3N^{\varepsilon_1}} \eta_2 \int_I \frac{\sigma}{(E' - \lambda_a)^2 + \sigma^2} \, dE$$

$$\leq C\eta_2 \frac{N\sigma}{N^{\varepsilon_1}} |I| N \sup_{E'} [\operatorname{Im} m_N(E' + iN^{\varepsilon_1}/N)]$$

$$\leq C\sigma \eta_2 N \leq \eta_2 N^{-\delta_2} \leq \eta_2 N^{\varepsilon_1}.$$

In the last inequality, we used (1.19) to bound $\operatorname{Im} m_N$. The term (3.12) is therefore bounded by $CN^{-1+\varepsilon_1+\delta_1-2\delta_2}$.

For the term (3.13), we again write

$$(3.17) \quad \int_I N|m_N(E' + i\eta_2)| \, dE \leq \sum_{a: |i-a| \leq 3N^{\varepsilon_1}} \int_I \frac{1}{|(E' - \lambda_a) + i\eta_2|} \, dE$$

$$(3.18) \quad + \sum_{a: |i-a| > 3N^{\varepsilon_1}} \int_I \frac{1}{|(E' - \lambda_a) + i\eta_2|} \, dE.$$

The integral in the first term can be bounded directly and we obtain for that sum the bound

$$(3.19) \quad CN^{\varepsilon_1} |\log(\eta_2)| \leq CN^{\varepsilon_1} \log(N).$$

For the second term, we have the bound $CN^{\varepsilon_1} \log(N)$ by rigidity (1.22), with overwhelming probability. Using these bounds, we conclude that (3.13) satisfies the bound

$$(3.20) \quad C\frac{\eta_2}{\eta_1} \int_0^{\eta_1} \left[\int_I N|m_N(E + e + i\eta_2)| \, dE \right] de \leq CN^{-1+\varepsilon_1-\delta_2} \log(N).$$

Hence, recalling (3.11),

$$(3.21) \quad \left| \int_I r(\operatorname{tr} f_E) \, dE - \int_I r(F_E) \, dE \right| \leq C\frac{N^{\varepsilon_1+\delta_1-2\delta_2}}{N} + CN^{-1+\varepsilon_1-\delta_2} \log(N).$$

Using (3.7), we have therefore proven that, with overwhelming probability,

$$(3.22) \quad \left| (\lambda_i - \gamma_j) - \int_I r(F_E) \, dE \right| \leq CN^{\varepsilon_1} \frac{(N^{\delta_1-2\delta_2} + N^{-\delta_1} + N^{-\delta_2})}{N} \log(N).$$

We define $\tilde{\lambda}_i$ by

$$(3.23) \qquad \tilde{\lambda}_i = \int_I r(F_E) \, dE + \gamma_j.$$

We now fix parameters $\delta_1 = \delta_2 = \delta$ and $\varepsilon_1 = \varepsilon/2$ and observe the first estimate of (2.7) holds for the $\tilde{\lambda}_i$ we have constructed.

3.2. *Derivative bounds.* In this subsection, we obtain estimates on derivatives of the $\tilde{\lambda}_i$ defined in the previous subsection with respect to the matrix entries. We first need some estimates on the Green's function. Since in the proof of Theorem 1.3 we need to bound remainder terms in Taylor expansions involving the matrices $\theta_{ab}H$ as defined above, we make the following definition.

DEFINITION 3.1. We say that rigidity and delocalization hold for a matrix X with parameters α and ε if

$$(3.24) \qquad \sup_{i \in [\![\alpha N, (1-\alpha)N]\!]} |\lambda_i(X) - \gamma_i| \leq \frac{N^{\varepsilon/10}}{N},$$

and for its eigenvectors u_i we have

$$(3.25) \qquad \sup_i \|u_i\|_\infty \leq \frac{N^{\varepsilon/10}}{\sqrt{N}}.$$

3.2.1. *Green's function bounds.* In this section, we consider self-adjoint matrices X with a Green's function we denote by G_{ab}. We have the following two a priori estimates. The first is immediate from [2], Lemma 10.2, and the spectral decomposition [2], (2.1).

LEMMA 3.2 ([2], Lemma 10.2). *Fix $\alpha \in (0,1)$ and $\varepsilon > 0$. Let $E \in [\gamma_{\alpha N/2}, \gamma_{(1-\alpha/2)N}]$ and $\eta > 0$. If rigidity and delocalization hold for X, then*

$$(3.26) \qquad |G_{ab}(E + i\eta)| \leq N^\varepsilon \left(\frac{1}{N\eta} + 1 \right).$$

We also have a better bound if we are integrating in E.

LEMMA 3.3. *Fix $\alpha \in (0,1)$ and $\varepsilon, \varepsilon_1 \in (0, 1/2)$ with $\varepsilon < \varepsilon_1/2$. Let $I \subseteq [\gamma_{\alpha N/2}, \gamma_{(1-\alpha/2)N}]$ be an interval such that $|I| \leq N^{\varepsilon_1}/N$. There exists a constant $C = C(\varepsilon, \alpha) > 0$ such that if rigidity and delocalization hold for X with parameters ε and α, then for any $\eta > 0$,*

$$(3.27) \qquad \int_I |G_{ab}(E + i\eta)| \, dE \leq C \frac{N^{\varepsilon + \varepsilon_1}}{N} (1 + |\log(\eta)|).$$

PROOF. By the spectral theorem and eigenvector delocalization,

$$(3.28) \qquad |G_{ab}(z)| = \left| \sum_n \frac{u_n(a)u_n(b)}{\lambda_n - z} \right| \leq \frac{N^{\varepsilon/5}}{N} \sum_n \frac{1}{|\lambda_n - z|}.$$

We write

$$
\begin{aligned}
\int_I |G_{ab}(E + i\eta)| \, dE &\leq \frac{N^{\varepsilon/5}}{N} \sum_{n:|n-i| \leq \log(N)N^\varepsilon} \int_I \frac{1}{|\lambda_i - z|} \, dE \\
(3.29) \qquad &+ \frac{N^{\varepsilon/5}}{N} \sum_{n:|n-i| > \log(N)N^\varepsilon} \int_I \frac{1}{|\lambda_i - z|} \, dE,
\end{aligned}
$$

where i is an index such that $\gamma_i \in I$. The first sum is bounded by

$$(3.30) \qquad \frac{N^{\varepsilon/5}}{N} \sum_{n:|n-i|<\log(N)N^\varepsilon} \int_I \frac{1}{|\lambda_i - z|} \, dE \leq C \frac{N^{\varepsilon+\varepsilon_1}}{N} (|1 + \log(\eta)|).$$

The second sum can be estimated using rigidity, and we obtain

$$(3.31) \qquad \frac{N^{\varepsilon/5}}{N} \sum_{n:|n-i|>\log(N)N^{\varepsilon_1}} \int_I \frac{1}{|\lambda_i - z|} \, dE \leq C N^{\varepsilon/5} \log(N)|I| \leq C \frac{N^{\varepsilon+\varepsilon_1}}{N}.$$

Combining (3.30) and (3.31) completes the proof. $\quad\square$

From the previous two lemmas, we quickly deduce the following.

LEMMA 3.4. *Let I, α, ε, and C be as in the previous lemma. If rigidity and delocalization hold for X with parameters ε and α, then*

$$(3.32) \qquad \int_I |G_{a_1 b_1}(z)| \cdots |G_{a_k b_k}(z)| \, dE \leq C \frac{N^{k\varepsilon+\varepsilon_1}}{N} (1 + |\log(\eta)|) \left(\frac{1}{N\eta} + 1 \right)^{k-1}.$$

3.2.2. *Differentiation.* We have seen that the quantity

$$(3.33) \qquad \int_I r(F_E) \, dE$$

is a good approximation for the fluctuations of the eigenvalue λ_i. We now estimate derivatives of this quantity with respect to matrix elements. We use the shorthand $\partial_{bc} = \partial_{X_{bc}}$ for such derivatives. Let us rewrite

$$F_E = \frac{N}{\pi} \int_{\mathbb{R}^2} (i f_E(e) \chi'(\sigma) - \sigma f_E'(e) \chi'(\sigma)) m_N(e + i\sigma) \, de \, d\sigma$$
$$+ \frac{2iN}{\pi} \int_{|\sigma|>\eta_2} \int_{\mathbb{R}} f_E'(e) \partial_\sigma (\sigma \chi(\sigma)) \operatorname{Re}[m_N(e + i\sigma)] \, de \, d\sigma = A_E + B_E.$$

We have the following estimate for these two quantities.

LEMMA 3.5. *Fix $\alpha \in (0, 1)$ and $\varepsilon, \varepsilon_1 \in (0, 1/2)$ with $\varepsilon < \varepsilon_1/2$. Suppose that I is an interval such that $I \subseteq [\gamma_{\alpha N/2}, \gamma_{(1-\alpha/2)N}]$ and $|I| \leq N^{\varepsilon_1}/N$. Suppose rigidity and delocalization hold for the matrix X with parameters α and ε, and $1 \leq k \leq 5$. Then for $E \in I$ we have*

$$(3.34) \qquad |\partial_{bc}^k A_E| \leq C N^{k\varepsilon}, \qquad |\partial_{bc}^k B_E| \leq C N^{k\varepsilon+k\delta},$$

and for the integral over I we have

$$(3.35) \qquad \int_I |\partial_{bc}^k A_E| \leq C N^{k\varepsilon}|I|, \qquad \int_I |\partial_{bc}^k B_E| \leq C \frac{N^{k\varepsilon+\varepsilon_1+(k-1)\delta}}{N},$$

where $C = C(\alpha, \varepsilon, \varepsilon_1) > 0$ is a constant.

PROOF. We recall the Green's function differentiation formula[4]

$$(3.36) \qquad \frac{\partial G_{ij}(z)}{\partial H_{kl}} = -G_{ik}(z) G_{lj}(z) - G_{il}(z) G_{kj}(z).$$

---

[4]This is a straightforward consequence of the resolvent expansion to first order. See [2], (2.3).

Using it, we find

(3.37)
$$|\partial_{bc} Nm(e + i\sigma)| \leq C \sum_a |G_{ab} G_{ca}| \leq C \sum_a |G_{ab}|^2 + |G_{ca}|^2$$
$$\leq \frac{C}{\sigma} (|G_{bb}(z)| + |G_{cc}(z)|),$$

where in the last line we used the Ward identity [2], (3.16).

From this and (3.26), we see that

(3.38)
$$|\partial_{bc} A_E| \leq C N^{\varepsilon}.$$

Similarly, we obtain

(3.39)
$$|\partial_{bc}^k Nm(z)| \leq \frac{C}{\sigma} (|G_{bb}(z)| + |G_{cc}(z)| + |G_{bc}(z)|)^k,$$

and so

(3.40)
$$|\partial_{bc}^k A_E| \leq C N^{k\varepsilon}$$

for any $1 \leq k \leq 5$. This proves the first bound in (3.34), and the first bound in (3.35) follows immediately.

Now we consider B_E. We have the bound

(3.41)
$$|\partial_{bc}^k B_E| \leq C \int_E^{E+\eta_1} \int_{\sigma > \eta_2} \frac{1}{\eta_1} \frac{C}{\sigma} (|G_{bb}(z)| + |G_{cc}(z)| + |G_{bc}(z)|)^k \, de \, d\sigma$$
$$+ C \int_{-11}^{-10} \int_{\sigma > \eta_2} \frac{C}{\sigma} (|G_{bb}(z)| + G_{cc}(z)| + G_{bc}(z)|)^k \, de \, d\sigma.$$

From this and (3.26), we see that

(3.42)
$$|\partial_{bc}^k B_E| \leq N^{k\varepsilon + k\delta_2}.$$

From (3.41) and Lemma 3.4, we have

(3.43)
$$\int_I |\partial_{bc}^k B_E| \leq \frac{N^{\varepsilon + \varepsilon_1}}{N} \int_{\sigma > \eta_2} \chi(\sigma) \frac{1}{\sigma (N\sigma)^{k-1}} \, d\sigma \leq \frac{N^{2\varepsilon + \varepsilon_1 + (k-1)\delta}}{N}.$$

This completes the proof. $\square$

PROOF OF LEMMA 2.2. The smoothed eigenvalues $\tilde{\lambda}_i$ were constructed in Section 3.1, and the first estimate in (2.7) was already derived as (3.23).

To control the derivatives of $\tilde{\lambda}_i$, it suffices to control the derivatives of (3.33) using Lemma 3.5. We start with $k = 1$, and recall that by Theorem 1.9, delocalization and rigidity hold with overwhelming probability for the generalized Wigner matrix H. We have

(3.44)
$$\left| \partial_{bc} \int_I r(F_E) \, dE \right| \leq C \int_I |\partial_{bc} A_E| + |\partial_{bc} B_E| \, dE \leq \frac{N^{\varepsilon_1 + \varepsilon}}{N}.$$

Now for $k = 2$ we have

(3.45)
$$\left| \partial_{bc}^2 \int_I r(F_E) \, dE \right| \leq C \int_I |\partial_{bc}^2 A_E| + |\partial_{bc}^2 B_E| + (|\partial_{bc} B_E| + |\partial_{bc} A_E|)^2 \, dE.$$

The first term is bounded by

(3.46)
$$\int_I |\partial_{bc}^2 A_E| \, dE \leq C \frac{N^{\varepsilon_1 + 2\varepsilon}}{N}.$$

For the second term, we use (3.35) to bound

$$(3.47) \qquad \int_E |\partial_{bc}^2 B_E| \, dE \leq C \frac{N^{\varepsilon_1 + 2\varepsilon + \delta_2}}{N}.$$

For the last term, we estimate

$$(3.48) \qquad \int_I (|\partial_{bc} B_E| + |\partial_{bc} A_E|)^2 \, dE \leq N^{\varepsilon + \delta_2} \int_I |\partial_{bc} B_E| + |\partial_{bc} A_E| \, dE \leq C \frac{N^{\varepsilon_1 + 2\varepsilon + \delta_2}}{N}.$$

Hence we have with overwhelming probability,

$$(3.49) \qquad \left| \partial_{bc} \int_I r(F_E) \, dE \right| \leq \frac{N^{\varepsilon + 2\varepsilon_1 + \delta}}{N}$$

for any $\varepsilon > 0$. In general, the kth derivative of the quantity (3.33) is bounded by

$$(3.50) \qquad \left| \partial_{bc}^k \int_I r(F_E) \, dE \right| \leq C \sum_{M_E} \int_I M_E \, dE,$$

where each M_E is a monomial in the terms $|\partial_{bc}^n B_E|$ and $|\partial_{bc}^m A_E|$. There are finitely many monomials in the sum. If a monomial has $l \leq k$ terms and if we denote the order of the derivative in the ith term in the monomial by n_i then $n_1 + n_2 + \cdots + n_l = k$. We have

$$(3.51) \qquad \int_I M_E \, dE \leq N^{\varepsilon + (k - n_1)\delta_2} \int_I |\partial^{n_1} X_E| \, dE,$$

where X_E is either A_E or B_E. We see that

$$(3.52) \qquad \int_I |\partial^{n_1} X_E| \, dE \leq \frac{N^{\varepsilon + \varepsilon_1 + (n_1 - 1)\delta_2}}{N}.$$

Hence,

$$(3.53) \qquad \int_I M_E \, dE \leq C N^{2\varepsilon + \varepsilon_1 + (k-1)\delta_2} N.$$

This proves eigenvalue derivative bounds for H. We now extend them to the perturbations $\theta_{ab} H$.

Fix $\kappa > 0$ and consider $\eta \geq N^{-1+\kappa}$. From Theorem 1.8 and a resolvent expansion to high order,[5] it is straightforward to prove that

$$(3.54) \qquad \sup_{0 \leq \theta \leq 1} \left| \frac{1}{N} \operatorname{tr} \frac{1}{H - z} - \frac{1}{N} \operatorname{tr} \frac{1}{\theta_{ab} H - z} \right| \leq \frac{N^{\varepsilon}}{N\eta}$$

and

$$(3.55) \qquad \sup_{0 \leq \theta \leq 1} \left| \left(\frac{1}{H - z} \right)_{ij} - \left(\frac{1}{\theta_{ab} H - z} \right)_{ij} \right| \leq \frac{N^{\varepsilon}}{\sqrt{N}}$$

for any $\varepsilon > 0$, with overwhelming probability. Hence, by standard arguments,[6] rigidity and delocalization hold simultaneously for $\theta_{ab} H$ for all choices of θ, and we can apply the above calculations, which give the second estimate of (2.10).

Finally, the almost sure bounds in (2.8) follow from using the trivial bound $|G(z)|_{ij} \leq \eta^{-1}$ in the above calculations. $\quad \square$

---

[5]See [2], (2.3).

[6]See [2], Theorem 2.10 and [2], Section 9.

4. Extension to maximal gap over an interval. In this section, we extend the method of the previous subsection to cover statistics of the form

$$(4.1) \qquad \max_{\lambda_i \in I} \lambda_{i+1} - \lambda_i,$$

where $I = [a, b]$ is compact subinterval of $(-2, 2)$, which proves Theorem 1.5. The main issue is that the indicator function $\mathbb{1}_{\{\lambda_i \in I\}}$ is not differentiable. We pass to a smoothed object with the help of the following Wegner-type estimate.

LEMMA 4.1. *Let H by a generalized Wigner matrix and $E \in (-2 + \kappa, 2 - \kappa)$ for fixed $\kappa > 0$. For every $\varepsilon > 0$, there is $\delta = \delta(\varepsilon) > 0$ such that*

$$(4.2) \qquad \mathbb{P}\big[\exists i : |\lambda_i - E| \le 2N^{-1-\varepsilon}\big] \le N^{-\delta}.$$

PROOF. This is a corollary of fixed energy universality for generalized Wigner ensembles with an effective polynomial rate of convergence. This was proved in [27]. Alternatively, this estimate was proved for Gaussian divisible ensembles in Section 7 of [28]. The estimate can then be transferred to all ensembles using the four moment approach of [33, 34] or the matrix continuity estimate of [5]. □

Fix a small $\varepsilon_w > 0$ with corresponding δ_w as in the statement of the lemma. Let ρ be a smooth test function such that $\rho(x) = 1$ for $|x| \le \frac{1}{2}$ and $\rho(x) = 0$ for $|x| \ge 1$. Let $r(x) :$ $\mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ be a smooth function that is 1 for $x \le \frac{1}{2}$ and 0 for $x \ge 1$. Let i_0 and j_0 be the indices of the closest classical eigenvalues to a and b. Fix a small $\varepsilon_r > 0$ and consider the index set $J_r = [\![i_0 - N^{\varepsilon_r}, j_0 + N^{\varepsilon_r}]\!]$. Consider the function

$$(4.3) \qquad f_1(\boldsymbol{\lambda}) = r\bigg(\sum_{i \in J_r} \rho(N^{1+\varepsilon_w}(\lambda_i - a))\bigg),$$

define $f_2(\boldsymbol{\lambda})$ similarly but with b instead of a. Note that on the event that no eigenvalue is within distance $N^{-1-\varepsilon_w}$ of a and b we have that $f_1(x) = f_2(x) = 1$. Moreover, if there is an eigenvalue with distance $N^{-1-\varepsilon_w}/2$ of a, then f_1 is zero, and similarly for b and f_2.

Let $\chi(x)$ be a smoothed out step function with $\chi(x) = 1$ for $x \le -\frac{1}{2}$ and $\chi(x) = 0$ for $x \ge \frac{1}{2}$. Consider the two functions

$$(4.4) \qquad g_1(x) = \chi\big(10N^{1+\varepsilon_w}(a - x)\big), \qquad g_2(x) = \chi\big(10N^{1+\varepsilon_w}(x - b)\big).$$

The important observation is that if $|\lambda_i - b| > N^{-1-\varepsilon_w}$, then $g_2(\lambda_i) = \mathbb{1}_{\{x < b\}}$, and moreover, that if $g_2^{(k)}(\lambda_i) \ne 0$ for any $k \ge 1$, then f_2 and all of its derivatives are 0. There is a similar consideration with g_1 and f_1.

We now fix $\beta = N^\gamma$ and consider

$$(4.5) \qquad \widehat{F}(\boldsymbol{\lambda}) = f_1(\boldsymbol{\lambda}) f_2(\boldsymbol{\lambda}) \frac{1}{\beta} \log \widehat{Z},$$

where

$$(4.6) \qquad \widehat{Z}(\boldsymbol{\lambda}) = \sum_{i \in J_r} g_1(\lambda_i) g_2(\lambda_i) \exp\big[\nu\beta(\lambda_{i+1} - \lambda_i)\big].$$

LEMMA 4.2. *Let $I = [a, b] \subset (-2, 2)$. Let ε_w, ε_r, and δ_w be as above. Let f_1, f_2 and g_2 and g_1 be as above. Let H be a generalized Wigner matrix with eigenvalues λ_i. Let $\tilde{\lambda}_i$ be the regularized eigenvalues from Lemma 2.2, with ε and δ given. Choose these parameters to satisfy*

$$(4.7) \qquad \varepsilon_w < \delta - \varepsilon, \qquad \varepsilon < \varepsilon_r/2.$$

Then with probability at least $1 - CN^{-\delta_w}$, we have the estimate

$$(4.8) \qquad \left| \max_{i:\lambda_i \in I} \nu(\lambda_{i+1} - \lambda_i) - \widehat{F}(\tilde{\boldsymbol{\lambda}}) \right| \leq C\nu \frac{N^\varepsilon}{N^{1+\delta}} + \frac{\log(N)}{N^\gamma}.$$

PROOF. By Lemma 4.1, there is an event with probability at least $1 - 2N^{-\delta_w}$ on which there is no eigenvalue λ_i within distance $2N^{-1-\varepsilon_w}$ of the interval endpoints a or b. Further, rigidity (1.22) holds for a sufficiently small $\varepsilon < \varepsilon_r/2$ with overwhelming probability, as does the first estimate of (2.7). In the latter, we choose α sufficiently small so that the classical eigenvalues corresponding to indices in $[\![\alpha N, (1-\alpha)N]\!]$ contain an interval which contains I.

When these three events hold, $\lambda_i \in I$ if and only if $\tilde{\lambda}_i \in I$, $i \in J_r = [\![i_0 - N^{\varepsilon_r}, j_0 + N^{\varepsilon_r}]\!]$, and no $\tilde{\lambda}_i$ is within distance $N^{-1-\varepsilon_w}$ of the interval endpoints a or b (recall (4.7)). From this discussion, we see first that

$$(4.9) \qquad \left| \max_{i:\lambda_i \in I} \nu(\lambda_{i+1} - \lambda_i) - \max_{i \in J_r: \tilde{\lambda}_i \in I,} \nu(\tilde{\lambda}_{i+1} - \tilde{\lambda}_i) \right| \leq C\nu \frac{N^\varepsilon}{N^{1+\delta}}$$

when these events hold.

Second, we see that on the event in question that $f_1(\tilde{\lambda}_i) = f_2(\tilde{\lambda}_i) = 1$, and for $i \in J_r$ that $g_1(\tilde{\lambda}_i)g_2(\tilde{\lambda}_i) = \mathbb{1}_{\{\tilde{\lambda}_i \in I\}}$. Hence, similar to Lemma 2.1, we see that

$$(4.10) \qquad \left| \max_{i:\tilde{\lambda}_i \in I} \nu(\tilde{\lambda}_{i+1} - \tilde{\lambda}_i) - \widehat{F}(\tilde{\boldsymbol{\lambda}}) \right| \leq \frac{\log(N)}{\beta}.$$

This yields the claim. □

We now extend Lemma 2.4 to $\widehat{F}$. Consider $\widehat{F}$ as a function on $\mathbb{R}^N$, $\widehat{F}(\boldsymbol{v})$, for $\boldsymbol{v} \in \mathbb{R}^N$. The key observation is that if $\partial_{v_j}^{(k)}(f_1 f_2) \neq 0$ for some $k \geq 1$, then there is no v_i within distance $N^{-1-\varepsilon_w}/2$ of a or b. If the latter holds, then $\partial_{v_i}(g_1 g_2) = 0$ by construction. This allows the estimation of derivatives of $\log(\widehat{Z})$ without worrying about quantities like g_1'/g_1 growing large.

LEMMA 4.3. *Let $\varepsilon_w, \varepsilon_r, \widehat{F}, \widehat{Z}$ be as above. Let $\mathcal{A} \subseteq \mathbb{R}^N$ be defined as*

$$(4.11) \qquad \mathcal{A} = \{\boldsymbol{v} \in \mathbb{R}^N : 2|v_i - a| \geq N^{-1-\varepsilon_w} \text{ and } 2|v_i - b| \geq N^{-1-\varepsilon_w}, \forall i \in J_r\}.$$

For every $d \geq 1$, we have

$$(4.12) \quad \sum_{j_1,\dots,j_d} |\partial_{j_1} \cdots \partial_{j_d}(f_1(\boldsymbol{v})f_2(\boldsymbol{v}))|$$
$$\leq C_d N^{d\varepsilon_w + d}(|\{i \in J_r : |v_i - a| \leq N^{-1-\varepsilon_r}\}| + |\{i \in J_r : |v_i - b| \leq N^{-1-\varepsilon_r}\}|)^d \mathbb{1}_{\mathcal{A}}$$

and

$$(4.13) \qquad |f_1 f_2| \leq \mathbb{1}_{\mathcal{A}}.$$

We have also

$$(4.14) \qquad \mathbb{1}_{\mathcal{A}} \sum_{j_1,\dots,j_d} |\partial_{j_1} \cdots \partial_{j_d} \log \widehat{Z}| \leq C_d(\beta\nu)^d.$$

PROOF. The first two estimates follow from the fact that $f_1 f_2$ is the constant function 0 on the set $\mathcal{A}^c$, and that $\rho^{(k)}(x)$ is supported on the set $|x| \leq 1$. For the proof of the final estimate, we observe that by the above discussion that if $\boldsymbol{v} \in \mathcal{A}$ we have $\partial_{v_j}^{(k)} g_1(v_j)g_2(v_j) = 0$

and $g_1(v_j)g_2(v_j) = \mathbb{1}_{\{v_j \in I\}}$. With this in mind, the calculations in Lemma 2.4 go through without change, and we find the final estimate. $\quad \square$

With these preparations, we are ready to prove Theorem 1.5.

PROOF OF THEOREM 1.5. Rigidity implies that with overwhelming probability,

$$(4.15) \qquad \left|\{i \in J_r : |v_i - a| \leq N^{-1-\varepsilon_r}\}\right| + \left|\{i \in J_r : |v_i - b| \leq N^{-1-\varepsilon_r}\}\right| \leq N^{\varepsilon}$$

for any $\varepsilon > 0$. From Lemma 4.3 and Lemma 2.2, we find the analog of the estimate (2.10),

$$(4.16) \qquad \sup_{\theta}\left|\partial_{ij}^d \widehat{F}(\theta_{ab}H)\right| \leq \frac{C_d N^{d\delta}}{\beta}(1 + N^{d(\gamma + \varepsilon - \delta)} + N^{2\varepsilon + d\varepsilon_w}),$$

which holds with overwhelming probability. We also find the same almost sure estimate (2.11). In the case $\ell = 1$, the theorem is proven using these estimates, similar to how Theorems 1.3 and 1.4 are deduced from Proposition 2.3. The extension to $\ell > 1$ requires only replacing $\widehat{Z}$ by quantities like

$$(4.17) \qquad \sum_{i_1,\dots i_\ell}\left(\prod_j g_1(\tilde{\lambda}_{i_j})g_2(\tilde{\lambda}_{i_j})\right)\exp\left[\beta v \sum_{k=1}^{\ell}\tilde{\lambda}_{i_k+1} - \tilde{\lambda}_{i_k}\right]$$

and proceeding as before. $\quad \square$

5. Universality corollaries.

We consider the two processes

$$(5.1) \qquad dx_i = \sqrt{\frac{2}{N\beta}}\,dB_i + \frac{1}{N}\sum_{j \neq i}\frac{1}{x_i - x_j}\,dt$$

and

$$(5.2) \qquad dy_i = \sqrt{\frac{2}{N\beta}}\,dB_i + \frac{1}{N}\sum_{j \neq i}\frac{1}{y_i - y_j}\,dt.$$

For the initial data, we take $x_i(0) = \lambda_i(H)$ for all $i \in [\![1, N]\!]$, where H is a generalized Wigner matrix, and $y_i(0) = \lambda_i(G)$, where G is the Gaussian ensemble of the same symmetry class. The parameter is $\beta = 1$ in the real case and $\beta = 2$ in the complex case. Then the proof of [27], Theorem 3.1, implies the following result. We comment on adapting the argument there to the simpler setting here in the Appendix. The lemma may also be directly cited from [3], Corollary 3.2, which appeared after this paper was written.

PROPOSITION 5.1. *Let $t = N^{\omega}/N$, and let $\alpha \in (0, 1/2)$. Assume that $0 < \omega < 1/2$. Then for all $i \in [\![\alpha N, (1 - \alpha)N]\!]$ and all $\varepsilon > 0$, we have with overwhelming probability*

$$(5.3) \qquad \left|x_{i+1}(t) - x_i(t) - (y_{i+1}(t) - y_i(t))\right| \leq \frac{N^{\varepsilon}}{N^{1+c}}$$

for some constant $c = c(\alpha, \omega) > 0$ (independent of ε).

REMARK. The above flow is not quite the same as (1.8). However, since $t \ll 1$, Proposition 5.1 holds also for the flow (1.8), as explained in [24], Section 2.3. We use this fact without comment in what follows.

PROOF OF COROLLARY 1.6. We first consider the statement for $\mathcal{T}_{\ell,J}$. By the results of [17], Section 16.2, there exists a Gaussian divisible ensemble X_t matching H to four moments, in the sense of (1.6), with t satisfying the hypotheses of Proposition 5.1. By this proposition, the largest gap among $\{\lambda_{i+1} - \lambda_i : i \in J\}$ for X_t is the same as the corresponding quantity for the GOE or GUE up to an error of size $o(N^{-1})$. We deduce that the conclusion of Corollary 1.6 holds for X_t. Then by Theorem 1.3, it holds for H, too.

For the quantity $\widehat{\mathcal{T}}_{\ell,I}$, there is a small issue, which is that even if (5.3) holds, it may not be true that $x_i(t) \in I$ if and only if $y_i(t) \in I$. However, by Lemma 4.1 we know that for every $\varepsilon_w > 0$ there is a $\delta_w > 0$ such that the event that there is no $y_i(t)$ within distance $N^{-1-\varepsilon_w}$ of the endpoints of I holds with probability at least $1 - N^{-\delta_w}$. If we take ε_w small enough so that $N^{-1-\varepsilon_w}$ is larger than the error on the right-hand side of (5.3), then we see that on this event, $x_i(t) \in I$ if and only if $y_i(t) \in I$. Therefore, with probability at least $1 - CN^{-\delta_w}$, for some $\delta_w > 0$, the statistics $\widehat{\mathcal{T}}_{\ell,I}$ of H_t match those of the corresponding Gaussian ensemble. The remainder of the argument is similar to $\mathcal{T}_{\ell,J}$ and we conclude using Theorem 1.5. □

REMARK. We are thankful to the referee for pointing out the following simpler proof of Corollary 1.6 for $\widehat{\mathcal{T}}_{\ell,I}$, which has the advantage of avoiding the need to first prove Theorem 1.5. If $I = [a, b] \subset (-2, 2)$, then by the rigidity estimate (1.22) the random set of indices $J(I) = \{i : \lambda_i \in I\}$ satisfies

$$\tag{5.4} [\![\gamma_a + N^\varepsilon, \gamma_b - N^\varepsilon]\!] \subset J \subset [\![\gamma_a - N^\varepsilon, \gamma_b + N^\varepsilon]\!]$$

with overwhelming probability, where γ_a and γ_b are the classical eigenvalue locations closest to a and b, respectively, and $\varepsilon > 0$ is arbitrary. We may clearly neglect the exceptional set where this does not occur, using the trivial bound $|S| \leq \|S\|_\infty$.

Using Corollary 1.6 for $\mathcal{T}_J$, and observing that this estimate is uniform in J as in the statement of Theorem 1.4, we obtain

$$\tag{5.5} \sup_J \left| \mathbb{E}_H S(\nu \mathcal{T}_{\ell,J}(\lambda)) - \mathbb{E}_G S(\nu \mathcal{T}_{\ell,J}(\lambda)) \right| \leq C \Big(\sup_{0 \leq d \leq 5} \|S^{(d)}\|_\infty \Big) N^{-c},$$

where the supremum is taken over all J satisfying (5.4). It then suffices to show

$$\tag{5.6} \sup_J \left| \mathbb{E}_G S(\nu \mathcal{T}_{\ell,J}(\lambda)) - \mathbb{E}_G S(\nu \mathcal{T}_{\ell,I}(\lambda)) \right| \leq C \Big(\sup_{0 \leq d \leq 5} \|S^{(d)}\|_\infty \Big) N^{-c},$$

for some $c > 0$ (possibly smaller than the constant in (5.5)). In turn it suffices to show that there exists $c > 0$ such that, with probability at least $1 - N^{-c}$, the maximal gap over such J does not correspond to eigenvalues with indices in

$$\tag{5.7} [\![\gamma_a - N^\varepsilon, \gamma_a + N^\varepsilon]\!] \cup [\![\gamma_b - N^\varepsilon, \gamma_b + N^\varepsilon]\!].$$

By the treatment of the display before (3.5) in [1], there exists a constant $c' > 0$ such that the maximum gap over I, and hence over the indices $[\![\gamma_a + N^\varepsilon, \gamma_b - N^\varepsilon]\!]$, is at least $c'\sqrt{\log N}/N$ with probability at least $1 - N^{-c'}$. On the other hand, we see using [1], Lemma 3.2, [1], Lemma 3.5, and a union bound that the maximal gap among the indices in (5.7) is greater than $c'\sqrt{\log N}/N$ with probability at most $N^{-c''}$, for some constant $c''(c', \varepsilon) > 0$, where we have chosen ε small enough. This proves (5.6), and hence the corollary, after bounding S using the trivial bound $|S| \leq \|S\|_\infty$ on the exceptional set where the maximal gap corresponds to indices in (5.7).

PROOF OF COROLLARY 1.7. The first result of Corollary 1.7 will follow from applying Corollary 1.6 with S a polynomial, and the corresponding convergence of the largest gaps for the GUE given by [1], Theorem 1.7. For notational convenience, set $T = \frac{MN}{\sqrt{32 \log N}} \widehat{\mathcal{T}}_{\ell_N, I}$.

We recall H is the generalized Wigner matrix under consideration, and let G be the Gaussian ensemble corresponding to the symmetry class of X.

Note that Corollary 1.6 may be applied nontrivially for any test function S satisfying $\sup_{0 \le d \le 3} \|S^{(d)}(x)\|_\infty \ll N^c$, where c is as in Corollary 1.6. We let S be an N^ε-smoothed regularization of the L^p norm, with $S(x) = |x - 1|^p$ when $|x| < 4N^\varepsilon$ and $\sup_{0 \le d \le 3} \|S^{(d)}(x)\|_\infty \le C N^{3p\varepsilon}$.

By Theorem 1.9, the event $\mathcal{A}$ on which $|\lambda_i - \gamma_i| < N^{-1+\varepsilon/2}$ for all i such that $\gamma_i \in (a - \varepsilon, b + \varepsilon)$ holds with overwhelming probability. We now claim that, for any $D > 0$, there exists a $C(D, p, \varepsilon) > 0$ such that

$$(5.8) \qquad \mathbb{E}_H\big[|T - 1|^p\big] \le \mathbb{E}_H\big[|T - 1|^p \mathbb{1}_{\mathcal{A}}\big] + \mathbb{E}_H\big[|T - 1|^p \mathbb{1}_{\mathcal{A}^c}\big]$$

$$(5.9) \qquad \le \mathbb{E}_G\big[|T - 1|^p\big] + \big(\mathbb{E}_H[S(T)] - \mathbb{E}_G[S(T)]\big) + C N^{2p - D}.$$

In the last inequality, we used the fact that $S(T) = |T - 1|^p$ on the event $\mathcal{A}$ to split the first term in (5.8). For the second term in (5.8), we used Cauchy–Schwarz and the fact that $\mathcal{A}$ holds with overwhelming probability to write

$$(5.10) \qquad \mathbb{E}_H\big[|T - 1|^p \mathbb{1}_{\mathcal{A}^c}\big] \le \sqrt{\mathbb{E}_H\big[|T - 1|^{2p}\big]\mathbb{E}_H[\mathbb{1}_{\mathcal{A}^c}]} = N^{-D}\sqrt{\mathbb{E}_H\big[|T - 1|^{2p}\big]},$$

for N large enough, and bounded $\mathbb{E}_H[|T - 1|^{2p}]$ as follows. Observe that T is is bounded by N times the sum of the absolute values of the largest and smallest eigenvalues. The latter quantity can be bounded by (a constant times) the Frobenius norm of H, $\|H\|_F$, and we quickly see that

$$(5.11) \qquad \mathbb{E}\|H\|_F \le \sqrt{\mathbb{E}\|H\|_F^2} \le C\sqrt{N}$$

from the moment growth and independence assumptions on the matrix entries made in Definition 1.1.

Then in (5.9), the first term is $o(1)$ by [1], Theorem 1.7, the second term is $C N^{3p\varepsilon - c}$ by Corollary 1.6, and the third term is $O(N^{-1})$ when D is chosen large enough. Therefore, it converges to zero.

The proof of the second result is similar. It follows from choosing S to be smooth functions bracketing indicator functions, after a shift and rescaling by some powers of $\log(N)$, which do not affect the polynomial error rate. For instance, take $S(x)$ such that $S(x) = 1$ for $x \in I_1 = [a_1, b_1]$, $S(x) = 0$ for $x \notin (a_1 - N^{-\varepsilon}, b_1 + N^{-\varepsilon})$, and

$$(5.12) \qquad \sup_{x \in (a_1 - N^{-\varepsilon}, a_1) \cup (b_1, b_1 + N^{-\varepsilon})} \big|S^{(d)}(x)\big| \le C N^{d\varepsilon}$$

for $0 \le d \le 3$. $\quad \square$

APPENDIX: EIGENVALUE COUPLING

The purpose of this Appendix is to indicate how Proposition 5.1 may be deduced from the existing literature. It is essentially a consequence of [27], Theorem 3.1. However, this theorem was proved in a more general context and so there are some assumptions that may seem out of place. We fix the parameter $\alpha > 0$ and restrict our considerations to indices $i \in [\![\alpha N, (1 - \alpha)N]\!]$. We also fix $\omega \in (0, 1/2)$ and set $t = N^{-1+\omega}$. The constants below depend on these choices, but we omit this in the notation for this section. (The dependencies are written explicitly in Proposition 5.1.)

In [27], the behavior of $x_i(t)$ was considered for more general initial data $x_i(0)$ than just the generalized Wigner setting. One of the key differences is that the empirical eigenvalue density was no longer a semicircle, but given by a different density coming from free probability,

called the *free convolution*. In that work, it is denoted by $\rho_{\mathrm{fc},t}$. The regularity of this density is controlled by the parameter t, and so in [27], the particles were allowed to evolve until a fixed time t_0 before the coupling between x_i and y_i was introduced. This allowed the density $\rho_{\mathrm{fc},t}$ to have some regularity before being manipulated.

In the present setting, both $x_i(t)$ and $y_i(t)$ are described by the semicircle density, but supported on the interval $\sqrt{1+t}[-2, 2]$. The semicircle density is smooth in the bulk, so allowing the particles to evolve until t_0 is unnecessary in the present setting. Errors on the right-hand side of the main estimate of Theorem 3.1 of [27] can then be replaced by setting $\omega_0 = 1$.[7]

The result of [27] studies the behavior of $x_i(t)$ near a particle index i_0. In the set-up of [27], it is possible that i_0 was very close to 1 or N, for instance $i_0 = o(N)$. To account for this, the coupling between x_i and y_i was such that x_{i_0} and $y_{N/2}$ shared the same Brownian motion terms; otherwise, it could be possible that y_{i_0} did not share the bulk GOE/GUE statistics, as it could have been close to the edge, and the argument of [27] would have failed. Aside from being macroscopically separated from the edge, the index $N/2$ was chosen for notational convenience. Moreover, the process x_i was scaled and shifted so that the particle x_{i_0} was close to $y_{N/2}$ (which is close to 0), and so that the local particle density near this point was the same as the coupled Gaussian ensemble, this being given by $\rho_{\mathrm{sc}}(0)$. This is reflected in the assumption about $\rho_{\mathrm{fc},t}$ and the location of the classical eigenvalue of index i_0 stated in Theorem 3.1 of [27]. In the present setting, we are only interested in indices separated from the edge by a macroscopic distance, and the densities of $x_i(t)$ and $y_i(t)$ match, as they are both given by the semicircle. We can therefore ignore this assumption in our application of Theorem 3.1, as the proof would go through in the present setting.[8]

Finally, [27], Theorem 3.1, implies that

$$(A.1) \qquad x_i(t) - y_i(t) = \sum_{|j-i| \leq t N^{1+\delta}} \xi_t\left(\frac{i-j}{N}\right)(x_j(0) - y_j(0)) + O(N^{-1-c_1})$$

for any $\delta > 0$ and an appropriate constant $c_1 = c_1(\delta) > 0$, with overwhelming probability. Here, $\xi_t(x)$ is a smooth function obeying estimates as given in Proposition 3.2 of [27]. We claim that applying (A.1) to i and $i + 1$ yields

$$(A.2) \qquad \left| x_i(t) - y_i(t) - (x_{i+1}(t) - y_{i+1}(t)) \right|$$

$$\leq \left| \sum_{|j-i| \leq t N^{1+\delta}} \xi_t\left(\frac{i-j}{N}\right)(x_j(0) - y_j(0)) \right.$$

$$(A.3)$$

$$\left. - \sum_{|j-(i+1)| \leq t N^{1+\delta}} \xi_t\left(\frac{i+1-j}{N}\right)(x_j(0) - y_j(0)) \right|$$

$$(A.4) \qquad \leq \frac{N^\varepsilon}{N^{1+c_2}}$$

---

[7]Alternatively, we could have allowed the particles to evolve up a fixed time t_0 in the present setting, before allowing the coupling between x_i and y_i, but we would not gain anything by doing so.

[8]An alternative route to the application of Theorem 3.1 is the following. Instead of applying it directly to x_i and y_i, one can construct a third process $z_i(t)$, also coming from a Gaussian ensemble, so that $z_{N/2}(t)$ is coupled to x_{i_0} and y_{i_0} of the original processes. The processes x_{i_0} and y_{i_0} can then be shifted and rescaled *by the same constants* so that x_{i_0} and y_{i_0} are close to $z_{N/2}$ and the densities match the semicircle there. Then Theorem 3.1 can be applied twice, once to x_i and z_i, and the separately to y_i and z_i. Taking the differences between these differences, the z_i dependence drops out and after undoing the scalings, one arrives at an appropriate estimate for x_i and y_i.

for some constant $c_2 > 0$ with overwhelming probability. This implies the estimate we need for (5.3).

The essential point is that taking differences of gaps weighted by the ξ_t kernel produces a cancellation that improves upon the naive bound of $N^{\varepsilon-1}$ obtained by applying the rigidity estimate (1.22) to each gap separately. To show this precisely, we write the differences of the ξ_t terms as an integral of $\partial_x \xi_t$ and use the derivative estimate [27], (3.12).[9] We have

$$\xi_t\left(\frac{i-j}{N}\right)(x_j(0) - y_j(0)) - \xi_t\left(\frac{i+1-j}{N}\right)(x_j(0) - y_j(0))$$

(A.5)

$$= (x_j(0) - y_j(0)) \int_{(i-j)/N}^{(i-j+1)/N} \partial_x \xi_t(x)\, dx.$$

We now use rigidity and the estimate[10] $\partial_x \xi_t(x) \le CN^{-1}(x^2 + t^2)^{-1}$ to obtain

(A.6) $$\text{(A.3)} \le \frac{N^\varepsilon}{N}\frac{1}{N}\int_{-tN^{1+\delta}}^{tN^{1+\delta}} \frac{1}{x^2+t^2}\, dx \le C\frac{N^\varepsilon}{N^2 t},$$

as desired.

REFERENCES

[1] BEN AROUS, G. and BOURGADE, P. (2013). Extreme gaps between eigenvalues of random matrices. *Ann. Probab.* **41** 2648–2681. MR3112927 https://doi.org/10.1214/11-AOP710
[2] BENAYCH-GEORGES, F. and KNOWLES, A. (2018). Lectures on the local semicircle law for Wigner matrices. Preprint. Available at arXiv:1601.04055.
[3] BOURGADE, P. (2019). Extreme gaps between eigenvalues of Wigner matrices. Preprint. Available at arXiv:1812.10376.
[4] BOURGADE, P., ERDŐS, L., YAU, H.-T. and YIN, J. (2016). Fixed energy universality for generalized Wigner matrices. *Comm. Pure Appl. Math.* **69** 1815–1881. MR3541852 https://doi.org/10.1002/cpa.21624
[5] BOURGADE, P. and YAU, H.-T. (2017). The eigenvector moment flow and local quantum unique ergodicity. *Comm. Math. Phys.* **350** 231–278. MR3606475 https://doi.org/10.1007/s00220-016-2627-6
[6] CHATTERJEE, S. (2006). A generalization of the Lindeberg principle. *Ann. Probab.* **34** 2061–2076. MR2294976 https://doi.org/10.1214/009117906000000575
[7] CHE, Z. and LOPATTO, P. (2019). Universality of the least singular value for sparse random matrices. *Electron. J. Probab.* **24** Paper No. 9, 53. MR3916329 https://doi.org/10.1214/19-EJP269
[8] DIACONIS, P. (2003). Patterns in eigenvalues: The 70th Josiah Willard Gibbs lecture. *Bull. Amer. Math. Soc. (N.S.)* **40** 155–178. MR1962294 https://doi.org/10.1090/S0273-0979-03-00975-3
[9] ERDŐS, L., KNOWLES, A., YAU, H.-T. and YIN, J. (2012). Spectral statistics of Erdős–Rényi Graphs II: Eigenvalue spacing and the extreme eigenvalues. *Comm. Math. Phys.* **314** 587–640. MR2964770 https://doi.org/10.1007/s00220-012-1527-7
[10] ERDŐS, L., KNOWLES, A., YAU, H.-T. and YIN, J. (2013). The local semicircle law for a general class of random matrices. *Electron. J. Probab.* **18** no. 59, 58. MR3068390 https://doi.org/10.1214/EJP.v18-2473

[9]This idea has appeared previously in [7], Section 3.6.
[10]Note that [27], Proposition 3.2, is missing the factor of N^{-1}. The statement of Lemma 3.12 later in the work, from which Proposition 3.2 is derived, is correct.

[11] ERDŐS, L., KNOWLES, A., YAU, H.-T. and YIN, J. (2013). Spectral statistics of Erdős–Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** 2279–2375. MR3098073 https://doi.org/10.1214/11-AOP734

[12] ERDŐS, L., PÉCHÉ, S., RAMÍREZ, J. A., SCHLEIN, B. and YAU, H.-T. (2010). Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63** 895–925. MR2662426 https://doi.org/10.1002/cpa.20317

[13] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2010). Wegner estimate and level repulsion for Wigner random matrices. *Int. Math. Res. Not. IMRN* **2010** 436–479. MR2587574 https://doi.org/10.1093/imrn/rnp136

[14] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2011). Universality of random matrices and local relaxation flow. *Invent. Math.* **185** 75–119. MR2810797 https://doi.org/10.1007/s00222-010-0302-7

[15] ERDŐS, L. and YAU, H.-T. (2012). Universality of local spectral statistics of random matrices. *Bull. Amer. Math. Soc. (N.S.)* **49** 377–414. MR2917064 https://doi.org/10.1090/S0273-0979-2012-01372-1

[16] ERDŐS, L. and YAU, H.-T. (2015). Gap universality of generalized Wigner and β-ensembles. *J. Eur. Math. Soc. (JEMS)* **17** 1927–2036. MR3372074 https://doi.org/10.4171/JEMS/548

[17] ERDŐS, L. and YAU, H.-T. (2017). *A Dynamical Approach to Random Matrix Theory. Courant Lecture Notes in Mathematics* **28**. Courant Institute of Mathematical Sciences, New York. MR3699468

[18] ERDŐS, L., YAU, H.-T. and YIN, J. (2012). Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields* **154** 341–407. MR2981427 https://doi.org/10.1007/s00440-011-0390-3

[19] ERDŐS, L., YAU, H.-T. and YIN, J. (2012). Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math.* **229** 1435–1515. MR2871147 https://doi.org/10.1016/j.aim.2011.12.010

[20] FENG, R., TIAN, G. and WEI, D. (2019). Small gaps of GOE. *Geom. Funct. Anal.* **29** 1794–1827. MR4034920 https://doi.org/10.1007/s00039-019-00520-5

[21] FENG, R. and WEI, D. (2018). Large gaps of CUE and GUE. Preprint. Available at arXiv:1807.02149.

[22] FENG, R. and WEI, D. (2018). Small gaps of circular β-ensemble. Preprint. Available at arXiv:1806.01555.

[23] FIGALLI, A. and GUIONNET, A. (2016). Universality in several-matrix models via approximate transport maps. *Acta Math.* **217** 81–176. MR3646880 https://doi.org/10.1007/s11511-016-0142-4

[24] HUANG, J., LANDON, B. and YAU, H.-T. (2015). Bulk universality of sparse random matrices. *J. Math. Phys.* **56** 123301, 19. MR3429490 https://doi.org/10.1063/1.4936139

[25] KNOWLES, A. and YIN, J. (2013). Eigenvector distribution of Wigner matrices. *Probab. Theory Related Fields* **155** 543–582. MR3034787 https://doi.org/10.1007/s00440-011-0407-y

[26] KORADA, S. B. and MONTANARI, A. (2011). Applications of the Lindeberg principle in communications and statistical learning. *IEEE Trans. Inf. Theory* **57** 2440–2450. MR2809100 https://doi.org/10.1109/TIT.2011.2112231

[27] LANDON, B., SOSOE, P. and YAU, H.-T. (2019). Fixed energy universality of Dyson Brownian motion. *Adv. Math.* **346** 1137–1332. MR3914908 https://doi.org/10.1016/j.aim.2019.02.010

[28] LANDON, B. and YAU, H.-T. (2017). Convergence of local statistics of Dyson Brownian motion. *Comm. Math. Phys.* **355** 949–1000. MR3687212 https://doi.org/10.1007/s00220-017-2955-1

[29] LOPATTO, P. and LUH, K. (2019). Tail bounds for gaps between eigenvalues of sparse random matrices. Preprint. Available at arXiv:1901.05948.

[30] LUH, K. and VU, V. (2018). Sparse random matrices have simple spectrum. Preprint. Available at arXiv:1802.03662.

[31] NGUYEN, H., TAO, T. and VU, V. (2017). Random matrices: Tail bounds for gaps between eigenvalues. *Probab. Theory Related Fields* **167** 777–816. MR3627428 https://doi.org/10.1007/s00440-016-0693-5

[32] ODLYZKO, A. M. (1987). On the distribution of spacings between zeros of the zeta function. *Math. Comp.* **48** 273–308. MR0866115 https://doi.org/10.2307/2007890

[33] TAO, T. and VU, V. (2010). Random matrices: Universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.* **298** 549–572. MR2669449 https://doi.org/10.1007/s00220-010-1044-5

[34] TAO, T. and VU, V. (2011). Random matrices: Universality of local eigenvalue statistics. *Acta Math.* **206** 127–204. MR2784665 https://doi.org/10.1007/s11511-011-0061-3

[35] VINSON, J. P. (2001). Closest Spacing of Eigenvalues. Ph.D. thesis, Princeton University, Princeton, NJ. MR2702214