The Annals of Applied Probability
2020, Vol. 30, No. 6, 2769–2795
https://doi.org/10.1214/20-AAP1572
© Institute of Mathematical Statistics, 2020

APPLICATIONS OF MESOSCOPIC CLTS IN RANDOM MATRIX THEORY

BY BENJAMIN LANDON[1] AND PHILIPPE SOSOE[2]

[1]*Department of Mathematics, Massachusetts Institute of Technology, blandon@mit.edu*
[2]*Department of Mathematics, Cornell University, ps934@cornell.edu*

We present some applications of central limit theorems on mesoscopic scales for random matrices. When combined with the recent theory of "homogenization" for Dyson Brownian motion, this yields the universality of quantities which depend on the behavior of single eigenvalues of Wigner matrices and β-ensembles. Among the results we obtain are the Gaussian fluctuations of single eigenvalues for Wigner matrices (without an assumption of 4 matching moments) and classical β-ensembles ($\beta = 1, 2, 4$), Gaussian fluctuations of the eigenvalue counting function, and an asymptotic expansion up to order $o(N^{-1})$ for the expected value of eigenvalues in the bulk of the spectrum. The latter result solves a conjecture of Tao and Vu.

1. Introduction. The object of this note is to show how two recent results in random matrix theory, central limit theorems for linear statistics of eigenvalues on mesoscopic scales (see [3, 21]), and homogenization for Dyson Brownian motion (DBM) introduced in [6] and refined in [25], can be used to derive a number of results concerning the fluctuations of single eigenvalues. In particular, we obtain the following results:

1. Gaussian fluctuations on scale $\frac{\sqrt{\log N}}{N}$ for individual eigenvalues of Wigner matrices in the bulk,
2. Gaussian fluctuation on scale $\frac{\sqrt{\log N}}{N}$ for individual eigenvalues of classical β-ensembles in the bulk ($\beta = 1, 2, 4$),
3. Gaussian fluctuations for eigenvalue counting functions,
4. asympotic expansion of the mean and variance of eigenvalues in the bulk, up to an error of order $o(N^{-1})$.

Here, N denotes the dimension of the random matrix ensemble under consideration. Item 1 was first obtained for the Gaussian unitary ensemble (GUE) by Gustavsson [20], and then for complex Hermitian Wigner matrices whose moments match the GUE to fourth order by Tao and Vu [34]. O'Rourke [29] extended this to the Gaussian orthogonal and symplectic ensembles (GOE/GSE) by applying a result of Forrester and Rains [19] (and then to Wigner matrices of these symmetry classes with four matching moments). Gustaffson originally first proved Item 3 and then obtained Item 1 via the duality

$$\big|\{i : \lambda_i < y\}\big| \geq k \quad \Longleftrightarrow \quad \lambda_k < y,$$

where the $\{\lambda_i\}_i$ are random matrix eigenvalues in increasing order. Dallaporta and Vu [11] used this to obtain the central limit theorem for the eigenvalue counting function of Wigner matrices (of those with four moments matching the Gaussian ensembles) and additionally computed the asymptotic expectation and variance of this quantity.

In the present work we remove all moment matching assumptions in Items 1 and 3. Moreover our approach also applies to β-ensembles and we obtain the Gaussian fluctuations for

Received December 2018; revised November 2019.
MSC2020 subject classifications. 60F05.
Key words and phrases. Random matrix theory, universality, mesoscopic linear statistics.

the eigenvalues of the classical β-ensembles. Our proof is based on *universality*, meaning that we show that the eigenvalue fluctuations are universal, coinciding with the Gaussian case. In the β-ensemble case, this approach applies for even nonclassical values of $\beta \geq 1$, and so if the analog of Gustavsson's result were known for general Gaussian β-ensembles, then we would obtain Item 2 for all $\beta \geq 1$.

The formula we obtain for the expected value of a single bulk eigenvalue in Item 4 solves a conjecture of Tao and Vu [33], Conjecture 1.7, and extends it to the real symmetric case.

1.1. *Homogenization of Dyson Brownian motion.* Our work relies on a recent technical innovation in the study of Dyson Brownian motion (DBM), known as "homogenization." DBM is a stochastic process on random matrices first introduced by Dyson [13]. It was first applied by Erdős, Schlein and Yau [16] who introduced DBM as a tool to study local eigenvalue fluctuations and prove Wigner–Dyson–Mehta conjecture for all symmetry classes.

Homogenization relies on a coupling between two Dyson Brownian motions $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$. The first process $\boldsymbol{x}(t)$ has initial data given by the Wigner matrix whose eigenvalues we wish to study, and $\boldsymbol{y}(t)$ comes from the equilibrium Gaussian ensemble that we would like to compare $\boldsymbol{x}(t)$ to. The difference between these two processes satisfies a discrete parabolic equation; the work [6] of Bourgade, Erdős, Yau and Yin establishes a homogenization theory for this parabolic equation, resulting in the estimate

$$x_i(t) - y_i(t) = \frac{1}{N}\zeta_x - \frac{1}{N}\zeta_y + \mathcal{O}\big(N^{-1-c}\big) \tag{1.1}$$

for some $c > 0$ with probability $1 - o(1)$. Here,

$$\zeta_x = \sum_{j=1}^{N} \Phi\big(x_j(0)\big) - \sum_{j=1}^{N} \Phi(\gamma_j), \qquad \zeta_y = \sum_{j=1}^{N} \Phi\big(y_j(0)\big) - \sum_{j=1}^{N} \Phi(\gamma_j) \tag{1.2}$$

for some explicit, N-dependent function Φ. Above, the γ_j are the quantiles of the semicircle density, the limit of the empirical eigenvalue density of $\boldsymbol{x}(0)$ and $\boldsymbol{y}(0)$, and so the second sums are roughly the expected value of the first sums. The function G is constructed from the heat kernel of a nonlocal operator that arises as a continuum approximation to the equation satisfied by $\boldsymbol{x} - \boldsymbol{y}$ (hence the name "homogenization").

The linear statistic (1.2) is on the scale t which must satisfy $1 \gg t \gg N^{-1}$ for the estimate (1.1) to hold. This scale is in between the microscopic and global scales, and so the quantity (1.2) is known as a mesoscopic linear statistic.

The representation (1.1) allows for a precise description of the fluctuations of quantities depending on single eigenvalues $x_i(t)$, provided one has good control of mesoscopic linear statistics, specifically, the quantities (1.2). At the time of the publication of [6], such a general result was not yet available, and the authors found that it sufficed, for their intended application, to obtain a result about mesoscopic linear statistics for the GOE. The intended application of homogenization in [6] was to the local correlation functions, which concern the scale N^{-1} and differs crucially from the $\sqrt{\log(N)}N^{-1}$ of the single eigenvalue fluctuations. For this reason, it was not necessary to treat the quantities (1.2) for the non-GOE case.

In [25], the authors of the present work (with Yau) revisited the argument in [6] and derived a refined homogenization result depending only on local assumptions on the initial data. We also obtained precise control on linear statistics of observables such as Φ in (1.2) in the case of deformed Wigner matrices. The derivation of our main results will follow from a combination of (1.1) and a strengthened version of the mesocopic CLT we proved in [25].

1.2. *Mesoscopic linear statistics.* Let $\lambda_1 \le \lambda_2 \cdots \le \lambda_N$, denote the eigenvalues of a Wigner matrix. It is by now classical that for smooth functions f, the quantity

$$\sum_{j=1}^{N} f(\lambda_j) - \mathbb{E}\left[\sum_{j=1}^{N} f(\lambda_j)\right]$$

is asymptotically Gaussian for smooth functions f. A general result applying to sufficiently regular functions appears in the work of Lytova and Pastur [28]. For compactly supported or rapidly decaying functions f, $0 < \alpha < 1$ and an energy $E \in (-2, 2)$,

$$\sum_{j=1}^{N} f\big(N^{\alpha}(\lambda_j - E)\big) - \mathbb{E}\left[\sum_{j=1}^{N} f\big(N^{\alpha}(\lambda_j - E)\big)\right] \tag{1.3}$$

is asymptotically Gaussian, with limiting variance

$$\frac{c_{\mathrm{sym}}}{2\pi^2} \int\int \left(\frac{f(x) - f(y)}{x - y}\right)^2 \mathrm{d}x\,\mathrm{d}y. \tag{1.4}$$

Here $c_{\mathrm{sym}} = \frac{1}{2}$ for Hermitian Wigner matrices, and $c_{\mathrm{sym}} = 1$ for real symmetric Wigner matrices. Because of the rescaling by N^{α}, the sums (1.3) typically involve about $N^{1-\alpha}$ eigenvalues, and consequently such linear statistics are refered to as *mesoscopic*. This result has been known in restricted cases for some time (see [8, 9, 27]). In [21], He and Knowles obtain the result for general Wigner matrices and general f. See also the work of Bekerman and Lodhia [3] for the case of β-ensembles. We gave proofs of more general results in [25] in the deformed Wigner case, and one-cut β-ensembles.

The method we use to derive our mesoscopic CLT represents a considerable simplification over the proof offered in [25]. Moreover, the result we derive here is stronger than that in [25], which had some restrictions on the scale of the function under consideration—this was an artifact of the proof given there, and absent from the present work.

This strengthening is in large part due to the fact that we are dealing with Wigner matrices, rather than the deformed Wigner matrices of our previous paper. Our result improves on the known results for Wigner matrices [3, 21] in that we allow the function f to not have compact support or decay quickly at infinity. As a consequence, the variance of the linear statistic may grow logarithmically with N. This is essential for our application to (1.2)—the function G turns out to roughly behave as a smoothed out step function on the mesoscopic scale t, and for such a function the quantity (1.4) is $\mathcal{O}(\log(N))$. Note that the size of the fluctuations of the linear statistics (1.2) and the single eigenvalues are the same, and so the mesoscopic central limit theorem is necessary for treating the single eigenvalue fluctuations via the homogenization approach. Such a CLT was not known prior to the present work.

The proof of the CLT, Proposition 4.1, is inspired by the method of Shcherbina [31], but uses the cumulant expansion (as in [21]) instead of the more intricate resolvent expansion in [25]. We also rely on the isotropic local law of [4, 23] to simplify the treatment of some error terms (this is not strictly necessary as such a result was not available for [25]).

1.3. *Statement of results.* For simplicity, we state all our results for Wigner matrices in the case when H is real and symmetric. The same proofs apply with only minor modification to random Hermitian matrices. Where relevant, we comment on the corresponding results for the complex Hermitian case. In this case however, many of results below can be derived using the Brézin–Hikami formula [10], requiring neither of the techniques highlighted in the Introduction.

Let ξ_o and ξ_d be two real centered random variables with bounded moments of all orders, with the variance of ξ_o being 1. A real symmetric Wigner matrix is an $N \times N$ self adjoint matrix so that the entries $\{H_{ij}\}_{i \le j}$ are independent, and

$$\sqrt{N} H_{ii} \sim \xi_d, \qquad \sqrt{N} H_{ij} \sim \xi_o, \quad i \neq j.$$

We also use the following notation for the cumulants of the matrix entries:

$$\begin{aligned} s_k &:= \frac{1}{i^k} \frac{\mathrm{d}^k}{\mathrm{d}^k t} \log \mathbb{E}[e^{it\xi_o}]\Big|_{t=0}, \\ s_k + a_k &:= \frac{1}{i^k} \frac{\mathrm{d}^k}{\mathrm{d}^k t} \log \mathbb{E}[e^{it\xi_d}]\Big|_{t=0}. \end{aligned} \tag{1.5}$$

Our notational choice for the second line of (1.5) is so that the cumulant expansion we use to derive our mesoscopic central limit theorem has a simple form (see the proof of Lemma 4.5).

We denote by $\lambda_1 \le \cdots \le \lambda_N$ be the ordered eigenvalues of H. As a consequence of the local semicircle law, Theorem 2.2, λ_i is typically very close to the ith N-quantile γ_i of the semicircle distribution, also known as the "classical location" of λ_i. The semicircle and classical eigenvalue locations are defined by

$$\rho_{\mathrm{sc}}(x) = \frac{1}{2\pi} \sqrt{(4 - x^2)_+}, \qquad \int_{-2}^{\gamma_i} \rho_{\mathrm{sc}}(x)\, \mathrm{d}x = \frac{i}{N}. \tag{1.6}$$

Our first result describes the fluctuations of λ_i about γ_i.

THEOREM 1.1. *Let $\kappa > 0$, and let λ_i be the eigenvalues of a symmetric Wigner matrix. For any sequence $i = i_N$ such that $\kappa N \le i \le (1-\kappa)N$ we have*

$$N \frac{\lambda_i - \gamma_i}{\sqrt{\frac{\log(N)}{1-\gamma_i^2/4}}} \xrightarrow{d} N(0,1). \tag{1.7}$$

In the case of complex Hermitian Wigner matrices, the same result holds with limiting variance $\frac{1}{2}$.

As explained in the Introduction, this has been obtained only for the Gaussian ensembles and those Wigner matrices whose first four moments match the Gaussian ensembles [20, 29, 34]. Our contribution is thus to remove the assumption of matching moments. Theorem 1.1 has also been obtained independently by Bourgade and Krishnan [7] by different methods.

The result of Gustavsson holds for any sequence of indices i_N such that $i_N \to \infty$ and $|N - i_N| \to \infty$, that is, for indices close to the edge. We expect that our universality result holds for any sequences such that $\kappa_N N \le i_N \le (1-\kappa_N)N$ with $\kappa_N \ge N^{-c}$ for some sufficiently small $c > 0$. We do not pursue this here in the interest of brevity. Results for the case of moderately growing i_N follow from edge universality. For example, Gaussian fluctuations in the case that $i_N \to \infty$ but $i_N \le N^c$ for some sufficiently small $c > 0$, follow from [5].

By a well-known argument (see [11, 20] or below), the result on fluctuations of eigenvalues implies a CLT for indicator functions:

COROLLARY 1.2. *For $E \in (-2, 2)$ let $\mathcal{N}(E) = \#\{i : \lambda_i \le E\}$ be the eigenvalue counting function of a real symmetric Wigner matrix. Then,*

$$\frac{\mathcal{N}(E) - N \int_{-\infty}^{E} \mathrm{d}\rho_{\mathrm{sc}}}{\pi^{-1}\sqrt{\log(N)}} \xrightarrow{d} N(0,1). \tag{1.8}$$

A similar method gives the following for classical β-ensembles in the one-cut case. We recall that a β-ensemble with potential $V : \mathbb{R} \to \mathbb{R}$ is the measure μ_V on the simplex $\{\lambda_1 < \lambda_2 < \cdots < \lambda_N \subseteq \mathbb{R}^N$ proportional to the Gibbs weight with Hamiltonian $\mathcal{H}$,

$$\mu_V \to \mathrm{e}^{-\beta N \mathcal{H}}, \qquad \mathcal{H} := \frac{1}{2}\sum_{j=1}^{N} V(\lambda_j) - \frac{1}{N}\sum_{i<j}\log(\lambda_j - \lambda_i)$$

where V is assumed to satisfy a growth condition so that this measure is finite (see, eq., [1]). For sufficiently regular V, the empirical measure is known to converge a.s. to a deterministic probability measure we call the *equilibrium measure* and denote by $\rho^{(V)}$ [1, 24]. The one-cut case is the case that this measure is supported on a single interval.

THEOREM 1.3. *Let V be a one-cut β-ensemble with equilibrium measure supported in $[A, B]$ and $V \in C^4$. Assume that V is regular in the sense of* [24]. *Denote the particles by λ_i. Let $i = i_N$ be a sequence of indices satisfying $\kappa N \le i \le (1-\kappa)N$ for fixed $\kappa > 0$. Then, for $\beta = 1, 2, 4$ we have*

$$\frac{\beta^{1/2} N \rho^{(V)}(\gamma_i^{(V)})\pi}{2\sqrt{\log(N)}}\big(\lambda_i - \gamma_i^{(V)}\big) \xrightarrow{d} N(0,1), \tag{1.9}$$

where $\rho^{(V)}$ is the equilibrium measure of V and $\gamma_i^{(V)}$ is the ith classical eigenvalue location.

If the result for $\beta \notin \{1, 2, 4\}$ were known for the Gaussian β-ensemble, then the above result would be true for all $\beta \ge 1$. Instead, we have the statement that for all $\beta \ge 1$ and smooth test functions F,

$$\begin{aligned} \lim_{N\to\infty} \big|&\mathbb{E}_V\big[F\big(N\rho^{(V)}(\gamma_i^{(V)})(\lambda_i - \gamma_i^{(V)})/\sqrt{\log(N)}\big)\big] \\ &- \mathbb{E}_{G\beta E}\big[F\big(N\rho_{\mathrm{sc}}(\gamma_i)(\lambda_i - \gamma_i)/\sqrt{\log(N)}\big)\big]\big| = 0. \end{aligned} \tag{1.10}$$

The expectation on the LHS is with respect to the β-ensemble with potential V and on the RHS we have the expectation with respect to the Gaussian β-ensemble.

As in Corollary 1.2 we also get a CLT for the eigenvalue counting function for β ensembles with classical values of β. A similar result for the indicator function $\mathbf{1}_{(a,b)}$ where $(a, b) \subseteq (A+\kappa, B-\kappa)$ was obtained by Shcherbina [32] for classical β, using Fredholm determinants.

Our method also allows for an asymptotic expansion of $\mathbb{E}[\lambda_i]$ for Wigner matrices up to order $o(N^{-1})$, resolving a conjecture of Tao and Vu [33]. A related result for Gaussian divisible ensembles in the complex Hermitian case was obtained by Edelman–Guionnet–Péché [14] using the Brézin–Hikami formula. This result finds an expansion for the quantiles of the expected density of states; however this result does not imply the corresponding expansion for the expected eigenvalue location.

THEOREM 1.4. *Fix $\kappa > 0$. Let H be a real symmetric Wigner matrix as above, and $\kappa N \le i \le (1-\kappa)N$. Then,*

$$N\mathbb{E}_H[\lambda_i - \gamma_i] = \frac{1}{2\pi\rho_{\mathrm{sc}}(\gamma_i)}\arcsin\left(\frac{\gamma_i}{2}\right) - \frac{1}{2\rho_{\mathrm{sc}}(\gamma_i)} + \frac{s_4}{4}(\gamma_i^3 - 2\gamma_i) + \frac{a_2 - 1}{2}\gamma_i + o(1) \tag{1.11}$$

and

$$\mathrm{Var}_H(\lambda_i) = \mathrm{Var}_{\mathrm{GOE}}(\lambda_i) + \frac{s_4}{8N^2}\gamma_i^2 + \frac{a_2 - 1}{N^2} + o(N^{-2}). \tag{1.12}$$

REMARK. Tao and Vu [33] conjectured in the complex Hermitian case a formula similar to (1.11). The formula we find sharpens their prediction, by identifying the quantity $C_{i,N}$ in their proposed asympotic $\mathbb{E}_{\mathrm{GUE}}[\lambda_i] = \gamma_i + N^{-1}C_{i,N} + o(N^{-1})$. Their main interest was the dependence of the expectation on the fourth moment of the entries, and we find exactly the dependence predicted in [33]. In the complex Hermitian analog of (1.11), the $\arcsin(\gamma_i)$ term is not present, and s_4 is replaced by the sum of the fourth cumulants of the real and imaginary parts of the off-diagonal entries of the matrix. The formula (1.12) appears new.

For a 1-cut β-ensemble with equilibrium distribution ρ_V supported on $[A, B]$ we obtain the following:

THEOREM 1.5. *Fix $\kappa > 0$. Let i be an index so that $\kappa N \le i \le (1-\kappa)N$. We have*

$$N\mathbb{E}_V\big[\rho_V\big(\gamma_i^{(V)}\big)\big(\lambda_i - \gamma_i^{(V)}\big)\big] = -\frac{1}{2} + \frac{1}{2\pi^2}\left(\frac{2}{\beta} - 1\right)\int_A^{\gamma_i^{(V)}} \mathrm{d}\nu_V(x), \tag{1.13}$$

where $\nu_V(x)$ is a specific signed measure, described in more detail below.

The mesoscopic central limit theorem we present, Theorem 4.3, is fairly robust. To illustrate this, we state an extension of some previous results on "partial linear statistics" of Bao, Pan and Zhou [2]. Specifically, we weaken the regularity assumptions on the functions f appearing there, and do not require the moment matching hypotheses of [2].

THEOREM 1.6. *Let H be a real symmetric Wigner matrix. Fix $u \in (-2, 2)$. Let f be a C^3 function such that $f'(x) \neq 0$ only for $x \in (-2+\kappa, 2-\kappa)$, and $f(u) = 0$. Then,*

$$\sum_i f(\lambda_i)\mathbf{1}_{\{\lambda_i \le u\}} - \mathbb{E}\left[\sum_i f(\lambda_i)\mathbf{1}_{\{\lambda_i \le u\}}\right] \xrightarrow{d} N\big(0, \sigma^2\big), \tag{1.14}$$

where σ^2 is an explicit function of f. In fact,

$$\sigma^2 = V\big(f(\cdot)\mathbf{1}_{\cdot \le u}\big),$$

where V is defined in (4.7).

Suppose $k/N \to u \in (-2, 2)$. Then

$$\sum_{i=1}^k f(\lambda_i) - \mathbb{E}\left[\sum_{i=1}^k f(\lambda_i)\right] \to N\big(0, \sigma'^2\big), \tag{1.15}$$

where $\sigma'^2 = V(f(\cdot)\mathbf{1}_{\cdot \le \tilde{\gamma}_u})$, where the function $\tilde{\gamma}_u$ is defined by

$$\int_{-2}^{\tilde{\gamma}_u} \rho_{\mathrm{sc}}(x)\,\mathrm{d}x = u.$$

REMARK. If $f(u) \neq 0$, the quantity (1.14) has variance of order $\log N$; a central limit theorem for this case follows from our result for the eigenvalue counting function. Further, we introduced the notation $\tilde{\gamma}_u$ as the inverse of the cumulative distribution function of the semicircle distribution so as to not conflict with our earlier notation for the N-quantiles γ_k. Note,

$$\gamma_k = \tilde{\gamma}_{k/N}.$$

2. The local semicircle law. Let H be a Wigner matrix as above. The empirical distribution of the eigenvalues of such a matrix converges to the semicircle distribution (1.6). The *local semicircle law*, which will use throughout without further comment represents a considerable strengthening of this statement. It is expressed in terms of the *resolvent matrix*

$$G(z) = (H - z)^{-1}, \quad \operatorname{Im} z > 0.$$

Of particular importance is the normalized trace of $G(z)$, because it equals the Stieltjes transform of the empirical eigenvalue distribution:

$$m_N(z) = \frac{1}{N} \operatorname{tr} G(z) = \frac{1}{N} \sum_{j=1}^{N} \frac{1}{\lambda_j - z}.$$

The semicircle law is equivalent to

$$m_N(z) = m(z) + o(1),$$

where

$$m(z) = \int \frac{1}{x - z} \rho_{\mathrm{sc}}(x)\, \mathrm{d}x = \frac{-z + \sqrt{z^2 - 4}}{2}.$$

The version of the local semicircle law we state here is taken from [22], Theorem 2.6, which is a detailed, pedagogical treatment of the semicircle law and its applications. To state it we introduce the following notion of overwhelming probability.

DEFINITION 2.1. We say that an event or family of events $\{\mathcal{A}_i\}_{i \in \mathcal{I}}$ hold with overwhelming probability if for all $D > 0$ there is an N_D, so that $\sup_{i \in \mathcal{I}} \mathbb{P}[A_i^c] \le N^{-D}$ for $N \ge N_D$.

THEOREM 2.2 (Local semicircle law). *Define*

$$\mathbf{S} = \{E + i\eta : |E| \le 10, 0 < \eta \le 10\}.$$

Then, for each $\epsilon > 0$ and each $D > 0$ large, and all N sufficiently large, we have

$$\max_{i,j} |G_{ij}(z) - \delta_{ij} m(z)| \le \sqrt{\frac{\operatorname{Im} m(z)}{N^{1-\epsilon}\eta}} + \frac{1}{N^{1-\epsilon}\eta}$$

and

$$|m_N(z) - m(z)| \le \frac{1}{N^{1-\epsilon}\eta}$$

uniformly in $z \in \mathbf{S}$ with overwhelming probability.

One consequence of the semicircle is that the eigenvalues λ_i are close to the classical location (1.6).

THEOREM 2.3 (Eigenvalue rigidity). *For each $\epsilon > 0$ and $D > 0$, we have*

$$|\lambda_i - \gamma_i| \le N^{-2/3+\epsilon} \min\{i, (N + 1 - i)\}^{-1/3}$$

uniformly in i with overwhelming probability.

3. Homogenization for DBM.

3.1. *Wigner matrices.* In this section we present the homogenization result of [6] for Wigner matrices. We will need first the following definition.

DEFINITION 3.1. We say that X is a Wigner matrix with Gaussian component of size t_0 if,

$$X = \mathrm{e}^{-t_0/2} X' + \sqrt{1-\mathrm{e}^{-t_0}}\, W, \tag{3.1}$$

where X' is a Wigner matrix and W is an independent GOE.

Given a Wigner matrix X (which we will eventually take to have a Gaussian component) and a GOE matrix W' independent of X, we define the following system of coupled SDEs. First we define,

$$\mathrm{d}x_i(t) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j\neq i}\frac{1}{x_i - x_j}\,\mathrm{d}t - \frac{x_i}{2}\,\mathrm{d}t, \quad x_i(0) = \lambda_i(X) \tag{3.2}$$

where the B_i are standard Brownian motions. Using the same Brownian motion terms, we define

$$\mathrm{d}y_i(t) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j\neq i}\frac{1}{y_i - y_j}\,\mathrm{d}t - \frac{y_i}{2}\,\mathrm{d}t, \quad y_i(0) = \lambda_i(W'). \tag{3.3}$$

It is well known that for each time t, the vector $x_i(t)$ is distributed as the eigenvalues of the matrix $\mathrm{e}^{-t/2}X + \sqrt{1-\mathrm{e}^{-t}}\,W''$ where W'' is an independent GOE matrix (see, e.g., the classic calcuation of Dyson [13] or Section 4.3 of [1] for a pedagogical treatment). Consequently, the vector $y_i(t)$ is distributed as the eigenvalues of a GOE matrix for every t.

The following homogenization result is Theorem 3.2 of [6].

THEOREM 3.2. *Let $\kappa > 0$ and fix an index i with $\kappa N \le i \le (1-\kappa)N$. There are constants $\tau_0 < 1/4$ and δ_1, δ_2 so that the following. Suppose that X is a Wigner matrix with Gaussian component of size $t_0 = N^{-\tau_0}$. Let W be a GOE matrix independent of X and consider the coupled system of SDEs defined above. Then, with probability at least $1 - N^{-\delta_1}$ and $t_1 = t_0/2$,*

$$x_i(t_1) - y_i(t_1) = \frac{1}{N}\sum_j p_{t_1}(\gamma_i, \gamma_j)\big(x_j(0) - y_j(0)\big) + \mathcal{O}\Big(\frac{1}{N^{1+\delta_2}}\Big). \tag{3.4}$$

The function $p_{t_1}(x, y)$ is smooth and its properties are given below.

REMARK. We could have also used instead the homogenization result of [25], which would have yielded the above estimate with much higher probability. However, the result of [25] is presented in a general setting, and it would take some exposition to specialize it to the simpler Wigner case. Moreover, we will not need the stronger result proved there.

The function $p_s(x, y)$ is defined on $[-2, 2]^2$ is defined explicitly in (3.22) of [6]. It is smooth and obeys the estimates [6],

$$0 \le p_s(\gamma_i, \gamma_j) \le \frac{Cs}{s^2 + (\gamma_i - \gamma_j)^2}, \qquad \big|\partial_x p_s(\gamma_i, x)\big| \le \frac{C}{s^2 + (\gamma_i - x)^2}. \tag{3.5}$$

In preparation for the application of the mesoscopic central limit theorem, we show how to rewrite the sum on the RHS as a linear statistic. Choose an $\varepsilon_1 > 0$ so that $0 < \varepsilon_1 < \tau_0/10$. Using rigidity, and the first estimate of (3.5) we see that,

$$\frac{1}{N}\sum_j p_{t_1}(\gamma_i,\gamma_j)(x_j(0)-\gamma_j) = \frac{1}{N}\sum_j p_{t_1}(\gamma_i,\gamma_j)\chi\left(\frac{\gamma_j-\gamma_i}{t_1 N^{\varepsilon_1}}\right)(x_j(0)-\gamma_j) + \mathcal{O}\left(\frac{1}{N^{1+\varepsilon_1/5}}\right) \tag{3.6}$$

with overwhelming probability. Here, χ is a smooth indicator function that is 1 for $|y| \le 1$ and 0 for $|y| \ge 2$. An explicit calculation using (3.22) of [6] yields,

$$p_{t_1}(x,y) = \frac{1}{\rho_{\mathrm{sc}}(\gamma_i)}\frac{t_1\rho_{\mathrm{sc}}(\gamma_i)}{(x-y)^2+(t_1\rho_{\mathrm{sc}}(\gamma_i))^2}(1+\mathcal{O}(N^{-\varepsilon_1/5}), \quad |x-\gamma_i|+|y-\gamma_i| \le 2t_1N^{\varepsilon_1}. \tag{3.7}$$

Note that we left uncancelled the factor $\rho_{\mathrm{sc}}(\gamma_i)$ to show the connection of $p_{t_1}(x,y)$ to the Poisson kernel $\frac{\eta}{(x-y)^2+\eta^2}$. Defining,

$$\Phi(y) = \int_{-2}^{y}\chi\left(\frac{x-\gamma_i}{t_1N^{\varepsilon_1}}\right)\frac{1}{\rho_{\mathrm{sc}}(\gamma_i)}\frac{t_1\rho_{\mathrm{sc}}(\gamma_i)}{(x-\gamma_i)^2+(t_1\rho_{\mathrm{sc}}(\gamma_i))^2}\,\mathrm{d}x, \tag{3.8}$$

we see that, by a similar argument as Section 4 of [6] (near (4.57–4.58)) that on the event that the conclusion of Theorem 3.2 holds, that

$$x_i(t_1) - y_i(t_1) = \frac{1}{N}\zeta_x - \frac{1}{N}\zeta_y + \mathcal{O}(N^{-1-c}) \tag{3.9}$$

for some $c > 0$, where,

$$\frac{1}{N}\zeta_x = \frac{1}{N}\sum_j \Phi(x_j) - \Phi(\gamma_j). \tag{3.10}$$

For later use, we note that for $|x-\gamma_i| > t_1N^{\varepsilon_1}$,

$$\Phi(x) = \frac{1}{\rho_{\mathrm{sc}}(\gamma_i)}\mathbf{1}_{\{x\ge\gamma_i\}} + o(1) \tag{3.11}$$

as well as the bounds,

$$\Phi(x) \le C, \qquad \|\Phi'(x)\|_1 \le C, \qquad \|\Phi''(x)\|_1 \le C\frac{N^{\varepsilon_1}}{t_1} \ll N. \tag{3.12}$$

3.2. *β-Ensembles*. In this section we discuss the analog of Theorem 3.2 for β-ensembles. Let us consider the following set-up. Fix an index i_0 with $\kappa N \le i_0 \le (1-\kappa)N$ with $\kappa > 0$, and two β-ensembles with potentials V and U. Denote their classical eigenvalue locations by $\gamma_i^{(V)}$ and $\gamma_i^{(U)}$, and equilibrium measure by $\rho^{(V)}$ and $\rho^{(U)}$. Assume that,

$$\gamma_{i_0}^{(V)} = \gamma_{i_0}^{(U)} = 0, \qquad \rho^{(V)}(0) = \rho^{(U)}(0) = \rho_{\mathrm{sc}}(0). \tag{3.13}$$

Let, $x_i(t)$ and $y_i(t)$ be the following coupled process:

$$\mathrm{d}x_i(t) = \sqrt{\frac{2}{N\beta}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j\ne i}\frac{1}{x_i-x_j}\,\mathrm{d}t - \frac{V'(x_i)}{2}\,\mathrm{d}t \tag{3.14}$$

and

$$\mathrm{d}y_i(t) = \sqrt{\frac{2}{N\beta}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j\neq i}\frac{1}{y_i - y_j}\,\mathrm{d}t - \frac{U'(y_i)}{2}\,\mathrm{d}t \tag{3.15}$$

where the B_i are standard Brownian motions, and the initial data are independent β-ensembles for potentials V and U, respectively. Note that the distribution of each of the sets of particles is invariant under t (of course the joint distribution of all $2N$ particles together is not).

The following theorem is a consequence of Section 8, specifically Theorem 8.15, of [25].

THEOREM 3.3. *Let $t_1 = N^{\omega_1}/N$ with $0 < \omega_1 < 1/10$. Then there are $\varepsilon, \delta_1, \delta_2 > 0$ so that the following estimate holds with probability at least $1 - N^{-\delta_1}$. We have*

$$x_{i_0}(t_1) - y_{i_0}(t_1) = \frac{1}{N}\sum_{|i_0 - j|\le N t_1 N^{\varepsilon}} \hat{p}_{t_1}\big(\gamma_{i_0}^{(\mathfrak{f})}, \gamma_j^{(\mathfrak{f})}\big)\big(x_j(0) - y_j(0)\big) + \mathcal{O}\big(N^{-1-\delta_2}\big). \tag{3.16}$$

In the above theorem, the function $\hat{p}$, as the theorem is written above, is not exactly the same as p appearing earlier, however it also obeys the estimates (3.5). Above $\gamma_j^{(\mathfrak{f})}$ are the "flattened" classical eigenvalue locations,

$$\gamma_i^{(\mathfrak{f})} = \frac{i}{N\rho_{\mathrm{sc}}(0)}. \tag{3.17}$$

Note that for $|j - i_0| \le \sqrt{N}$, the quantities $\gamma_j^{(\mathfrak{f})}$, $\gamma_j^{(U)}$ and $\gamma_j^{(V)}$ all differ by less than C/N. Similar arguments as to those in the Wigner case show that the sum on the RHS of (3.16) can be written as

$$\frac{1}{N}\sum_{|i-j|\le N t_1 N^{\varepsilon}} \hat{p}_{t_1}\big(\gamma_i^{(\mathfrak{f})}, \gamma_j^{(\mathfrak{f})}\big)\big(x_i(0) - y_i(0)\big) = \frac{1}{N}(\hat{\zeta}_x - \hat{\zeta}_y) + \mathcal{O}\big(N^{-1-c}\big), \tag{3.18}$$

where

$$\hat{\zeta}_x = \frac{1}{N}\sum_j \hat{\Phi}(x_i) - \hat{\Phi}\big(\gamma_i^{(V)}\big) \tag{3.19}$$

and similarly for $\hat{\zeta}_y$. Here $\hat{\Phi}$ obeys also (3.11) (with $\rho_{\mathrm{sc}}(\gamma_i)$ replaced by $\rho_{\mathrm{sc}}(0)$) and (3.12).

4. Linear statistics. In this section, we derive the mesoscopic central limit theorem we will use to prove our main results. Let H be a real symmetric Wigner matrix, and let a_k and s_k be as in (1.5). We will use frequently the local semicircle law discussed in Section 2.

4.1. *Statements.* Let $f_N(x) \in C^3$ be a sequence of test functions. We will drop the N-dependence from the notation and write $f = f_N$. We assume that there are $c, C > 0$ so that

$$\|f''\|_1 \le N^{1-c}, \qquad \|f'\|_1 + \|f\|_1 \le C, \tag{4.1}$$

where $\|f\|_p$ denotes the L^p norm. We also assume that

$$f'(x) \neq 0 \quad \text{only if} \quad x \in (-2+\kappa, 2-\kappa) \tag{4.2}$$

for some $\kappa > 0$. For definiteness we assume $\kappa < \frac{1}{10}$. Define, for λ,

$$\psi(\lambda) = \mathbb{E}\big[e(\lambda)\big], \quad e(\lambda) = \exp\big\{\mathrm{i}\lambda\big(\operatorname{tr} f(H) - \mathbb{E}\big[\operatorname{tr} f(H)\big]\big)\big\}. \tag{4.3}$$

The function ψ is the characteristic function of the centered linear statistic. Let $\chi(y)$ be a smooth function that is 1 for $|y| \le 1$ and 0 for $|y| \ge 2$. We use the quasi-analytic extension of f,

$$\tilde{f}(x+\mathrm{i}y) := \chi(y)\big(f(x) + \mathrm{i}y f'(x)\big). \tag{4.4}$$

In this section we make the following calculation.

PROPOSITION 4.1. *We have for any $\varepsilon > 0$,*

$$\partial_\lambda \psi(\lambda) = -\lambda\psi(\lambda)V(f) + \mathcal{O}\big(N^\varepsilon(1+|\lambda|^5)N^{-1/2}(1+\|f''\|_1^{1/2})\big), \tag{4.5}$$

where

$$\begin{aligned} V(f) := -\frac{1}{\pi^2}\int\int_{\mathbb{C}^2} \bar{\partial}_z \tilde{f}(z)\bar{\partial}_{z'}\tilde{f}(z')\frac{1}{z+2m(z)}\bigg\{2\partial_{z'}\frac{m(z')-m(z)}{z'-z} \\ + 2s_4 m^2(z)m(z')m'(z') + (a_2-1)m'(z')m(z)\bigg\}\,\mathrm{d}z\,\mathrm{d}z'. \end{aligned} \tag{4.6}$$

We can further calculate the variance and expectation.

LEMMA 4.2. *We have*

$$\begin{aligned} V(f) = \frac{1}{2\pi^2}\int_{-2}^{2}\int_{-2}^{2}\frac{(f(y)-f(x))^2}{(x-y)^2}\frac{4-xy}{\sqrt{4-x^2}\sqrt{4-y^2}}\,\mathrm{d}x\,\mathrm{d}y \\ + \frac{a_2-1}{4\pi^2}\bigg(\int_{-2}^{2} f(x)\frac{x}{\sqrt{4-x^2}}\,\mathrm{d}x\bigg)^2 \\ + \frac{s_4}{2\pi^2}\bigg(\int_{-2}^{2} f(x)\frac{2-x^2}{\sqrt{4-x^2}}\,\mathrm{d}x\bigg)^2. \end{aligned} \tag{4.7}$$

For the expectation we have for any $\varepsilon > 0$,

$$\begin{aligned} \mathbb{E}[\operatorname{tr} f(H)] = N\int_{-2}^{2} f(x)\rho_{\mathrm{sc}}(x)\,\mathrm{d}x + \frac{-1}{2\pi}\int_{-2}^{2}\frac{f(x)}{\sqrt{4-x^2}}\,\mathrm{d}x + \frac{f(2)+f(-2)}{4} \\ + \frac{1-a_2}{2\pi}\int_{-2}^{2} f(x)\frac{2-x^2}{\sqrt{4-x^2}}\,\mathrm{d}x \\ + \frac{s_4}{2\pi}\int_{-2}^{2} f(x)\frac{x^4-4x^2+2}{\sqrt{4-x^2}}\,\mathrm{d}x \\ + \mathcal{O}\big(N^\varepsilon N^{-1/2}(\|f''\|_1^{1/2}+1)\big). \end{aligned} \tag{4.8}$$

REMARK. The form of the variance (4.7) has appeared before in the literature. In the case that the Wigner ensemble matches the GOE to four moments, $a_2 = 1$ and $s_4 = 0$, and the formula matches that of Lytova and Pastur [28]. In the general case, this formula appeared in the work of Shcherbina [31].

Integrating (4.5) gives us the following theorem.

THEOREM 4.3. *Let f be as above and assume $V(f) \ge c$. Then*

$$\frac{\operatorname{tr} f(H) - \mathbb{E}[\operatorname{tr} f(H)]}{V(f)} \tag{4.9}$$

converges to a Gaussian with mean 0 *and variance* 1.

4.2. *Proof of Proposition* 4.1. We recall here the Hellfer–Sjöstrand formula (see [12] or Appendix C of [22]) which states that, we have

$$\begin{aligned}\operatorname{tr} f(H) - \mathbb{E}[\operatorname{tr} f(H)] &= \frac{1}{\pi}\int_{\mathbb{C}^2}\big(\partial_{\bar z}\tilde f(z)\big)\,\mathrm{d}z\,\mathrm{d}\bar z \\ &= \frac{1}{2\pi}\int_{\mathbb{R}^2}\big(\mathrm{i}y\chi(y)f''(x) + \mathrm{i}\big(f(x)+\mathrm{i}f'(x)y\big)\chi'(y)\big) \\ &\quad\times N\big(m_N(z) - \mathbb{E}[m_N(z)]\big)\,\mathrm{d}x\,\mathrm{d}y,\end{aligned} \tag{4.10}$$

where $\tilde f(x+\mathrm{i}y) := (f(x)+\mathrm{i}yf'(x))\chi(y)$ is a quasi-analytic extension of f and $\chi(y)$ is a smooth cut-off function that is 1 for $|y|<1$ and 0 for $|y|>2$. Fix a small $\mathfrak{a}>0$ and define the domain

$$\Omega_{\mathfrak{a}} := \{(x,y)\in\mathbb{R}^2 : |y| > N^{\mathfrak{a}-1}\}. \tag{4.11}$$

Using the fact that $y \to \operatorname{Im}[m_N(x+\mathrm{i}y)]y$ is an increasing function and that $m_N(\bar z) = \bar m_N(z)$ we get,

$$\begin{aligned}\operatorname{tr} f(H) - \mathbb{E}[\operatorname{tr} f(H)] &= \frac{1}{2\pi}\int_{\Omega_{\mathfrak{a}}}\big(\mathrm{i}y\chi(y)f''(x) + \mathrm{i}\big(f(x)+\mathrm{i}f'(x)y\big)\chi'(y)\big) \\ &\quad\times N\big(m_N(z) - \mathbb{E}[m_N(z)]\big)\,\mathrm{d}x\,\mathrm{d}y \\ &\quad + \mathcal{O}\big(N^{2\mathfrak{a}-1}\|f''\|_1\big)\end{aligned} \tag{4.12}$$

with overwhelming probability. Using this, we have

$$\begin{aligned}\psi'(\lambda) &= \frac{\mathrm{i}}{2\pi}\int_{\Omega_{\mathfrak{a}}}\big(\mathrm{i}y\chi(y)f''(x) + \mathrm{i}\big(f(x)+\mathrm{i}yf'(x)\big)\chi'(y)\big)E(z)\,\mathrm{d}x\,\mathrm{d}y \\ &\quad + \mathcal{O}\big(N^{2\mathfrak{a}-1}\|f''\|_1\big),\end{aligned} \tag{4.13}$$

where

$$E(z) = N\mathbb{E}\big[e(\lambda)\big(m_N(z) - \mathbb{E}[m_N(z)]\big)\big] = \sum_i \mathbb{E}\big[e(\lambda)\big(G_{ii} - \mathbb{E}[G_{ii}]\big)\big]. \tag{4.14}$$

The following lemma will be used repeatedly.

LEMMA 4.4. *Let $H(z)$ be a holomorphic function on $\mathbb{C}\setminus\mathbb{R}$. Suppose that the estimate*

$$|H(z)| \le \frac{K}{|\operatorname{Im}[z]|^s} \tag{4.15}$$

for some $1\le s\le 2$ whenever $|z-2|>\kappa/2$ or $|z+2|>\kappa/2$, where κ is as in the definition of f. There is a $C>0$ so that,

$$\begin{aligned}&\left|\int_{\Omega_{\mathfrak{a}}}\big(\mathrm{i}y\chi(y)f''(x) + \mathrm{i}\big(f(x)+\mathrm{i}yf'(x)\big)\chi'(y)\big)H(x+\mathrm{i}y)\,\mathrm{d}x\,\mathrm{d}y\right| \\ &\qquad \le CK\log(N)\big(1+\|f''\|_1\big)^{s-1}.\end{aligned} \tag{4.16}$$

PROOF. We have only to deal with the term involving $f''(x)$. For this, we fix a scale η_1 and integrate in parts for $|y|>\eta_1$, to find

$$\left|\int_{|y|>\eta_1} y\chi(y)f''(x)H(x+\mathrm{i}y)\,\mathrm{d}x\,\mathrm{d}y\right| = \left|\int_{|y|>\eta_1}\chi(y)f'(x)y(\partial_z H)(x+\mathrm{i}y)\,\mathrm{d}x\,\mathrm{d}y\right|, \tag{4.17}$$

where we used the Cauchy–Riemann equations. Since H is holomorphic away from the real axis we have by the Cauchy integral formula,

$$\left|\partial_z H(x+\mathrm{i}y)\right| \le \frac{CK}{|\operatorname{Im}[z]|^{s+1}}. \tag{4.18}$$

Hence,

$$\begin{aligned}\left|\int_{\Omega_{\mathfrak{a}}} y\chi(y)f''(x)H\,\mathrm{d}x\,\mathrm{d}y\right| &\le \left|\int_{|y|>\eta_1} \chi(y)f'(x)y(\partial_z H)(x+\mathrm{i}y)\,\mathrm{d}x\,\mathrm{d}y\right| \\ &\quad + \left|\int_{\Omega_{\mathfrak{a}}\cap\{|y|\le\eta_1\}} y\chi(y)f''(x)H\,\mathrm{d}x\,\mathrm{d}y\right| \\ &\le CK\log(N)\big(\|f'\|_1(\eta_1)^{1-s} + \|f''\|_1(\eta_1)^{2-s}\big) \\ &\le C'K\log(N)\big((\eta_1)^{1-s} + \|f''\|_1(\eta_1)^{2-s}\big),\end{aligned} \tag{4.19}$$

where we used the assumption that $\|f'\|_1 \le C$. The result follows by taking $\eta_1 = \|f''\|_1^{-1}$. □

REMARK. Note the integration by parts technique in the above proof; before the partial integration, the integral over x contributes $\|f''\|_1$. Integrating by parts contributes $\mathcal{O}(|y|^{-1})$ from the derivative of $H(z)$, which is smaller when $|y| > \|f''\|_1^{-1}$. The integral in the variable y is always estimated by power counting. This argument of integration by parts and power counting will be used repeatedly in what follows.

Let now

$$\begin{aligned}e_{\mathfrak{a}}(\lambda) = \exp\Bigg[\frac{\mathrm{i}\lambda}{\pi}\int_{\Omega_{\mathfrak{a}}} &\big(\mathrm{i}y\chi(y)f''(x) + \mathrm{i}\big(f(x)+\mathrm{i}f'(x)y\big)\chi'(y)\big) \\ &\times N\big(m_N(z) - \mathbb{E}[m_N(z)]\big)\,\mathrm{d}x\,\mathrm{d}y\Bigg],\end{aligned} \tag{4.20}$$

so that by (4.12),

$$\left|e(\lambda) - e_{\mathfrak{a}}(\lambda)\right| \le C|\lambda|N^{2\mathfrak{a}-1}\|f''\|_1. \tag{4.21}$$

Let $E_{\mathfrak{a}}(z)$ be the same as $E(z)$ but with $e_{\mathfrak{a}}(\lambda)$ in place of $e(\lambda)$. Applying Lemma 4.4 with $H = E - E_{\mathfrak{a}} = \mathcal{O}(|\lambda|N^{\varepsilon+2\mathfrak{a}-1}\|f''\|_1\eta^{-1})$, we see that

$$\begin{aligned}\psi'(\lambda) = \frac{\mathrm{i}}{\pi}\int_{\Omega_{\mathfrak{a}}} &\big(iy\chi(y)f''(x) + \mathrm{i}\big(f(x)+\mathrm{i}yf'(x)\big)\chi'(y)\big)E_{\mathfrak{a}}(z)\,\mathrm{d}x\,\mathrm{d}y \\ &+ \mathcal{O}\big(N^{\varepsilon+2\mathfrak{a}-1}\|f''\|_1\big)\end{aligned} \tag{4.22}$$

for any $\varepsilon > 0$.

We will use a cumulant expansion on the right-hand side of (4.22). Let ∂_{ab} denote differentiation with respect to H_{ab}. By a direct calculation (similar to (4.36) below) using Lemma 4.4 one can see that with overwhelming probability,

$$\left|\partial_{ab}^k e_{\mathfrak{a}}(\lambda)\right| \le \big(1+|\lambda|\big)^k N^{\varepsilon} \tag{4.23}$$

for any $\varepsilon > 0$ and N large enough.

LEMMA 4.5. *For $z \in \Omega_{\mathfrak{a}}$ we have*

$$
\begin{aligned}
(z+2m(z))E_{\mathfrak{a}}(z) &= \sum_{ia} \frac{1+a_2\delta_{ia}}{N}\mathbb{E}[\partial_{ia}e_{\mathfrak{a}}(\lambda)(G_{ia})] - \sum_{a\neq i}\frac{s_4}{2N^2}\mathbb{E}[\partial_{ai}^2 e_{\mathfrak{a}}(\lambda)m^2] \\
&\quad + \mathcal{O}\big(N^{\varepsilon}(1+|\lambda|^4) \\
&\quad \times ((N\eta)^{-1/2}(1+\|f''\|_1) + (N\eta^2)^{-1} + N^{-1/2}\eta^{-1})\big).
\end{aligned} \tag{4.24}
$$

PROOF. By the cumulant expansion (see, e.g., Lemma 3.2 of [26] and the discussion therein for use of the cumulant expansion), we derive

$$
\begin{aligned}
z\mathbb{E}[e_{\mathfrak{a}}(\lambda)(G_{ii}-\mathbb{E}[G_{ii}])] &= \sum_a \mathbb{E}[e_{\mathfrak{a}}(\lambda)(G_{ia}H_{ai} - \mathbb{E}[G_{ia}H_{ai}])] \\
&= \sum_a \frac{1+\delta_{ia}a_2}{N}\mathbb{E}[\partial_{ia}e_{\mathfrak{a}}(\lambda)G_{ia}] && (4.25)\\
&\quad - \sum_a \frac{1+\delta_{ia}a_2}{N}\mathbb{E}[e_{\mathfrak{a}}(\lambda)(G_{ii}G_{aa} - \mathbb{E}[G_{ii}G_{aa}])] && (4.26)\\
&\quad - \sum_a \frac{1+\delta_{ia}}{N}\mathbb{E}[e_{\mathfrak{a}}(\lambda)(G_{ia}^2 - \mathbb{E}[G_{ia}^2])] && (4.27)\\
&\quad + \Big\{\sum_a \frac{s_3+\delta_{ia}a_3}{2N^{3/2}}\Big[\mathbb{E}[\partial_{ai}^2 e_{\mathfrak{a}}(\lambda)G_{ai}] + 2\mathbb{E}[\partial_{ai}e_{\mathfrak{a}}(\lambda)\partial_{ai}G_{ai}] \\
&\quad + \mathbb{E}[e_{\mathfrak{a}}(\lambda)(\partial_{ai}^2 G_{ai} - \mathbb{E}[\partial_{ai}^2 G_{ai}])]\Big]\Big\} && (4.28)\\
&\quad + \Big\{\sum_a \frac{s_4+\delta_{ia}a_4}{6N^2}\Big[\mathbb{E}[\partial_{ai}^3 e_{\mathfrak{a}}(\lambda)G_{ai}] + \cdots \\
&\quad + \mathbb{E}[e_{\mathfrak{a}}(\lambda)(\partial_{ai}^3 G_{ai} - \mathbb{E}[\partial_{ai}^3 G_{ai}])]\Big]\Big\} && (4.29)\\
&\quad + \mathcal{O}(N^{\varepsilon-3/2}(1+|\lambda|^4)), && (4.30)
\end{aligned}
$$

any $\varepsilon > 0$. We begin calculating each term. Starting with (4.26) we have

$$
\begin{aligned}
\sum_{i,a}\frac{1+\delta_{ia}a_2}{N}\mathbb{E}[e_{\mathfrak{a}}(\lambda)(G_{ii}G_{aa} - \mathbb{E}[G_{ii}G_{aa}])] &= 2\sum_i m\mathbb{E}[e_{\mathfrak{a}}(\lambda)(G_{ii} - \mathbb{E}[G_{ii}])] \\
&\quad + \mathcal{O}(N^{\varepsilon}((N\eta)^{-1/2} + (N\eta^2)^{-1})).
\end{aligned} \tag{4.31}
$$

This term contributes the $2mE_{\mathfrak{a}}$ to the LHS of (4.24). For (4.27) we have

$$
\begin{aligned}
\sum_{i,a}\frac{1-\delta_{ia}}{N}\mathbb{E}[e_{\mathfrak{a}}(\lambda)(G_{ia}^2 - \mathbb{E}[G_{ia}^2])] &= \mathbb{E}[e_{\mathfrak{a}}(\lambda)\partial_z(m_N - \mathbb{E}[m_N])] + \mathcal{O}(N^{\varepsilon}(N\eta)^{-1/2}) \\
&= \mathcal{O}(N^{\varepsilon}((N\eta)^{-1/2} + (N\eta^2)^{-1})).
\end{aligned} \tag{4.32}
$$

We now start with the terms in (4.28). For a combinatorial factor K,

$$
\begin{aligned}
\sum_{i,a}\frac{s_3+\delta_{ia}a_3}{N^{3/2}}\mathbb{E}[e(\lambda)(\partial_{ai}^2 G_{ai} - \mathbb{E}[\partial_{ai}^2 G_{ai}])] &= Km^2\frac{1}{N^{3/2}}\sum_{ia}\mathbb{E}[e(\lambda)(G_{ia} - \mathbb{E}[G_{ia}])] \\
&\quad + \mathcal{O}(N^{\varepsilon}(N^{-1/2} + N^{-1/2}\eta^{-1})).
\end{aligned} \tag{4.33}
$$

By the isotropic local law (Theorem 2.12 of [4]),

$$\left|\frac{1}{N^{1/2}}\sum_i G_{ia}\right| \le \frac{N^\varepsilon}{\sqrt{N\eta}} \tag{4.34}$$

with overwhelming probability, and so

$$\frac{1}{N^{3/2}}\sum_{ia}\mathbb{E}\big[e(\lambda)\big(G_{ia}-\mathbb{E}[G_{ia}]\big)\big] = \mathcal{O}\big(N^\varepsilon(N\eta)^{-1/2}\big). \tag{4.35}$$

In preparation for the second term of (4.28) we calculate, for $i \neq a$,

$$\partial_{ia}e_{\mathfrak{a}}(\lambda) = -\frac{2\mathrm{i}\lambda}{\pi}e_{\mathfrak{a}}(\lambda)\int_{\Omega_{\mathfrak{a}}}\bar\partial_{z'}\tilde f(z')\sum_b G_{bi}G_{ab} = -\frac{2\mathrm{i}\lambda}{\pi}e_{\mathfrak{a}}(\lambda)\int_{\Omega_{\mathfrak{a}}}\bar\partial_{z'}\tilde f(z')\partial_{z'}G_{ia}. \tag{4.36}$$

Since with overwhelming probability,

$$\sum_b G_{bi}G_{ba} = \partial_{z'}G_{ai} = \mathcal{O}\big(N^\varepsilon(N\eta')^{-1/2}(\eta')^{-1}\big), \tag{4.37}$$

we find that (using Lemma 4.4)

$$\big|\partial_{ia}e_{\mathfrak{a}}(\lambda)\big| \le N^\varepsilon(1+|\lambda|)\big(N^{-1/2}(1+\|f''\|_1)^{1/2}\big). \tag{4.38}$$

For the second term of (4.28) we have

$$\begin{aligned}\sum_{ia}\frac{s_3+\delta_{ia}a_3}{N^{3/2}}\mathbb{E}\big[\partial_{ai}e_{\mathfrak{a}}(\lambda)\partial_{ai}G_{ai}\big] &= -\frac{s_3}{N^{3/2}}\sum_{i\neq a}\mathbb{E}\big[\partial_{ai}e_{\mathfrak{a}}(\lambda)G_{aa}G_{ii}\big]\\ &\quad+\mathcal{O}\big(N^\varepsilon(1+|\lambda|)\big(N^{-1/2}+N^{-1/2}\eta^{-1}\big)\big).\end{aligned} \tag{4.39}$$

We see, using (4.38),

$$\begin{aligned}\frac{s_3}{N^{3/2}}\sum_{i\neq a}\mathbb{E}\big[\partial_{ai}e_{\mathfrak{a}}(\lambda)G_{aa}G_{ii}\big] &= \frac{s_3m^2}{N^{3/2}}\sum_{i\neq a}\mathbb{E}\big[\partial_{ai}e_{\mathfrak{a}}(\lambda)\big]\\ &\quad+\mathcal{O}\big(N^\varepsilon\eta^{-1/2}N^{-1/2}(1+\|f'\|_1^{1/2})(1+|\lambda|)\big).\end{aligned} \tag{4.40}$$

The derivative in the above expression together with the sum over i,a gives us $N^{-3/2}\times\partial_{z'}\sum_{i,a}G_{ia} = N^{-1/2}\partial_{z'}m + \mathcal{O}(N^{-3/2}(\eta')^{-2})$ by the isotropic local law. So, using Lemma 4.4,

$$\frac{1}{N^{3/2}}\sum_{i\neq a}\mathbb{E}\big[\partial_{ai}e_{\mathfrak{a}}(\lambda)\big] = \mathcal{O}\big(N^\varepsilon(1+|\lambda|)N^{-1/2}\big). \tag{4.41}$$

For the first term of (4.28) we further calculate, for $i \neq a$,

$$\begin{aligned}\partial_{ia}^2e_{\mathfrak{a}}(\lambda) &= e_{\mathfrak{a}}(\lambda)(i\lambda)^2\pi^{-2}\bigg(\int_{\Omega_{\mathfrak{a}}}\bar\partial_{z'}\tilde f(z')\sum_b G_{bi}G_{ab}\,\mathrm{d}z'\,\mathrm{d}\bar z'\bigg)^2\\ &\quad+2e_{\mathfrak{a}}(\lambda)\frac{\mathrm{i}\lambda}{\pi}\int_{\Omega_{\mathfrak{a}}}\bar\partial_{z'}\tilde f(z')\partial_{z'}(G_{aa}G_{ii}+G_{ia}^2)\\ &= e_{\mathfrak{a}}(\lambda)\frac{\mathrm{i}\lambda}{\pi}\int_{\Omega_{\mathfrak{a}}}\bar\partial_{z'}\tilde f(z')4m(z')m'(z')\,\mathrm{d}z'\,\mathrm{d}\bar z'\\ &\quad+\mathcal{O}\big(N^\varepsilon(1+|\lambda|)^2N^{-1/2}(\|f''\|_1^{1/2}+1)\big)\end{aligned} \tag{4.42}$$

and so

$$\begin{aligned}
&\sum_{a,i} \frac{s_3+\delta_{ia}a_3}{N^{3/2}}\mathbb{E}\big[\partial_{ai}^2 e_{\mathfrak{a}}(\lambda)G_{ai}\big] \\
&\quad = \sum_{a,i} \frac{s_3}{N^{3/2}}\mathbb{E}\Big[e_{\mathfrak{a}}(\lambda)\mathrm{i}\lambda\pi^{-1}\int_{\Omega_{\mathfrak{a}}} \bar\partial_{z'}\tilde f(z')4m(z')m'(z')G_{ai}\,\mathrm{d}z'\,\mathrm{d}\bar z'\Big] \\
&\qquad + \mathcal{O}\big(N^{\varepsilon}(1+|\lambda|)^2(N\eta)^{-1/2}(\|f''\|_1^{1/2}+1)\big) \\
&\quad = \mathcal{O}\big(N^{\varepsilon}(1+|\lambda|)^2(N\eta)^{-1/2}(\|f''\|_1^{1/2}+1)\big)
\end{aligned} \tag{4.43}$$

where in the last line we used the isotropic law again.

Finally, for the terms (4.29) we see that (using (4.23) and the Green's function estimates of Theorem 2.2 repeatedly)

$$\begin{aligned}
&\sum_{a,i} \frac{s_4+\delta_{ia}a_4}{N^2}\mathbb{E}\big[(\partial_{ai}^3 e_{\mathfrak{a}})G_{ai}\big] + \cdots + \mathbb{E}\big[e_{\mathfrak{a}}(\lambda)\big(\partial_{ai}^3 G_{ai} - \mathbb{E}[\partial_{ai}^3 G_{ai}]\big)\big] \\
&\quad = -\sum_{a\neq i} \frac{3s_4}{N^2}\mathbb{E}\big[\partial_{ai}^2 e_{\mathfrak{a}}(\lambda)m^2\big] + \mathcal{O}\big(N^{\varepsilon}(1+|\lambda|^3)(N\eta)^{-1/2}\big).
\end{aligned} \tag{4.44}$$

□

We further calculate the terms on the right side of (4.24).

LEMMA 4.6. *We have*

$$\begin{aligned}
&\sum_{ia} \frac{1+\delta_{ia}}{N}\mathbb{E}\big[(\partial_{ia}e_{\mathfrak{a}}(\lambda))(G_{ia})\big] \\
&\quad = -\frac{2\mathrm{i}\lambda}{\pi}\int_{\Omega_{\mathfrak{a}}} \bar\partial_{z'}\tilde f(z')\partial_{z'}\frac{1}{N}\mathbb{E}\big[e(\lambda)\operatorname{tr}G(z')G(z)\big]\,\mathrm{d}z'\,\mathrm{d}\bar z'
\end{aligned} \tag{4.45}$$

and

$$\begin{aligned}
\frac{1}{N}\sum_i \mathbb{E}\big[\partial_{ii}e_{\mathfrak{a}}(\lambda)G_{ii}\big] &= -\frac{\mathrm{i}\lambda\mathbb{E}[e_{\mathfrak{a}}(\lambda)]}{\pi}\int_{\Omega_{\mathfrak{a}}} \bar\partial_{z'}\tilde f(z')m'(z')m(z)\,\mathrm{d}z'\,\mathrm{d}\bar z' \\
&\quad + \mathcal{O}\big(N^{\varepsilon}(1+|\lambda|)(N\eta)^{-1/2}\big)
\end{aligned} \tag{4.46}$$

and

$$\begin{aligned}
&\sum_{a\neq i} \frac{s_4}{2N^2}\mathbb{E}\big[\partial_{ai}^2 e_{\mathfrak{a}}(\lambda)m(z)^2\big] \\
&\quad = 2s_4\mathbb{E}\big[e_{\mathfrak{a}}(\lambda)\big]\mathrm{i}\lambda\int_{\Omega_{\mathfrak{a}}} \bar\partial_{z'}\tilde f(z')m(z')m'(z')m(z)^2\,\mathrm{d}z'\,\mathrm{d}\bar z' \\
&\qquad + \mathcal{O}\big(N^{\varepsilon}(1+|\lambda|)^2N^{-1/2}(\|f''\|_1^{1/2}+1)\big).
\end{aligned} \tag{4.47}$$

PROOF. We see that,

$$\begin{aligned}
&\sum_{ia} \frac{1+\delta_{ia}}{N}\mathbb{E}\big[(\partial_{ia}e(\lambda))(G_{ia})\big] \\
&\quad = -\frac{2\mathrm{i}\lambda}{\pi N}\int \bar\partial_{z'}\tilde f(z')\mathbb{E}\sum_{iab} e(\lambda)G_{ab}(z')G_{bi}(z')G_{ia}(z)\,\mathrm{d}z'\,\mathrm{d}\bar z' \\
&\quad = -\frac{2\mathrm{i}\lambda}{\pi}\int \bar\partial_{z'}\tilde f(z')\partial_{z'}\frac{1}{N}\mathbb{E}\big[e(\lambda)\operatorname{tr}G(z')G(z)\big]\,\mathrm{d}z'\,\mathrm{d}\bar z'.
\end{aligned} \tag{4.48}$$

The second estimate is similar, and the last estimate follows from (4.42). □

Putting together the last two lemmas we obtain,

$$
\begin{aligned}
E_{\mathfrak{a}}(z) = {} & \frac{1}{z+2m(z)}\frac{\mathrm{i}\lambda}{\pi}\int_{\Omega_{\mathfrak{a}}} \mathrm{d}z'\,\mathrm{d}\bar{z}'\bar{\partial}_{z'}\tilde{f}(z') \\
& \times \Big\{\partial_{z'}\frac{-2}{N}\mathbb{E}\big[e_{\mathfrak{a}}(\lambda)\operatorname{tr}G(z')G(z)\big] \\
& - \mathbb{E}\big[e_{\mathfrak{a}}(\lambda)\big]\big(2s_4 m(z)^2 m(z')\partial_{z'}m(z') - (a_2-1)m'(z')m(z)\big)\Big\} \\
& + \mathcal{O}\big(N^{\varepsilon}(1+|\lambda|^4) \\
& \times \big((N\eta)^{-1/2}(1+\|f''\|_{L^1(\mathbb{R})}^{1/2}) + (N\eta^2)^{-1} + N^{-1/2}\eta^{-1}\big)\big).
\end{aligned}
\tag{4.49}
$$

We now complete the proof of Proposition 4.1. We see that with overwhelming probability, if z and z' are on different half-planes,

$$
\begin{aligned}
\partial_{z'}\frac{1}{N}\operatorname{tr}G(z')G(z) &= \partial_{z'}\frac{1}{N}\operatorname{tr}\frac{G(z')-G(z)}{z'-z} \\
&= \partial_{z'}\frac{m(z')-m(z)}{z'-z} \\
&\quad + \mathcal{O}\big(N^{\varepsilon}(\eta')^{-1}\big((N\eta')^{-1}+(N\eta)^{-1}\big)(\eta+\eta')^{-1}\big)
\end{aligned}
\tag{4.50}
$$

whereas if they are on the same half-plane, divide into the cases $\eta < 2\eta'$ and $\eta > 2\eta'$. In the former, write $m(z) - m(z') = \int m'(s)\,\mathrm{d}s$ to get

$$
\begin{aligned}
\partial_{z'}\frac{1}{N}\operatorname{tr}G(z')G(z) &= \partial_{z'}\frac{m(z')-m(z)}{z'-z} + \mathcal{O}\big(N^{\varepsilon}(\eta')^{-1}\big((N\eta^2)^{-1}+(N(\eta')^2)^{-1}\big)\big) \\
&= \partial_{z'}\frac{m(z')-m(z)}{z'-z} + \mathcal{O}\big(N^{\varepsilon}(\eta')^{-1}(N\eta^2)^{-1}\big).
\end{aligned}
\tag{4.51}
$$

Whereas if $\eta > 2\eta'$ directly estimate $m_N(z') - m_N(z) = m(z') - m(z) + \mathcal{O}(N^{\varepsilon}(N\eta')^{-1}))$ to get,

$$
\partial_{z'}\frac{1}{N}\operatorname{tr}G(z')G(z) = \partial_{z'}\frac{m(z')-m(z)}{z'-z} + \mathcal{O}\big(N^{\varepsilon}\big((N\eta^2\eta')^{-1}+(N(\eta')^2\eta)^{-1}\big)\big).
\tag{4.52}
$$

Integrating all of this (i.e., using the integration by parts and power counting technique of Lemma 4.4) we see that for any $\varepsilon > 0$,

$$
\frac{\mathrm{d}}{\mathrm{d}\lambda}\psi(\lambda)
\tag{4.53}
$$

$$
\begin{aligned}
&= \lambda\mathbb{E}\big[e_{\mathfrak{a}}(\lambda)\big]\Big[\frac{1}{\pi^2}\int_{\Omega_{\mathfrak{a}}^2}\bar{\partial}_z\tilde{f}(z)\bar{\partial}_{z'}\tilde{f}(z')\frac{1}{z+2m(z)} \\
&\quad \times \Big(2\partial_{z'}\frac{m(z')-m(z)}{z'-z} \\
&\quad + 2s_4 m^2(z)m(z')m'(z') + (a_2-1)m'(z')m(z)\Big)\,\mathrm{d}z\,\mathrm{d}\bar{z}\,\mathrm{d}z'\,\mathrm{d}\bar{z}'\Big]
\end{aligned}
\tag{4.54}
$$

$$
+ \mathcal{O}\big(N^{\varepsilon}(1+|\lambda|^4)\big[N^{-1/2}\|f''\|_1 + N^{2\mathfrak{a}-1}\|f''\|_1\big]\big).
\tag{4.55}
$$

In order to complete the proof, we need only estimate the error in first replacing $\mathbb{E}[e_{\mathfrak{a}}(\lambda)]$ by $\mathbb{E}[e(\lambda)]$ and then restoring to the integral the region $\mathbb{C}^2 \setminus \Omega_{\mathfrak{a}}^2$. First we estimate the size of the integral appearing above, and show that it is at most $\mathcal{O}(\log(N)^2)$. We can then replace

$\mathbb{E}[e_{\mathfrak{a}}]$ by $\mathbb{E}[e(\lambda)] = \psi(\lambda)$ using (4.21). The terms which need to be estimated involve $f''(x)$; the contribution from the other terms are bounded, due to the assumptions (4.1) and the appearance of $\chi'(y)$ in these terms which is nonzero only for y of order 1. We turn to the terms with $f''(x)$. The second two terms in the line (4.54) are bounded functions, so we can integrate by parts in x to estimate this contribution by $\mathcal{O}(1)$.

When z and z' are in the same half-spaces, $(m(z') - m(z))/(z' - z)$ is bounded, and so the derivative is bounded by $C/\operatorname{Im}[z']$. Using then (the proof of) Lemma 4.4 we can estimate this contribution by $\mathcal{O}(\log(N)$. When z and z' are in separate half-spaces we use instead the estimate

$$\left|\partial_{z'} \frac{m(z) - m(z')}{z - z'}\right| \le \frac{C}{(\operatorname{Im}[z'])^2 + (\operatorname{Im}[z])^2}. \tag{4.56}$$

The largest contribution is from when both $\operatorname{Im}[z']$ and $\operatorname{Im}[z] > \|f''\|_1^{-1}$, where we integrate by parts in both $\operatorname{Re}[z]$ and $\operatorname{Re}[z']$ and find a contribution of order $\mathcal{O}(\log(N)^2)$. The other regions are at most $\mathcal{O}(\log(N))$.

Finally, we argue that we can restore to the integral the region $\mathbb{C}^2 \setminus \Omega_{\mathfrak{a}}^2$ at an error of at most $\mathcal{O}(\log(N) N^{2\mathfrak{a}-1} \|f''\|_1)$. The region $(\mathbb{C} \setminus \Omega_{\mathfrak{a}})^2$ is easy to control as the second line of (4.54) is bounded by $C/(\operatorname{Im}[z]\operatorname{Im}[z'])$, and gives a contribution of $\mathcal{O}(N^{2\mathfrak{a}-2}\|f''\|_1^2)$. The "cross terms" $\Omega_{\mathfrak{a}} \times (\mathbb{C} \setminus \Omega_{\mathfrak{a}})$ (and vice versa) must be handled by integration by parts again in the region where $\operatorname{Im}[z]$ or $\operatorname{Im}[z'] > \|f''\|_1^{-1}$ as above, and the cases of same and different half-spaces are treated similarly, for example, using the estimate (4.56). The cross terms are found to contribute $\mathcal{O}(N^{\mathfrak{a}-1}\|f''\|_1 \log(N))$. The claimed estimate of Proposition 4.1 follows after taking $\mathfrak{a} = \varepsilon$. $\square$

4.3. *Proof of Lemma 4.2.* We apply Green's theorem to each term, which states

$$\int_\Omega \bar\partial_z F(z)\,\mathrm{d}z\,\mathrm{d}\bar z = \frac{-\mathrm{i}}{2}\int_{\partial\Omega} F(z)\,\mathrm{d}z. \tag{4.57}$$

We write $z' = y \pm \mathrm{i}0$ and $z = x \pm \mathrm{i}0$. First, we integrate by parts to find

$$\begin{aligned} &\frac{-1}{\pi^2}\int_{\mathbb{C}^2} \bar\partial_z \tilde f(z) \bar\partial_{z'} \tilde f(z') \frac{1}{z + 2m(z)}\left(2\partial_{z'}\frac{m(z') - m(z)}{z' - z}\right)\mathrm{d}z\,\mathrm{d}\bar z\,\mathrm{d}z'\,\mathrm{d}\bar z' \\ &\quad = \frac{1}{\pi^2}\int_{\mathbb{C}^2} \bar\partial_z \tilde f(z) \bar\partial_{z'} \partial_{\operatorname{Re}[z']} \tilde f(z') \frac{1}{z + 2m(z)}\left(2\frac{m(z') - m(z)}{z' - z}\right)\mathrm{d}z\,\mathrm{d}\bar z\,\mathrm{d}z'\,\mathrm{d}\bar z'. \end{aligned} \tag{4.58}$$

Note that $x \pm i0 + 2m(x \pm \mathrm{i}0) = \pm\mathrm{i}\sqrt{4 - x^2}\mathbf{1}_{\{|x|\le 2\}}$. Hence,

$$\begin{aligned} &\frac{2}{x + 2m(x \pm \mathrm{i}0)}\left(\frac{m(y + \mathrm{i}0) - m(x \pm \mathrm{i}0)}{y - x} - \frac{m(y - \mathrm{i}0) - m(x \pm \mathrm{i}0)}{y - x}\right) \\ &\quad = \pm 2\frac{\sqrt{4 - y^2}}{\sqrt{4 - x^2}(y - x)}. \end{aligned} \tag{4.59}$$

Hence, we find (note that some of the integrals over the real axes require intepretation as principal values)

$$\begin{aligned} &\frac{1}{\pi^2}\int_{\mathbb{C}^2} \bar\partial_z \tilde f(z) \bar\partial_{z'} \partial_{\operatorname{Re}[z']} \tilde f(z') \frac{1}{z + 2m(z)}\left(2\frac{m(z') - m(z)}{z' - z}\right)\mathrm{d}z\,\mathrm{d}\bar z\,\mathrm{d}z'\,\mathrm{d}\bar z' \\ &\quad = \frac{-1}{\pi^2}\int_{-2}^{2}\int_{-2}^{2} f(x) f'(y) \frac{\sqrt{4 - y^2}}{\sqrt{4 - x^2}(y - x)}\,\mathrm{d}x\,\mathrm{d}y \end{aligned} \tag{4.60}$$

$$= \frac{1}{\pi^2}\int_{-2}^{2} f'(y)\sqrt{4-y^2}\int_{-2}^{2}\frac{f(y)-f(x)}{\sqrt{4-x^2}(y-x)}\,\mathrm{d}x\,\mathrm{d}y$$
$$= \frac{-1}{2\pi^2}\int_{-2}^{2}\int_{-2}^{2}\frac{(f(y)-f(x))^2}{(y-x)^2}\frac{4-xy}{\sqrt{4-x^2}\sqrt{4-y^2}}\,\mathrm{d}x\,\mathrm{d}y,$$

where the last line follows by integration by parts in y and the second last line from the identity

$$\int_{-2}^{2}\frac{1}{(x-y)\sqrt{4-x^2}}\,\mathrm{d}x = 0. \tag{4.61}$$

For the other terms, note that $(z+2m)m = m^2-1$ and since $m' = m^2/(1-m^2)$,

$$\frac{m}{z+2m} = -m'. \tag{4.62}$$

Therefore,

$$\begin{aligned}
&\frac{-1}{\pi^2}\int_{\mathbb{C}^2}\bar\partial_z\tilde f(z)\bar\partial_{z'}\tilde f(z')\frac{m'(z')m(z)}{z+2m(z)}\,\mathrm{d}z\,\mathrm{d}\bar z\,\mathrm{d}z'\,\mathrm{d}\bar z' \\
&\quad= \frac{1}{\pi^2}\int_{\mathbb{C}^2}\bar\partial_z\tilde f(z)\bar\partial_{z'}\tilde f(z')m'(z')m'(z)\,\mathrm{d}z\,\mathrm{d}\bar z\,\mathrm{d}z'\,\mathrm{d}\bar z' \\
&\quad= \frac{1}{4\pi^2}\biggl(\int_{-2}^{2} f(x)\frac{x}{\sqrt{4-x^2}}\,\mathrm{d}x\biggr)^2.
\end{aligned} \tag{4.63}$$

For the final term we have

$$\begin{aligned}
&\frac{-1}{\pi^2}\int_{\mathbb{C}^2}\bar\partial_z\tilde f(z)\bar\partial_{z'}\tilde f(z')\frac{m^2(z)m'(z)m(z)}{z+2m(z)}\,\mathrm{d}z\,\mathrm{d}\bar z\,\mathrm{d}z'\,\mathrm{d}\bar z' \\
&\quad= \frac{1}{\pi^2}\int_{\mathbb{C}^2}\bar\partial_z\tilde f(z)\bar\partial_{z'}\tilde f(z')m'(z')m(z')m'(z)m(z)\,\mathrm{d}z\,\mathrm{d}\bar z\,\mathrm{d}z'\,\mathrm{d}\bar z' \\
&\quad= \frac{1}{4\pi^2}\biggl(\int_{-2}^{2} f(x)\frac{2-x^2}{\sqrt{4-x^2}}\,\mathrm{d}x\biggr)^2.
\end{aligned} \tag{4.64}$$

This completes the calculation of $V(f)$. To calculate the expectation, we use the cumulant expanson on $\mathbb{E}[G_{ii}]$. Let $\Omega_{\mathfrak{a}}$ and $\mathfrak{a} > 0$ be as earlier. For $z \in \Omega_{\mathfrak{a}}$ we find for any $\varepsilon > 0$,

$$\begin{aligned}
1 + z\mathbb{E}[G_{ii}] &= -\frac{1}{N}\sum_{a\neq i}\mathbb{E}\bigl[G_{ii}G_{aa} + G_{ia}^2\bigr] - \frac{1+a_2}{N}\mathbb{E}\bigl[m^2\bigr] \\
&\quad + \frac{s_3 m^2}{2N^{3/2}}\sum_{a\neq i}\mathbb{E}[G_{ia}] - \frac{s_4}{N}m^4 \\
&\quad + \mathcal{O}\bigl(N^{\varepsilon}N^{-1}(N\eta^2)^{-1/2}\bigr).
\end{aligned} \tag{4.65}$$

The term with s_3 can be absorbed into the error term using the isotropic local law. We therefore find,

$$\begin{aligned}
\sum_i (z+2m)\mathbb{E}[G_{ii} - m] &= -\partial_z m - (a_2-1)m^2 - s_4 m^4 \\
&\quad + \mathcal{O}\bigl(N^{\varepsilon}\bigl((N\eta^2)^{-1} + (N\eta^2)^{-1/2}\bigr)\bigr).
\end{aligned} \tag{4.66}$$

Integrating, we find

$$\mathbb{E}[f(H)] - N\int f(x)\,\mathrm{d}\rho_{\mathrm{sc}}(x)\,\mathrm{d}x = \frac{1}{\pi}\int \bar{\partial}_z \tilde{f}(z)\frac{(1-a_2)m^2 - s_4 m^4 - \partial_z m}{z+2m}\,\mathrm{d}z\,\mathrm{d}\bar{z}' + \mathcal{O}\big(N^{\varepsilon}N^{-1/2}(\|f''\|_1 + 1) + N^{\mathfrak{a}+\varepsilon-1}\|f''\|_1\big). \tag{4.67}$$

Note that we had to remove and then return the domain $\mathbb{C}\setminus\Omega_{\mathfrak{a}}$ as in the proof of Proposition 4.1; this is similar to what was done above. We take $\mathfrak{a} = \varepsilon$. We calculate using Green's theorem as above,

$$\begin{aligned}\frac{1}{\pi}\int_{\mathbb{C}} \bar{\partial}_z \tilde{f}(z)\frac{m^2}{z+2m}\,\mathrm{d}z\,\mathrm{d}\bar{z}' &= -\frac{1}{\pi}\int_{\mathbb{C}} \bar{\partial}_z \tilde{f}(z) m'(z) m(z)\,\mathrm{d}z\,\mathrm{d}\bar{z}' \\ &= \frac{1}{2\pi}\int_{-2}^{2} f(x)\frac{2-x^2}{\sqrt{4-x^2}}\,\mathrm{d}x,\end{aligned} \tag{4.68}$$

and

$$\begin{aligned}\frac{-1}{\pi}\int_{\mathbb{C}} \bar{\partial}_z \tilde{f}(z)\frac{m^4}{z+2m}\,\mathrm{d}z\,\mathrm{d}\bar{z}' &= \frac{1}{\pi}\int_{\mathbb{C}} \bar{\partial}_z \tilde{f}(z) m^3(z) m'(z)\,\mathrm{d}z\,\mathrm{d}\bar{z}' \\ &= \frac{1}{2\pi}\int_{-2}^{2} f(x)\frac{x^4-4x^2+2}{\sqrt{4-x^2}}\,\mathrm{d}x.\end{aligned} \tag{4.69}$$

Finally,

$$\frac{-1}{\pi}\int_{\mathbb{C}} \bar{\partial}_z \tilde{f}(z)\frac{m'(z)}{z+2m}\,\mathrm{d}z\,\mathrm{d}\bar{z}' = \frac{-1}{2\pi}\int_{-2}^{2}\frac{f(x)}{\sqrt{4-x^2}}\,\mathrm{d}x + \frac{f(2)+f(-2)}{4}. \tag{4.70}$$

This completes the proof. □

5. Derivation of the main results.

5.1. *Reduction to Gaussian divisible ensembles.* We will only provide full details for the proofs of Theorems 1.1 and 1.4 for those Wigner matrices with Gaussian components (see Definition 3.1). The reduction from general Wigner matrices to those with a Gaussian component is well known in the random matrix literature, and so we will be brief in our discussion. In our setting, this reduction will be a consequence of the four moment method of Tao and Vu [34]. For the specific case of single eigenvalue fluctuations, similar arguments were made in [34] and [29]. First, we have

PROPOSITION 5.1. *Let H and W be two Wigner matrices so that*

$$\big|\mathbb{E}[H_{ij}^s] - \mathbb{E}[W_{ij}^s]\big| \le N^{-2-c}, \quad 1 \le s \le 4 \tag{5.1}$$

for some $c > 0$. Then if the results of Theorems 1.1 or 1.4 hold for W, they hold for H.

This is essentially a consequence of Theorem 15 of [34], specialized to the real symmetric setting. The only extension is that the original Theorem asks for equality of the first four moments rather then the fact that they are only approximately equal. This extension is well known; see, for example, [15].

The additional required input is that given a Wigner matrix, one can find a matching Gaussian divisible ensemble. The following is a consequence of, for example, Lemma 3.4 of [18].

LEMMA 5.2. *Let H be a Wigner matrix, and let $\tau_0 > 0$. Then there is a Wigner matrix W with Gaussian component of size $N^{-\tau_0}$ so that (5.1) holds with some $c > 0$.*

The consequence of these two results is that if there is a $\tau_0 > 0$ so that we can prove our main results for Wigner matrices with Gaussian components of size $N^{-\tau_0}$, then our results extend to all Wigner matrices. The τ_0 we take is the one so that Theorem 3.2 holds.

5.2. *Proof of Theorem* 1.1. Let F be a smooth function, which we moreover take to be of compact support. Let $\mathbf{x}(t)$ and $\mathbf{y}(t)$ be as in Section 3.1, so that the initial data $\mathbf{x}(0)$ are the eigenvalues of a Wigner matrix with Gaussian component of size $N^{-\tau_0}$. We will need to introduce a third process $\mathbf{z}(t)$ coupled to the same Brownian motions as $\mathbf{x}(t)$ and $\mathbf{y}(t)$, with initial data an independent GOE matrix. We let

$$Z_i(\mathbf{x}) = N\frac{x_i(H_t) - \gamma_i}{\sqrt{\frac{\log N}{(1-\gamma_i^2/4)}}},$$

and $Z_i(\mathbf{y})$ be the corresponding quantity for a GOE matrix. It suffices to show

$$\mathbb{E}\big[F\big(Z_i\big(\mathbf{x}(t)\big)\big)\big] = \mathbb{E}\big[F\big(Z_i\big(\mathbf{z}(t)\big)\big)\big] + o(1),$$

where the time parameter $t = t_1$ as in Theorem 3.2. By (3.9) we have

$$\mathbb{E}\big[F\big(Z_i\big(\mathbf{x}(t)\big)\big)\big] = \mathbb{E}\big[F\big(Z_i\big(\mathbf{y}(t)\big) + \alpha_N(\zeta_{\mathbf{x}} - \zeta_{\mathbf{y}})\big)\big] + \mathcal{O}(N^{-c_1}),$$

where

$$\alpha_N = \frac{2\pi}{\sqrt{\log N}} \times \sqrt{1 - \gamma_i^2/4}.$$

Denoting the characteristic function of $\alpha_N\zeta_x$ by $\psi_x(\lambda)$, we have

$$\mathbb{E}\big[F\big(Z_i\big(\mathbf{y}(t)\big) + \alpha_N(\zeta_x - \zeta_y)\big)\big] = \int \psi(\lambda)\widehat{F}(\lambda)\mathbb{E}\big[e^{i\lambda(Z_i(\mathbf{y}(t)) - \alpha_N\zeta_y)}\big]\,\mathrm{d}\lambda.$$

Repeating the same argument for $\mathbf{z}(t)$ we see that

$$\begin{aligned}\mathbb{E}\big[F\big(Z_i\big(\mathbf{x}(t)\big)\big)\big] &- \mathbb{E}\big[F\big(Z_i\big(\mathbf{z}(t)\big)\big)\big] \\ &= \int \big(\psi_x(\lambda) - \psi_z(\lambda)\big)\hat{F}(\lambda)\mathbb{E}\big[e^{i\lambda(Z_i(\mathbf{y}(t)) - \alpha_N\zeta_y)}\big]\,\mathrm{d}\lambda + o(1).\end{aligned} \tag{5.2}$$

Denote by $V_x(\Phi)$ and $V_z(\Phi)$ the functionals appearing in Lemma 4.2 for matrices X and Z. Then we have

$$\big|V_x(\Phi) - V_z(\Phi)\big| \le C. \tag{5.3}$$

From this and Proposition 4.1, we obtain that

$$\psi_x(\lambda) = e^{-\alpha\lambda^2} + o(1), \quad |\lambda| \le \log(N)^{1/4},$$

where α is independent of the choice of x or z. Since $\hat{F}$ is a Schwartz function, this yields the claim. □

5.3. *Proof of Corollary* 1.2. The argument leading from Theorem 1.1 to the Corollary 1.2 was detailed in [20], Theorem 1.1. It suffices to notice that

$$\mathbb{P}\bigg(\frac{\mathcal{N}(E) - N\int_{-2}^{E}\rho_{\mathrm{sc}}(u)\,\mathrm{d}u}{\sqrt{\frac{1}{2\pi^2}\log N}} \le x\bigg) = \mathbb{P}(\lambda_{i_N} \le E) = \mathbb{P}\bigg(N\frac{\lambda_{i_N} - \gamma_{i_N}}{\sqrt{\log N}} \le x_N\bigg),$$

where i_N is the integer part of

$$N\int_{-2}^{E}\rho_{\mathrm{sc}}(y)\,\mathrm{d}y - x\sqrt{\frac{1}{2\pi^2}\log N},$$

and x_N is defined by the equality. Note that,

$$\frac{i_N}{N} \to \int_{-2}^{E} \rho_{\mathrm{sc}}(y)\,\mathrm{d}y.$$
$$x_N \to x.$$

The result now follows directly from the convergence in distribution of λ_{i_N}. $\square$

5.4. *Proof of Theorem* 1.3. Recall the following result for mesoscopic linear statistics, [25], Theorem 6.18.

THEOREM 5.3. *Suppose f is a C^2 function with support in $[A, B]$, where* $\operatorname{supp}\rho_V = [-2, 2]$, *satisfying the assumptions in Section* 4.1. *Let λ_i, $1 \le i \le N$ be sampled from a β-ensemble with potential V. Then, there exists $\epsilon > 0$ such that*

$$\mathbb{E}\big[e^{i\lambda(\sum_{i=1}^{N} f(\lambda_i) - N\int f(x)\rho_V)}\big] = \exp\Big(-\frac{\lambda^2}{2}V(f) + i\delta(f)\Big) + \mathcal{O}(N^{-1+10\epsilon})\|f''\|_{L^1(\mathbb{R})} + \mathcal{O}(N^{-\epsilon}), \tag{5.4}$$

where

$$V(f) = \frac{1}{2\beta\pi^2}\int_A^B\int_A^B\Big(\frac{f(x)-f(y)}{x-y}\Big)^2 \frac{-AB - xy + \frac{1}{2}(A+B)(x+y)}{\sqrt{(x-A)(B-x)}\sqrt{(y-A)(B-y)}}\,\mathrm{d}x\,\mathrm{d}y,$$
$$\delta(f) = \frac{1}{2\pi^2}\Big(\frac{2}{\beta} - 1\Big)\int_A^B f(x)\,\mathrm{d}\nu_V(x)$$

where ν_V is a signed measure.

REMARK. The measure ν_V is characterized explicitly in the book of Pastur and Shcherbina [30] (see Theorem 11.3.2) in the case of analytic V. In general it can be realized as a boundary value of the Stieltjes transform of the equilbrium measure and its derivatives. In the case of sufficiently regular equilibrium measures one can show that it is a sum of delta functions at the spectral edges and a continuous density in the interior.

The proof of Theorem 1.3 is now nearly identical to the proof of Theorem 1.1, replacing Propostion 4.1 with Theorem 5.3. We apply the homogenization result Theorem 3.3 to the re-scaled β-ensembles,

$$\frac{\rho^{(V)}(\gamma_i^{(V)})}{\rho_{\mathrm{sc}}(0)}(x_j - \gamma_i^{(V)}), \qquad \frac{\rho^{(U)}(\gamma_i^{(U)})}{\rho_{\mathrm{sc}}(0)}(y_j - \gamma_i^{(U)}), \tag{5.5}$$

where x_i is the β-ensemble under consideration and y_j is a Gaussian β-ensemble.

The only thing that needs to be checked is that the variance of $V(\hat{\Phi})$ does not depend on the parameters A, B (up to $o(\log(N))$ terms). Using the fact that $\hat{\Phi}' \neq 0$ only near $\gamma_i^{(V)}$, a short calculation using the bounds (3.12) (which also hold for $\hat{\Phi}$), one sees that for any $\kappa > 0$,

$$V(\hat{\Phi}) = \frac{1}{2\beta\pi^2}\int_{|y-\gamma_i^{(V)}|\le\kappa}\int_{|x-\gamma_i^{(V)}|\le\kappa}\Big(\frac{\hat{\Phi}(x)-\hat{\Phi}(y)}{x-y}\Big)^2\,\mathrm{d}x\,\mathrm{d}y + \mathcal{O}(1). \tag{5.6}$$

The double integral does not depend on A and B, and so the proof proceeds as before.

5.5. *Proof of Theorem 1.4 and Theorem 1.5.* We note the following elementary definite integrals, which correspond to the terms on the right side of (4.8).

PROPOSITION 5.4. *Let $\gamma \in (-2, 2)$ and f be the indicator function*:

$$f(x) = \mathbf{1}_{\{x \le \gamma\}}.$$

Then,

$$\frac{1}{2\pi}\int_{-2}^{2} f(x)\frac{x^4 - 4x^2 + 2}{\sqrt{4 - x^2}}\,\mathrm{d}x = \frac{\sqrt{4-\gamma^2}}{8\pi}(2\gamma - \gamma^3), \tag{5.7}$$

$$\frac{1}{2\pi}\int_{-2}^{2} f(x)\frac{2 - x^2}{\sqrt{4 - x^2}}\,\mathrm{d}x = \frac{\sqrt{4-\gamma^2}}{4\pi}\gamma, \tag{5.8}$$

$$\frac{1}{2\pi}\int_{-2}^{2} f(x)\frac{x}{\sqrt{4 - x^2}}\,\mathrm{d}x = -\frac{\sqrt{4-\gamma^2}}{2\pi}. \tag{5.9}$$

In order to prove the expansion (1.11), we will use the following estimate, which is a consequence of gap universality [17]. This estimate states that there is a constant $\mathfrak{a} > 0$ so that for all indices i, j with $\kappa N \le i, j \le (1-\kappa)N$ where $\kappa > 0$,

$$|\mathbb{E}_W[\rho_{\mathrm{sc}}(\gamma_i)(\lambda_{i+1} - \lambda_i)] - \mathbb{E}_{\mathrm{GOE}}[\rho_{\mathrm{sc}}(\gamma_j)(\lambda_{j+1} - \lambda_j)] \le \frac{C}{N^{1+\mathfrak{a}}}, \tag{5.10}$$

where the first expectation is with respect to a Wigner matrix and the second is with respect to a GOE matrix. Taking now a fixed Wigner matrix with eigenvalues λ_i, we fix a small $\omega > 0$ and write

$$\lambda_{i_0} = \frac{1}{2N^\omega + 1}\sum_{|i - i_0| \le N^\omega} \lambda_i + \frac{1}{2N^\omega + 1}\sum_{|i - i_0| \le N^\omega} \lambda_{i_0} - \lambda_i. \tag{5.11}$$

We compare the expectation of $\lambda_{i_0} - \lambda_{i_0+k}$ to $\lambda_{i_0} - \lambda_{i_0-k}$. We can rewrite each of these quantities as a telescoping sum of gaps of consecutive eigenvalues and apply (5.10). As a consequence, using the smoothness of the semicircle distribution we see that by taking ω small enough, depending on $\mathfrak{a}$, that

$$\left|\mathbb{E}\frac{1}{2N^\omega + 1}\sum_{|i - i_0| \le N^\omega} \lambda_{i_0} - \lambda_i\right| \le \frac{C}{N^{1+\mathfrak{a}/2}}. \tag{5.12}$$

Hence, it suffices to take the expectation of a local average of eigenvalues. Fix a small $\varepsilon_1 < \omega/10$ and consider the following smooth function φ. We let

$$\begin{aligned}
&\varphi(x) = x - \gamma_{i_0}, \quad \gamma_{i_0-(N^\omega+N^{\varepsilon_1})} \le x \le \gamma_{i_0+N^\omega+N^{\varepsilon_1}},\\
&\varphi(x) = \gamma_{i_0+N^\omega+2N^{\varepsilon_1}} - \gamma_{i_0}, \quad x \ge \gamma_{i_0+N^\omega+3N^{\varepsilon_1}},\\
&\varphi(x) = \gamma_{i_0-(N^\omega-2N^{\varepsilon_1})} - \gamma_{i_0}, \quad x \le \gamma_{i_0-(N^\omega+3N^{\varepsilon_1})}.
\end{aligned} \tag{5.13}$$

We let φ smoothly interpolate between these values so that $|\varphi'(x)| \le C$, and $|\varphi^{(k)}(x)| \le CN^{(k-1)(1-\varepsilon_1)}$ for $k = 2, 3$. Note that if

$$f(x) = \frac{N}{2N^\omega + 1}\varphi(x) \tag{5.14}$$

then $\|f(x)\|_1 \le C$, $\|f'(x)\|_1 \le C$ and $\|f''(x)\|_1 \le N^\omega/N$ and so we can apply Lemma 4.2 to f.

Let I denote the interval

$$I = [\gamma_{i_0 - N^\omega - 1}, \gamma_{i_0 + N^\omega}]. \tag{5.15}$$

By rigidity we have with overwhelming probability that

$$\frac{1}{2N^\omega + 1}\left| \sum_{|i - i_0| > N^\omega} \varphi(\lambda_i) - N \int_{x \notin I} \varphi(x)\rho_{\mathrm{sc}}(x)\,\mathrm{d}x \right| \le \frac{CN^{2\varepsilon_1}}{N^{1+\omega}} \le \frac{C}{N^{1+\omega/2}} \tag{5.16}$$

by our choice of ε_1. Using the fact that

$$\gamma_i - \gamma_j = \frac{i-j}{N\rho_{\mathrm{sc}}(\gamma_i)} + \mathcal{O}\left(\frac{(i-j)^2}{N^2}\right) \tag{5.17}$$

one sees that

$$\begin{aligned} \frac{N}{2N^\omega + 1}\int_I \varphi(x)\rho_{\mathrm{sc}}(x)\,\mathrm{d}x &= \frac{N}{2N^\omega + 1}\int_I (x - \gamma_{i_0})\rho_{\mathrm{sc}}(x)\,\mathrm{d}x \\ &= \frac{-1}{2\rho(\gamma_{i_0})N} + \mathcal{O}(N^{2\omega - 2}). \end{aligned} \tag{5.18}$$

Hence, we see that for some $c > 0$ we have

$$N\mathbb{E}[\lambda_{i_0} - \gamma_{i_0}] = \sum_i \mathbb{E}[f(\lambda_i)] - N\int f(x)\rho_{\mathrm{sc}}(x)\,\mathrm{d}x - \frac{1}{2\rho_{\mathrm{sc}}(\gamma_{i_0})} + \mathcal{O}(N^{-c}). \tag{5.19}$$

Note that,

$$f(x) \to \frac{1}{2\rho_{\mathrm{sc}}(\gamma_{i_0})}(\mathbf{1}_{\{x \ge \gamma_{i_0}\}} - \mathbf{1}_{\{x \le \gamma_{i_0}\}}), \tag{5.20}$$

and so we see from Lemma 4.2 that

$$\begin{aligned} &\sum_i \mathbb{E}[f(\lambda_i)] - N\int f(x)\rho_{\mathrm{sc}}(x)\,\mathrm{d}x \\ &\quad = \frac{1}{2\pi\rho_{\mathrm{sc}}(\gamma_i)}\arcsin\left(\frac{\gamma_i}{2}\right) - \frac{1 - a_2}{2}\gamma_i + \frac{s_4}{4}(\gamma_i^3 - 2\gamma_i) + o(1). \end{aligned} \tag{5.21}$$

This yields the claim.

To obtain the result for the variance, we rearrange (3.9) into:

$$\begin{aligned} x_i(t) - \mathbb{E}[x_i(t)] + \frac{1}{N}\zeta_y &= y_i(t) - \mathbb{E}[y_i(t)] \\ &\quad + \frac{1}{N}(\zeta_x - \mathbb{E}[\zeta_x]) \\ &\quad + \frac{1}{N}\mathbb{E}[\zeta_x] + \mathbb{E}[y_i(t)] - \mathbb{E}[x_i(t)] + \mathcal{O}(N^{-1-c}). \end{aligned}$$

Squaring and using the independence between $x_i(t)$ and ζ_y, and $y_i(t)$ and ζ_x, we have

$$\begin{aligned} &\mathrm{Var}(x_i(t)) + \frac{1}{N^2}\mathrm{Var}(\zeta_y) + \frac{1}{N^2}(\mathbb{E}[\zeta_y])^2 \\ &\quad = \mathrm{Var}(y_i(t)) + \frac{1}{N^2}\mathrm{Var}(\zeta_x) + \left(\mathbb{E}\left[\frac{1}{N}\zeta_x - x_i(t) + y_i(t)\right]\right)^2 + \mathcal{O}(N^{-2-2c}). \end{aligned}$$

By (1.1), we have

$$\frac{1}{N^2}(\mathbb{E}[\zeta_y])^2 = \left(\mathbb{E}\left[\frac{1}{N}\zeta_x - x(t) + y(t)\right]\right)^2 + \mathcal{O}(N^{-2-2c}),$$

so

$$\begin{aligned}\operatorname{Var}(x_i(t)) &= \operatorname{Var}(y_i(t)) + \frac{1}{N^2}(\operatorname{Var}(\zeta_x) - \operatorname{Var}(\zeta_{\mathbf{y}})) + \mathcal{O}(N^{-2-2c}) \\ &= \operatorname{Var}(y_i(t)) + \frac{a_2 - 1}{4\pi^2 N^2}\left(\int_{-2}^{2} \Phi(x)\frac{x}{\sqrt{4-x^2}}\,\mathrm{d}x\right)^2 \\ &\quad + \frac{s_4}{2\pi^2 N^2}\left(\int_{-2}^{2} \Phi(x)\frac{2-x^2}{\sqrt{4-x^2}}\,\mathrm{d}x\right)^2 + \mathcal{O}(N^{-2-2c}).\end{aligned}$$

Using (3.11) and Proposition 5.4, we have

$$\begin{aligned}&\frac{a_2 - 1}{4\pi^2}\left(\int_{-2}^{2} \Phi(x)\frac{x}{\sqrt{4-x^2}}\,\mathrm{d}x\right)^2 + \frac{s_4}{2\pi^2}\left(\int_{-2}^{2} \Phi(x)\frac{2-x^2}{\sqrt{4-x^2}}\,\mathrm{d}x\right)^2 \\ &\quad = (a_2 - 1)\gamma_i^2 + \frac{s_4}{8} + o(1).\end{aligned}$$

The gap universality result of [17] also holds for β-ensembles, and so the proof of Theorem 1.5 is similar to the Wigner matrix case. □

5.6. *Proof of Theorem* 1.6. First we consider the case (1.14). We will smooth out the indicator function $\mathbf{1}_{\{x \le u\}}$. Let $\chi(x)$ be a smooth function so that $\chi(x) = 1$ for $x \le u$ and $\chi(x) = 0$ for $x \ge u + N^{\omega-1}$ where $0 < \omega < 1/10$, and $|\chi^{(k)}(x)| \le CN^{k(\omega-1)}$, $k = 1, 2, 3$. Let i_0 be the index of the classical eigenvalue closest to u. By rigidity for any $\varepsilon > 0$, we have with overwhelming probability

$$\left|\sum_i f(\lambda_i)\mathbf{1}_{\{\lambda_i \le u\}} - \sum_i f(\lambda_i)\chi(\lambda_i)\right| \le \sum_{|i-i_0| \le N^{\omega+\varepsilon}} |f(\lambda_i)| \le C\frac{N^{2\omega+2\varepsilon}}{N}, \tag{5.22}$$

where we used $f(u) = 0$, $|f'| \le C$ and $|\lambda_i - u| \le CN^{\omega+\varepsilon}/N$ for i appearing in the rightmost sum. It then suffices to apply Proposition 4.1 to the function $f(x)\chi(x)$.

The proof of (1.15) is similar. First, by subtracting a constant we may assume that $f(\gamma_k) = 0$. Let $\chi(x)$ be the indicator function as above, but for $u = \gamma_k$. Then, for any $\varepsilon > 0$, we have with overwhelming probability,

$$\left|\sum_{i=1}^{k} f(\lambda_i) - \sum_{i=1}^{N} f(\lambda_i)\chi(\lambda_i)\right| \le \sum_{|i-k| \le N^{\omega+\varepsilon}} |f(\lambda_i)| \le C\frac{N^{2\omega+2\varepsilon}}{N}. \tag{5.23}$$

The result follows from applying Proposition 4.1 to $f(x)\chi(x)$.

Acknowledgements. P. Sosoe is supported in part by NSF Grant DMS-1811093.

We thank the anonymous referee for a careful reading of this work and numerous comments improving the presentation.

REFERENCES

[1] ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2010). *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

[2] BAO, Z., PAN, G. and ZHOU, W. (2013). Central limit theorem for partial linear eigenvalue statistics of Wigner matrices. *J. Stat. Phys.* **150** 88–129. MR3018879 https://doi.org/10.1007/s10955-012-0663-y

[3] BEKERMAN, F. and LODHIA, A. (2018). Mesoscopic central limit theorem for general β-ensembles. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 1917–1938. MR3865662 https://doi.org/10.1214/17-AIHP860

[4] BLOEMENDAL, A., ERDŐS, L., KNOWLES, A., YAU, H.-T. and YIN, J. (2014). Isotropic local laws for sample covariance and generalized Wigner matrices. *Electron. J. Probab.* **19** Art. ID 33. MR3183577 https://doi.org/10.1214/ejp.v19-3054

[5] BOURGADE, P., ERDÖS, L. and YAU, H.-T. (2014). Edge universality of beta ensembles. *Comm. Math. Phys.* **332** 261–353. MR3253704 https://doi.org/10.1007/s00220-014-2120-z

[6] 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

[7] BOURGADE, P. and MODY, K. (2019). Gaussian fluctuations of the determinant of Wigner matrices. *Electron. J. Probab.* **24** Art. ID 96. MR4017114 https://doi.org/10.1214/19-ejp356

[8] BOUTET DE MONVEL, A. and KHORUNZHY, A. (1999). Asymptotic distribution of smoothed eigenvalue density. I. Gaussian random matrices. *Random Oper. Stoch. Equ.* **7** 1–22. MR1678012 https://doi.org/10.1515/rose.1999.7.1.1

[9] BOUTET DE MONVEL, A. and KHORUNZHY, A. (1999). Asymptotic distribution of smoothed eigenvalue density. II. Wigner random matrices. *Random Oper. Stoch. Equ.* **7** 149–168. MR1689027 https://doi.org/10.1515/rose.1999.7.2.149

[10] BRÉZIN, E. and HIKAMI, S. (1996). Correlations of nearby levels induced by a random potential. *Nuclear Phys. B* **479** 697–706. MR1418841 https://doi.org/10.1016/0550-3213(96)00394-X

[11] DALLAPORTA, S. and VU, V. (2011). A note on the central limit theorem for the eigen-value counting function of Wigner matrices. *Electron. Commun. Probab.* **16** 314–322. MR2819655 https://doi.org/10.1214/ECP.v16-1634

[12] DAVIES, E. B. (1995). The functional calculus. *J. Lond. Math. Soc.* (2) **52** 166–176. MR1345723 https://doi.org/{10.1112/jlms/52.1.166?}doi

[13] DYSON, F. J. (1962). A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3** 1191–1198. MR0148397 https://doi.org/10.1063/1.1703862

[14] EDELMAN, A., GUIONNET, A. and PÉCHÉ, S. (2016). Beyond universality in random matrix theory. *Ann. Appl. Probab.* **26** 1659–1697. MR3513602 https://doi.org/10.1214/15-AAP1129

[15] ERDŐS, L., RAMÍREZ, J., SCHLEIN, B., TAO, T., VU, V. and YAU, H.-T. (2010). Bulk universality for Wigner Hermitian matrices with subexponential decay. *Math. Res. Lett.* **17** 667–674. MR2661171 https://doi.org/10.4310/MRL.2010.v17.n4.a7

[16] 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

[17] 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

[18] ERDŐS, L., YAU, H.-T. and YIN, J. (2011). Universality for generalized Wigner matrices with Bernoulli distribution. *J. Comb.* **2** 15–81. MR2847916 https://doi.org/10.4310/JOC.2011.v2.n1.a2

[19] FORRESTER, P. J. and RAINS, E. M. (2001). Interrelationships between orthogonal, unitary and symplectic matrix ensembles. In *Random Matrix Models and Their Applications*. *Math. Sci. Res. Inst. Publ.* **40** 171–207. Cambridge Univ. Press, Cambridge. MR1842786

[20] GUSTAVSSON, J. (2005). Gaussian fluctuations of eigenvalues in the GUE. *Ann. Inst. Henri Poincaré Probab. Stat.* **41** 151–178. MR2124079 https://doi.org/10.1016/j.anihpb.2004.04.002

[21] HE, Y. and KNOWLES, A. (2017). Mesoscopic eigenvalue statistics of Wigner matrices. *Ann. Appl. Probab.* **27** 1510–1550. MR3678478 https://doi.org/10.1214/16-AAP1237

[22] KNOWLES, A. and BENAYCH-GEORGES, F. (2016). Lectures on the local semi-circle law for Wigner matrices. SMF volume Panoramas et Syntheses.

[23] KNOWLES, A. and YIN, J. (2013). The isotropic semicircle law and deformation of Wigner matrices. *Comm. Pure Appl. Math.* **66** 1663–1750. MR3103909 https://doi.org/10.1002/cpa.21450

[24] KUIJLAARS, A. B. J. and MCLAUGHLIN, K. T.-R. (2000). Generic behavior of the density of states in random matrix theory and equilibrium problems in the presence of real analytic external fields. *Comm. Pure Appl. Math.* **53** 736–785. MR1744002 https://doi.org/10.1002/(SICI)1097-0312(200006)53:6<736::AID-CPA2>3.0.CO;2-5

[25] 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

[26] LEE, J. O. and SCHNELLI, K. (2018). Local law and Tracy–Widom limit for sparse random matrices. *Probab. Theory Related Fields* **171** 543–616. MR3800840 https://doi.org/10.1007/s00440-017-0787-8

[27] LODHIA, A. and SIMM, N. J. (2015). Mesoscopic linear statistics of Wigner matrices. Preprint. Available at arXiv:1503.03533.

[28] LYTOVA, A. and PASTUR, L. (2009). Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. *Ann. Probab.* **37** 1778–1840. MR2561434 https://doi.org/10.1214/09-AOP452

[29] O'ROURKE, S. (2010). Gaussian fluctuations of eigenvalues in Wigner random matrices. *J. Stat. Phys.* **138** 1045–1066. MR2601422 https://doi.org/10.1007/s10955-009-9906-y

[30] PASTUR, L. and SHCHERBINA, M. (2011). *Eigenvalue Distribution of Large Random Matrices*. *Mathematical Surveys and Monographs* **171**. Amer. Math. Soc., Providence, RI. MR2808038 https://doi.org/10.1090/surv/171

[31] SHCHERBINA, M. (2011). Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices. *Zh. Mat. Fiz. Anal. Geom.* **7** 176–192, 197, 199. MR2829615

[32] SHCHERBINA, M. (2016). Fluctuations of the eigenvalue number in the fixed interval for β-models with $\beta = 1, 2, 4$. In *Theory and Applications in Mathematical Physics* 131–146. World Sci. Publ., Hackensack, NJ. MR3468249

[33] TAO, T. and VU, V. (2011). Random matrices: Localization of the eigenvalues and the necessity of four moments. *Acta Math. Vietnam.* **36** 431–449. MR2908537

[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