AIP | Journal of Mathematical Physics

Local deformed semicircle law and complete delocalization for Wigner matrices with random potential

Ji Oon Lee and Kevin Schnelli

Citation: Journal of Mathematical Physics **54**, 103504 (2013); doi: 10.1063/1.4823718
View online: http://dx.doi.org/10.1063/1.4823718
View Table of Contents: http://scitation.aip.org/content/aip/journal/jmp/54/10?ver=pdfcov
Published by the AIP Publishing

JOURNAL OF MATHEMATICAL PHYSICS **54**, 103504 (2013)

Local deformed semicircle law and complete delocalization for Wigner matrices with random potential

Ji Oon Lee[1,a)] and Kevin Schnelli[2,b)]
[1]*Department of Mathematical Sciences, KAIST, Daejeon 305-701, South Korea*
[2]*School of Mathematics, Institute for Advanced Study, Princeton, New Jersey 08540, USA*

(Received 5 June 2013; accepted 16 September 2013; published online 10 October 2013)

We consider Hermitian random matrices of the form $H = W + \lambda V$, where W is a Wigner matrix and V is a diagonal random matrix independent of W. We assume subexponential decay for the matrix entries of W and we choose $\lambda \sim 1$ so that the eigenvalues of W and λV are of the same order in the bulk of the spectrum. In this paper, we prove for a large class of diagonal matrices V that the local deformed semicircle law holds for H, which is an analogous result to the local semicircle law for Wigner matrices. We also prove complete delocalization of eigenvectors and other results about the positions of eigenvalues. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4823718]

I. INTRODUCTION

Consider large matrices whose entries are random variables. Famous examples of such matrices are Wigner matrices: a Wigner matrix is an $N \times N$ real or complex matrix $W = (w_{ij})$ whose entries are independent random variables with mean zero and variance 1/*N*, subject to the symmetry constraint $w_{ij} = \overline{w}_{ji}$. The eigenvalues of Wigner matrices are highly correlated; the empirical density of eigenvalues converges to the Wigner semicircle law[40] in the large *N* limit. Under some additional moment assumptions on the entries this convergence also holds on very small scales: denoting by $G_W(z) = (W - z)^{-1}$, $z \in \mathbb{C}^+$, the resolvent or Green function of W, the convergence of the empirical eigenvalue distribution on scale η at an energy $E \in \mathbb{R}$ is equivalent to the convergence of the averaged Green function $m_W(z) = N^{-1}\mathrm{Tr}\, G_W(z)$, $z = E + \mathrm{i}\eta$. The convergence of $m_W(z)$ at the optimal scale N^{-1}, up to logarithmic corrections, the so-called *local semicircle law*, was established for Wigner matrices in a series of papers,[16–18] where it was also shown that the eigenvectors of Wigner matrices are completely delocalized. The proof is based on a self-consistent equation for $m_W(z)$ and the continuity of the Green function $G(z)$ in the spectral parameter z. In Refs. 22 and 23, convergence of Green function entries was established on optimal scales. Precise estimates on the averaged Green function $m_W(z)$ and on the eigenvalue locations are essential ingredients for proving bulk universality[19,20] and edge universality[24] for Wigner matrices.

Diagonal matrices with independent and identically distributed (i.i.d.) random entries are another example of random square matrices. Their eigenvalues are independent, hence uncorrelated, and their eigenvectors are localized. Physically, the diagonal matrix may represent an on-site random potential on a lattice system. Compared to the mean-field nature of Wigner matrices, which are in the weak disorder or the delocalization regime, the diagonal randomness also provides a good example of the strong disorder or localization regime.

In this paper we consider the interpolation of the two, i.e., the $N \times N$ random matrix

$$H = \lambda V + W\,, \qquad \lambda \in \mathbb{R}\,, \tag{1.1}$$

---

a)Electronic mail: jioon.lee@kaist.edu
b)Electronic mail: kschnelli@ias.edu

where V is a real diagonal random matrix, or a "random potential," and W is a real symmetric or complex Hermitian Wigner matrix independent of V. The matrix V is properly normalized so that the typical eigenvalues of V and W are of the same order. (See Definition 2.1 for a precise statement.) The parameter λ determines the relative strength of each part in this model.

For $\lambda \sim 1$ the eigenvalue density is not solely determined by V or W in the limit $N \to \infty$, but can be described by a functional equation for the Stieltjes transforms of the limiting eigenvalue distributions of V and W; see Refs. 33 and 34. In general, this limiting eigenvalue distribution, referred to as the *deformed semicircle law*, is different from the semicircle distribution. The equal strength of V and W makes it non-trivial to find the spectral properties of H. We remark that there are some results related to this model.[4,8,29]

When W belongs to the Gaussian Unitary Ensemble (GUE), H is called the *deformed GUE*, and it can describe Dyson Brownian motion on the real line; see, e.g., Refs. 10 and 27. There has been much important work on various scales of λ: Related to symmetry-breaking, transition statistics for eigenvalues in the bulk, especially the nearest neighbor spacing, were studied in Refs. 25 and 32 for $\lambda \sim N^{1/2}$. In this situation, the diagonal part λV controls the average density, while the GUE part induces fluctuation of eigenvalues. For $\lambda \lesssim 1$, it was shown in Ref. 35 that universality of eigenvalue correlation functions holds in the bulk of the spectrum. Concerning the edge behaviour, it was shown in Ref. 28 that the transition from the Tracy-Widom to the standard Gaussian distribution occurs on the scale $\lambda \sim N^{-1/6}$. For $\lambda \ll N^{-1/6}$, the Tracy-Widom distribution for the edge eigenvalues was established by Shcherbina.[36]

In this paper, we prove, for $\lambda \lesssim 1$ and a large class of random potentials, convergence of the empirical density of eigenvalues down to the optimal scale $1/N$, i.e., we show a *local deformed semicircle law* for the averaged Green function $m_H(z) = N^{-1}\,\mathrm{Tr}(H - z)^{-1}$, $z = E + \mathrm{i}\eta$, for all $\eta \gg N^{-1}$. Unlike in the Wigner case, the diagonal disorder of V prevents the diagonal Green function entries from concentrating around $m_H(z)$ for $\lambda \neq 0$. Following Ref. 24 we derive a self-consistent equation for $m_H(z)$, whose analysis requires a stability estimate that forces interesting conditions on V and λ. As an intermediate result, we obtain a weak local deformed semicircle law for $m_H(z)$ and complete delocalization of the eigenvectors of H up to the edge. In Ref. 23 a "fluctuation average lemma" was proven that yielded optimal rigidity estimates on the location of the eigenvalues of H in the bulk[23] and up to the edge.[24] Combining the weak deformed semicircle law with the "fluctuation average lemma"[23] we obtain the convergence of $m_H(z)$ on the optimal scale. However, the self-averaging mechanism of the Wigner matrix W in the bulk is only observed after the leading fluctuations stemming from V are subtracted. For example, for $\lambda \sim 1$, the rigidity of eigenvalue location is weaker than in the Wigner case, but we show that the eigenvalue spacing is rigid in the bulk on intermediate scales.

The paper is organized as follows: In Sec. II, we introduce the precise definition and assumptions of the model, state the main results, and give a short outline of the proofs. Our assumptions on λV mainly depend on the behaviour of the deformed semicircle law as described in Lemmas 2.4 and 2.7 (see also Lemma 3.2 in Sec. III). For similar results on the deformed semicircle law; see Refs. 5 and 36. In Sec. III, we prove a weak local (deformed) semicircle law and complete delocalization of eigenvectors. The proof of the local deformed semicircle law follows closely the proof of the weak local semicircle law for sparse random matrices given in Ref. 12. In Sec. IV, we give a proof of the average fluctuation lemma (Lemma 4.1). The proof is inspired by Erdős,[14] where fluctuation averages are considered for generalized Wigner and random band matrices. The combination of the fluctuation average lemma with the weak local (deformed) semicircle law, yields a proof of the strong local deformed semicircle law as in Refs. 12 and 24. In Sec. V, we identify the leading corrections due to the random diagonal part to the strong local semicircle law on scale $N^{-1/2}$, in the bulk of the spectrum. In Sec. VI, we establish, following the Helffer-Sjöstrand argument given in Ref. 15, estimates on the density of states and the rigidity of eigenvalues. Using the results obtained in Sec. V, we obtain estimates on the rigidity of the eigenvalue spacing on intermediate scales in the bulk of the spectrum. Technical details about the square root behaviour and the stability bounds for the deformed semicircle law are given in the Appendix.

II. DEFINITION AND RESULTS

In this section, we define our model and state our main results.

A. Free convolution

As first shown in Ref. 33 the limiting spectral distribution of the interpolating model (1.1) is given by the *(additive) free convolution measure* of the limiting distribution of the entries of λV and μ_{sc}, the semicircular measure. In a more general setting, the free convolution measure, $\mu_1 \boxplus \mu_2$, of two probability measures μ_1 and μ_2, is defined as the distribution of the sum of two freely independent non-commutative random variables, having distributions μ_1, μ_2, respectively; we refer to Refs. 1, 26, 30, and 39. The (additive) free convolution may also be described in terms of the Stieltjes transform: Let μ be a probability measure on $\mathbb{R}$, then we define the Stieltjes transform of μ by

$$m_\mu(z) := \int_{\mathbb{R}} \frac{\mathrm{d}\mu(x)}{x-z}\,, \qquad z \in \mathbb{C}^+\,. \tag{2.1}$$

Note that $m_\mu(z)$ is an analytic function in the upper half plane, satisfying $\lim_{y\to\infty} \mathrm{i}y m_\mu(\mathrm{i}y) = 1$. As shown in Refs. 6 and 38, the free convolution has the following property: Denote by m_{μ_1}, m_{μ_2}, $m_{\mu_1\boxplus\mu_2}$, the Stieltjes transforms of μ_1, μ_2, $\mu_1 \boxplus \mu_2$, respectively. Then there exist two analytic functions ω_1, ω_2, from $\mathbb{C}^+$ to $\mathbb{C}^+$, satisfying $\lim_{y\to\infty} \omega_i(\mathrm{i}y)/\mathrm{i}y = 1$, $(i = 1, 2)$, such that

$$\begin{aligned} m_{\mu_1\boxplus\mu_2}(z) &= m_{\mu_1}(\omega_1(z)) = m_{\mu_2}(\omega_2(z))\,, \\ \omega_1(z) + \omega_2(z) &= z - \frac{1}{m_{\mu_1\boxplus\mu_2}(z)}\,, \end{aligned} \tag{2.2}$$

for $z \in \mathbb{C}^+$. The functions ω_i are referred to as subordination functions. Note that (2.2) also shows that $\mu_1 \boxplus \mu_2 = \mu_2 \boxplus \mu_1$. It was pointed out in Refs. 3 and 9 that the system (2.2) may be used as an alternative definition of the free convolution. In particular, given μ_1, μ_2, the system (2.2) has a unique solution $(m_{\mu_1\boxplus\mu_2}, \omega_1, \omega_2)$.

The system (2.2) has been used in Ref. 34 to exploit the limiting eigenvalue distributions for random matrices of the form $A + UBU^*$, with A, B deterministic or random $N \times N$ matrices and U an $N \times N$ random Haar unitary matrix. Free probability theory turned out to be a natural setting for studying global laws for such ensembles; see, e.g., Refs. 1 and 39. For more recent treatments, including local laws, we refer to Refs. 4, 8, and 29.

In case we choose the measure μ_2 as the standard semicircular law $\mathrm{d}\mu_{sc}(E) = \frac{1}{2\pi}\sqrt{(4-E^2)_+}\mathrm{d}E$, a simple computation reveals that the Stieltjes transform, $m_{\mu_{sc}} \equiv m_{sc}$, satisfies

$$m_{sc}(z) = -\frac{1}{z + m_{sc}(z)}\,, \qquad z \in \mathbb{C}^+.$$

Using this information, we can reduce the system (2.2), to the self-consistent equation

$$m_{fc}(z) = \int \frac{\mathrm{d}\mu(x)}{x - z - m_{fc}(z)}\,, \qquad z \in \mathbb{C}^+, \tag{2.3}$$

$\operatorname{Im} m_{fc}(z) \geq 0$, for $z \in \mathbb{C}^+$, with $\lim_{y\to\infty} \mathrm{i}y\, m_{fc}(\mathrm{i}y) = 1$, where we have abbreviated $\mu \equiv \mu_1$. Equation (2.3) is often called the *Pastur relation*. A slightly modified version of the functional equation (2.3) is the starting point of the analysis given in Ref. 33 and also in the present paper; see (2.9).

The (unique) solution of (2.3) has first been studied in details in Ref. 5. In particular, it has been shown that $\limsup_{\eta\searrow 0} \operatorname{Im} m_{fc}(E + \mathrm{i}\eta) < \infty$, $E \in \mathbb{R}$, and hence the free convolution measure $\mu_{fc} \equiv \mu \boxplus \mu_{sc}$ is absolutely continuous (for simplicity we denote the density also with μ_{fc}) and we conclude from the Stieltjes inversion formula that

$$\mu_{fc}(E) = \lim_{\eta\searrow 0} \frac{1}{\pi} \operatorname{Im} m_{fc}(E + \mathrm{i}\eta)\,, \qquad E \in \mathbb{R}\,.$$

Moreover, it was shown in Ref. 5 that the density μ_{fc} is analytic in the interior of the support of μ_{fc}. We refer to, e.g., Ref. 2 for further results on the regularity of the free convolution measure.

B. Assumptions

In this section, we define the model (1.1) in detail and list our main assumptions.

1. Definition of the model

Definition 2.1. Let W be an $N \times N$ random matrix, whose entries, (w_{ij}), are independent, up to the symmetry constraint $w_{ij} = \overline{w}_{ji}$, centered, complex random variables with variance N^{-1} and subexponential decay, i.e.,

$$\mathbb{P}\left(\sqrt{N}|w_{ij}| > x\right) \leq C_0\,\mathrm{e}^{-x^{1/\theta}}\,, \tag{2.4}$$

for some positive constants C_0 and $\theta > 1$. In particular,

$$\mathbb{E}w_{ij} = 0\,, \qquad \mathbb{E}|w_{ij}|^p \leq C\frac{(\theta p)^{\theta p}}{N^{p/2}}\,, \qquad (p \geq 3)\,, \tag{2.5}$$

and,

$$\mathbb{E}w_{ii}^2 = \frac{1}{N}\,, \qquad \mathbb{E}|w_{ij}|^2 = \frac{1}{N}\,, \qquad \mathbb{E}w_{ij}^2 = 0\,, \qquad (i \neq j)\,. \tag{2.6}$$

Remark 2.2. We remark that all our methods also apply to *symmetric* Wigner matrices, i.e., when (w_{ij}) are centered, real random variables with variance N^{-1}, and subexponential decay. In this case, (2.6) gets replaced by

$$\mathbb{E}w_{ii}^2 = \frac{2}{N}\,, \qquad \mathbb{E}w_{ij}^2 = \frac{1}{N}\,, \qquad (i \neq j)\,. \tag{2.7}$$

Let $V = (v_i)$ be an $N \times N$ diagonal random matrix, whose entries (v_i) are real, centered, i.i.d. random variables, independent of $W = (w_{ij})$, with law μ. More assumptions on μ will be stated below. For $\lambda \in \mathbb{R}$, we consider the random matrix

$$H = (h_{ij}) := \lambda V + W\,. \tag{2.8}$$

In Sec. II B 3, we will choose μ, such that $\operatorname{supp}\mu = [-1, 1]$, but we observe that varying λ is equivalent to changing the support of μ.

We define the resolvent, or *Green function*, $G(z)$, and the averaged Green function, $m(z)$, of H by

$$G(z) = (G_{ij}(z)) := \frac{1}{\lambda V + W - z}\,, \qquad m(z) := \frac{1}{N}\operatorname{Tr} G(z)\,, \qquad z \in \mathbb{C}^+\,.$$

Frequently, we abbreviate $G \equiv G(z)$, $m \equiv m(z)$, etc.

2. Free convolution

Following the discussion in Subsection II A, we define m_{fc}^{λ} as the solution to

$$m_{fc}^{\lambda}(z) = \int \frac{\mathrm{d}\mu(v)}{\lambda v - z - m_{fc}^{\lambda}(z)}\,, \qquad z \in \mathbb{C}^+\,, \tag{2.9}$$

with $\operatorname{Im} m_{fc}^{\lambda}(z) \geq 0$, $z \in \mathbb{C}^+$. We denote by μ_{fc}^{λ} the corresponding probability measure. For simplicity, we discard the superscript λ from our notation. Let us list some easy examples:

i. Choosing $\mu = \delta_1$, one directly sees that μ_{fc} is a semicircle law of radius 2 centered at λ.

ii. For the choice $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$, (2.9) reduces to a cubic equation and the support of the measure μ_{fc} can be inferred from a simple analysis of the discriminant of that equation. As it turns out, the support of μ consists of a single interval for $\lambda \leq 1$ and of two intervals for

$\lambda > 1$. For simplicity, we will exclude the possibility of μ_{fc} having support on several disjoint intervals in the following. However, some of our results can be generalized to this setting.

iii. If μ is the standard Gaussian measure, no closed expression for μ_{fc} exists, but the moments of μ_{fc} can be computed recursively; see Ref. 7. Moreover, the density of μ_{fc} is a smooth function with Gaussian tails. Although the Gaussian case is important, we will not deal with measures of unbounded support, but comment on the Gaussian case in Remarks 2.11 and 2.16.

3. Assumptions on λV and μ

We state our assumptions on μ the distributions of the entries (v_i) of V. From now on, we choose μ with $\mu = \text{supp}[-1, 1]$. Depending on the size of the "perturbation" parameter λ, we have to distinguish two cases: For $|\lambda| \le 1$, we will assume the following:

Assumption 2.3. [Small λ] The entries of the diagonal matrix $V = (v_i)$ are centered, real, i.i.d. random variables, independent of $W = (w_{ij})$. For $|\lambda| \le 1$, we assume that the distribution of (v_i) has a continuous density $\mu(v)$, such that $\mu(v) > 0$, $v \in (-1, 1)$, and $\mu(v) = 0$, $v \notin [-1, 1]$.

This assumption ensures that the deformed semicircle law μ_{fc} is supported on a single interval $[L_1, L_2]$, with a square root behaviour at the edges. More precisely, we have the following result:

Lemma 2.4. Let $\lambda \le 1$ and assume that μ satisfies Assumption 2.3. Then there are $-\infty < L_1 < 0 < L_2 < \infty$, such that supp $\mu_{fc} = [L_1, L_2]$. *Moreover, denoting by κ_E the distance to the endpoints of the support of μ_{fc}, i.e.,*

$$\kappa_E := \min\{|E - L_1|, |E - L_2|\}, \qquad E \in \mathbb{R}, \tag{2.10}$$

there exists $C \ge 1$ such that

$$C^{-1}\sqrt{\kappa_E} \le \mu_{fc}(E) \le C\sqrt{\kappa_E}, \qquad E \in [L_1, L_2]. \tag{2.11}$$

In a slightly different setting this lemma has been proven in Ref. 36; see also Refs. 5 and 31. In the Appendix we explain how to adopt the proof in Ref. 36 to our setting.

Remark 2.5. Above, we have chosen λ to be independent of N. However, we may choose $\lambda = CN^{-\delta}$, for some constants C and $\delta > 0$. In this setting *all* our results hold true as well, in particular, in all the bounds one may simply replace λ by $CN^{-\delta}$.

Remark 2.6. Determining the endpoints L_1 and L_2 of the support of μ_{fc} explicitly is, in general, not possible, since it involves solving an implicit equation. However, for λ sufficiently small, one can show that $L_1 = -2\sqrt{1+\lambda^2} + \mathcal{O}(\lambda^3)$ and $L_2 = 2\sqrt{1+\lambda^2} + \mathcal{O}(\lambda^3)$. Also the measure μ_{fc} is $\mathcal{O}(\lambda)$-close to the semicircular measure of radius $2\sqrt{1+\lambda^2}$ in an appropriate distance, but we refrain from going into the details of this "perturbative approach."

For $|\lambda| > 1$, we have to strengthen the above assumptions, since the square root behaviour at the endpoint of the support of μ_{fc} may fail, for λ large enough. We call a probability measure μ a *Jacobi* measure if it is given by a density of the form

$$\mu(v) = Z^{-1}(1+v)^{\mathrm{a}}(1-v)^{\mathrm{b}} d(v) \mathbb{1}_{[-1,1]}(v), \tag{2.12}$$

where $d \in C^1([-1, 1])$, with $d(v) > 0$, $v \in [-1, 1]$, $-1 < \mathrm{a}, \mathrm{b} < \infty$, and Z a normalization constant.

We have the following result.

Lemma 2.7. Let μ be a centered Jacobi measure; see (2.12). *Then, for any $\lambda \in \mathbb{R}$, there are $-\infty < L_1 < 0 < L_2 < \infty$, such that* supp $\mu_{fc} = [L_1, L_2]$. *Moreover,*

(1) *for $-1 < \mathrm{a}, \mathrm{b} \le 1$, for any $\lambda \in \mathbb{R}$, μ_{fc} has the square root behaviour* (2.11);

(2) *for $1 < \mathrm{a}, \mathrm{b} < \infty$, there exists $\lambda_1 \equiv \lambda_1(\mu) > 1$ and $\lambda_2 \equiv \lambda_2(\mu) > 1$ such that,*

  (2a) *for $|\lambda| < \lambda_1$, $|\lambda| < \lambda_2$, μ_{fc} has the square root behaviour at both endpoints*;

(2b) *for* $|\lambda| < \lambda_1$, $|\lambda| > \lambda_2$, μ_{fc} *has the square root behaviour at the lower endpoint of the support (i.e., for* $E \in [L_1, 0]$*), but there is* $C \geq 1$, *such that*

$$C^{-1}(L_2 - E)^{\mathrm{b}} \leq \mu_{fc}(E) \leq C(L_2 - E)^{\mathrm{b}}\,, \qquad E \in [0, L_2]\,. \tag{2.13}$$

Analogue statements hold for $|\lambda| > \lambda_1$, $|\lambda| < \lambda_2$, *etc.*

The proof of the lemma is given in the Appendix.

For our methods to work, we have to exclude situation (2b) of Lemma 2.7. For $|\lambda| > 1$, we will thus assume:

Assumption 2.8. [Large λ] The entries of the diagonal matrix $V = (v_i)$ are centered, real, i.i.d. random variables, independent of $W = (w_{ij})$. For $|\lambda| > 1$, we assume that the distribution of the (v_i) is given by a centered Jacobi measure, and λ and a, b are chosen as in (1) or ($2a$) of Lemma 2.7.

4. Notations and conventions

To state our main results, we need some more notations and conventions. For high probability estimates we use two parameters $\xi \equiv \xi_N$ and $\varphi \equiv \varphi_N$: We assume that

$$a_0 < \xi \leq A_0 \log\log N\,, \qquad \varphi = (\log N)^C\,, \tag{2.14}$$

for some fixed constants $a_0 > 2, A_0 \geq 10, C \geq 1$. These constants are chosen such that large deviation estimates in Lemma 3.5 hold. They only depend on θ and C_0 in (2.4) and will be kept fixed in the following.

Definition 2.9. For $\nu > 0$, *we say an event* Ω *has* (ξ, ν)*-high probability, if*

$$\mathbb{P}(\Omega^c) \leq \mathrm{e}^{-\nu(\log N)^{\xi}}\,,$$

for N sufficiently large.

Similarly, for a given event Ω_0 *we say an event* Ω *holds with* (ξ, ν)*-high probability on* Ω_0, *if*

$$\mathbb{P}(\Omega_0 \cap \Omega^c) \leq \mathrm{e}^{-\nu(\log N)^{\xi}}\,,$$

for N sufficiently large.

For brevity, we occasionally say an event holds with high probability, when we mean (ξ, ν)-high probability. We do not keep track of the explicit value of ν in the following, allowing ν to decrease from line to line such that $\nu > 0$. From our proof it becomes apparent that such reductions occur only finitely many times.

We use the symbols $\mathcal{O}(\,\cdot\,)$ and $o(\,\cdot\,)$ for the standard big-O and little-o notation. The notations $\mathcal{O}$, o, $\ll$, $\gg$, usually refer to the limit $N\to\infty$. Here $a \ll b$ means $a = o(b)$. We use c and C to denote positive constants that do not depend on N. Their value may change from line to line. Finally, we write $a \sim b$, if there is $C \geq 1$ such that $C^{-1}|b| \leq |a| \leq C|b|$, and, occasionally, we write for N-dependent quantities $a_N \lesssim b_N$, if there exist constants $C, c > 0$ such that $|a_N| \leq C(\varphi_N)^{c\xi}|b_N|$.

C. Results

In this subsection we state our main results. The presentation of our results follows the one mentioned in Ref. 12.

Since we choose the measure μ to be centered, we may assume that $\lambda \geq 0$, without loss of generality in the following. Fix some $\lambda_0 > 0$, then we assume that the perturbation parameter λ is in the domain

$$\mathcal{D}_{\lambda_0} := \{\lambda \in \mathbb{R}^+ \,:\, \lambda \leq \lambda_0\}\,. \tag{2.15}$$

Here λ_0 is an arbitrary constant, but recall that in case $\lambda_0 > 1$, Assumption 2.8 may not be satisfied for a, b > 1.

We define the spectral parameter $z = E + \mathrm{i}\eta$, with $E \in \mathbb{R}$ and $\eta > 0$. Let $E_0 \geq 3 + \lambda_0$ and define the domain

$$\mathcal{D}_L := \{z = E + \mathrm{i}\eta \in \mathbb{C} \,:\, |E| \leq E_0 \,, (\varphi_N)^L \leq N\eta \leq 3N\} \,, \tag{2.16}$$

with $L \equiv L(N)$, such that $L \geq 12\xi$. Here, we chose E_0 bigger than $3 + \lambda$, since we know that the spectrum of W lies in the set $\{E \in \mathbb{R} \,:\, |E| \leq 3\}$ with high probability. Thus spectral perturbation theory implies that the spectrum of H is contained in $\{E \in \mathbb{R} \,:\, |E| \leq 3 + \lambda\}$, with high probability.

Recall the definition of κ_E, the distance to the endpoints of the support of μ_{fc}, in (2.10). In the following, we often abbreviate $\kappa \equiv \kappa_E$.

1. Local laws

Theorem 2.10. [Strong local law] *Let $H = \lambda V + W$, where W satisfies the assumptions in Definition 2.1 and λV satisfies Assumption 2.3 or 2.8. Let*

$$\xi = \frac{A_0 + o(1)}{2} \log\log N \,. \tag{2.17}$$

Then there are constants $\nu > 0$ and c_1, depending on the constants θ and C_0 in (2.4), λ_0 in (2.15), E_0 in (2.16), A_0 in (2.17), and the measure μ, such that for $L \geq 40\xi$, the events

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ \left| m(z) - m_{fc}(z) \right| \leq (\varphi_N)^{c_1 \xi} \left(\min\left\{ \frac{\lambda^{1/2}}{N^{1/4}}, \frac{\lambda}{\sqrt{\kappa + \eta}} \frac{1}{\sqrt{N}} \right\} + \frac{1}{N\eta} \right) \right\} \tag{2.18}$$

and

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ \max_{i \neq j} |G_{ij}(z)| \leq (\varphi_N)^{c_1 \xi} \left(\sqrt{\frac{\operatorname{Im} m_{fc}(z)}{N\eta}} + \frac{1}{N\eta} \right) \right\} \tag{2.19}$$

both have (ξ, ν)-high probability.

Remark 2.11. If we choose the entries of $V = (v_i)$ to be independent standard Gaussian random variables, we have the following result: Under the same assumptions as in Theorem 2.10 (except Assumption 2.3 or 2.8) and with similar constants, the events

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ \left| m(z) - m_{fc}(z) \right| \leq (\varphi_N)^{c_1 \xi} \left(\frac{\lambda}{\sqrt{N}} + \frac{1}{N\eta} \right) \right\} \tag{2.20}$$

and

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ \max_{i \neq j} |G_{ij}(z)| \leq (\varphi_N)^{c_1 \xi} \left(\sqrt{\frac{\operatorname{Im} m_{fc}(z)}{N\eta}} + \frac{1}{N\eta} \right) \right\} \tag{2.21}$$

both have (ξ, ν)-high probability. Note, however, that the result only applies to the compact domain $\mathcal{D}_L$ of the spectral parameter z (i.e., for some fixed E_0), but the limiting spectrum of $H = \lambda V + W$ has unbounded support. The proof of the estimates (2.20) and (2.21) is similar to the proof of Theorem 2.10 and we refrain from stating it explicitly.

For $\lambda = 0$, we have $m_{fc} = m_{sc}$, where m_{sc} is the Stieltjes transform of the standard semicircle law. In this case stronger estimates have been obtained; see, e.g., Ref. 11. Roughly speaking, in this situation we have the high probability bounds

$$|m(z)-m_{sc}(z)|\lesssim\frac{1}{N\eta}\qquad\text{and}\qquad|G_{ij}(z)-\delta_{ij}m(z)|\lesssim\sqrt{\frac{\operatorname{Im}m_{sc}(z)}{N\eta}}+\frac{1}{N\eta}\,,\qquad(\lambda=0)\,,\tag{2.22}$$

(up to logarithmic corrections), within the range of admitted parameters.

This suggests that the bound on $G_{ij}(z)$, $(i\neq j)$, in (2.19) is optimal. However, for $\lambda\neq 0$, $G_{ii}(z)$ strongly depends on v_i, $i\in\{1,\dots,N\}$, and the diagonal resolvent entries do not concentrate round their mean $m(z)$. This becomes apparent from Schur's complement formula (see, e.g., (3.17)) and one easily establishes that $|G_{ii}(z)-m(z)|\le C\lambda+o(1)$, with high probability.

Comparing the estimate on $m-m_{fc}$ in (2.18) with the corresponding estimate in (2.22), one may suspect that the leading correction terms in (2.18) stem from fluctuations of the random variables (v_i). The next theorem asserts that this is indeed true, at least in the bulk of the spectrum: There are random variables, $\zeta_0\equiv\zeta_0^N(z)$, which depend on the random variables (v_i), but are independent of the random variables (w_{ij}), such that $|m(z)-m_{fc}(z)-\zeta_0(z)|\lesssim(N\eta)^{-1}$ with high probability in the bulk of the spectrum; see (2.23). Concerning the spectral edge, we remark that the estimate in (2.18) is optimal for $\lambda\ll N^{-1/6}$, but it is not known whether $\lambda^{1/2}N^{-1/4}$ is the optimal rate for $\lambda\gg N^{-1/6}$.

To state our next result, we define the domain

$$\mathcal{B}_L:=\mathcal{D}_L\cap\{z=E+\mathrm{i}\eta\in\mathbb{C}\,:\,\sqrt{\kappa_E+\eta}\ge(\varphi_N)^{L\xi}N^{-1/4}\}\,.$$

We have the following result:

Theorem 2.12. *Let $H=\lambda V+W$, where W satisfies the assumptions in Definition 2.1 and λV satisfies Assumption 2.3 or 2.8. Then, for any $z\in\mathcal{D}_L$, $\lambda\in\mathcal{D}_{\lambda_0}$, there exist random variables $\zeta_0(z)\equiv\zeta_0^N(z)$, depending only on (v_i) such that, with the same constants as in Theorem 2.10, the event*

$$\bigcap_{\substack{z\,\in\,\mathcal{B}_L\\ \lambda\,\in\,\mathcal{D}_{\lambda_0}}}\left\{\left|m(z)-m_{fc}(z)-\zeta_0(z)\right|\le(\varphi_N)^{c_1\xi}\frac{1}{N\eta}\right\}\tag{2.23}$$

has (ξ,ν)-high probability. The random variables $\zeta_0(z)$ have the following property: The event

$$\bigcap_{\substack{z\,\in\,\mathcal{D}_L\\ \lambda\,\in\,\mathcal{D}_{\lambda_0}}}\left\{|\zeta_0(z)|\le(\varphi_N)^{c_1\xi}\min\left\{\frac{\lambda^{1/2}}{N^{1/4}},\frac{\lambda}{\sqrt{\kappa+\eta}}\frac{1}{\sqrt{N}}\right\}\right\}\tag{2.24}$$

has (ξ,ν)-high probability.

Remark 2.13. The estimates in (2.23) and (2.24) need some explanation: Choosing E in the bulk of the spectrum, i.e., $\kappa_E\ge\varkappa$, for some $\varkappa>0$, we have

$$|m(z)-m_{fc}(z)-\zeta_0(z)|\le(\varphi_N)^{c_1\xi}\frac{1}{N\eta}\,,\qquad|m(z)-m_{fc}(z)|\le(\varphi_N)^{c_1\xi}\left(\frac{\lambda}{\sqrt{N}}+\frac{1}{N\eta}\right)\,,\tag{2.25}$$

with high probability and the estimate seems to be optimal. In particular, on microscopic scales, $\eta\ll N^{-1/2}$, the local fluctuations stem from the Wigner matrix W, whereas on intermediate scales, $\eta\sim N^{-1/2}$, the fluctuations due to the Wigner matrix are of the same size as the fluctuations due to the diagonal matrix V. Finally, on macroscopic scales, $\eta\sim 1$, the fluctuations are dominated by the matrix V.

Remark 2.14. The random variable $\zeta_0 \equiv \zeta_0(z)$ is defined as the solution to a quadratic equation; see (5.2) below. In the bulk of the spectrum, we can approximate ζ_0 by

$$\widetilde{\zeta}_0(z) := \left(1 - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2}\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^{N} \frac{1}{\lambda v_i - z - m_{fc}(z)} - \int \frac{\mathrm{d}\mu(v)}{\lambda v - z - m_{fc}(z)}\right), \quad z \in \mathcal{D}_L, \tag{2.26}$$

that only depends on (v_i). Note that $\widetilde{\zeta}_0(z)$ embodies, up to a deterministic (z-dependent) prefactor, the expected fluctuation corresponding to the $\lambda N^{-1/2}$ term in (2.18). For $\kappa_E \geq \varkappa$, for some fixed $\varkappa > 0$, we will argue in Subsection VI B 1, that $|\zeta_0(z) - \widetilde{\zeta}_0(z)| \ll (N\eta)^{-1}$ and we may thus replace ζ_0 in (2.25) by $\widetilde{\zeta}_0$ without changing the bounds; cf. Lemma 6.3.

2. Eigenvector delocalization

Next, let $\mu_1 \leq \ldots \leq \mu_N$ denote the eigenvalues of $H = \lambda V + W$, and let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$ denote the associated eigenvectors. We use the notation $\boldsymbol{u}_\alpha = (u_\alpha(i))_{i=1}^N$ for the vector components. All eigenvectors are ℓ^2-normalized. The next theorem asserts that, with high probability, all eigenvectors of $H = \lambda V + W$ are completely delocalized:

Theorem 2.15. [Eigenvector delocalization] *Assume that $H = \lambda V + W$ satisfies the assumptions in Definition 2.1 and Assumption 2.3 or 2.8. Then there is a constant $\nu > 0$, depending on θ and C_0 in (2.4), λ_0 in (2.15), E_0 in (2.16), A_0 in (2.17), and the measure μ, such that, for any ξ satisfying (2.14), we have*

$$\max_{1\leq\alpha\leq N} \max_{1\leq i\leq N} |u_\alpha(i)| \leq \frac{(\varphi_N)^{4\xi}}{\sqrt{N}},$$

with (ξ, ν)-high probability.

Remark 2.16. In case the entries of $V = (v_i)$ are independent Gaussian random variables, the situation is more subtle. For any finite E_0, there exists a constant c_{E_0}, independent of N, and a constant ν, depending on A_0, E_0, θ, and C_0, such that, for any ξ satisfying (2.14), the following holds: Let $\alpha \in \{1, \ldots, N\}$ be such that the eigenvalue μ_α satisfies $|\mu_\alpha| \leq E_0$. Then we have

$$\max_{1\leq i\leq N} |u_\alpha(i)| \leq c_{E_0} \frac{(\varphi_N)^{4\xi}}{\sqrt{N}}, \tag{2.27}$$

with (ξ, ν)-high probability. However, $c_{E_0} \to \infty$ and $\nu \to 0$, as $E_0 \to \infty$.

3. Density of states

Next, we state our main results about the *local density of states* of $H = \lambda V + W$. For $E_1 < E_2$, we define the counting functions

$$\mathfrak{n}(E_1, E_2) := \frac{1}{N}|\{\alpha : E_1 < \mu_\alpha \leq E_2\}|, \qquad n_{fc}(E_1, E_2) := \int_{E_1}^{E_2} \mathrm{d}x\, \rho_{fc}(x), \tag{2.28}$$

where we denote by ρ_{fc} the density of the free convolution measure μ_{fc}.

Theorem 2.17. [Local density of states] *Let $H = \lambda V + W$, where W satisfies the assumptions in Definition 2.1 and λV satisfies Assumption 2.3 or 2.8. Let ξ satisfy (2.17). Then there are constants $\nu > 0$ and c, depending on θ and C_0 in (2.4), λ_0 in (2.15), E_0 in (2.16), A_0 in (2.17), and the measure μ, such that, for $L \geq 40\xi$, the following holds: For any E_1, E_2, satisfying $-E_0 \leq E_1 < E_2 \leq E_0$, $E_2 > E_1 + (\varphi_N)^L N^{-1}$, and any $\lambda \in \mathcal{D}_{\lambda_0}$, the estimate*

$$|\mathfrak{n}(E_1, E_2) - n_{fc}(E_1, E_2)| \leq (\varphi_N)^{c\xi}\left(\frac{1}{N} + \frac{\lambda(E_2 - E_1)}{\sqrt{\kappa + (E_2 - E_1)}}\frac{1}{\sqrt{N}}\right), \tag{2.29}$$

holds with (ξ, ν)-high probability.

Moreover, let $\varkappa > 0$. Then, there exists a constant $C_\varkappa$, depending only on $\varkappa$, such that, for any E_1, E_2, satisfying $L_1 + \varkappa \le E_1 < E_2 \le L_2 - \varkappa$, $E_2 > E_1 + (\varphi_N)^L N^{-1}$, and any $\lambda \in \mathcal{D}_{\lambda_0}$, the estimate

$$|\mathfrak{n}(E_1, E_2) - n_{fc}(E_1, E_2)| \le C_\varkappa (\varphi_N)^{c\xi} \left(\frac{1}{N} + \frac{\lambda^2 (E_2 - E_1)^2}{\sqrt{N}} \right) , \tag{2.30}$$

holds with (ξ, ν)-high probability.

We remark, however, that the estimate in (2.30) deteriorates at the edge: $C_\varkappa \to \infty$, as $\varkappa \to 0$.

4. Rigidity of eigenvalue spacing

Recall that we denote by $\mu_1 \le \mu_2 \le \ldots \le \mu_N$, the eigenvalues of $H = \lambda V + W$. The estimates on the density of states in Theorem 2.17 imply the following result on the rigidity of eigenvalue spacing, which establishes the relation between $|\mu_i - \mu_j|$ and $|i - j|$.

Theorem 2.18. *Assume that $H = \lambda V + W$ satisfies the assumptions in Definition 2.1 and Assumption 2.3 or 2.8. Consider $\mu_i < \mu_j$, with $i \ge \epsilon N$ and $j \le (1 - \epsilon)N$, for some constant $\epsilon > 0$. Assume that $|i - j| \ge (\varphi_N)^{C'\xi}$, for some constant $C' > C$, where C is the constant in (2.30). Then, there exist constants C_1, C_2, depending only on ϵ, $\max\{\rho_{fc}(x) \colon \mu_i \le x \le \mu_j\}$ and $\min\{\rho_{fc}(x) \colon \mu_i \le x \le \mu_j\}$, such that the following holds: For ξ and $\nu > 0$, as in Theorem 2.17, the estimate*

$$C_1 \frac{|i - j|}{N} \le |\mu_i - \mu_j| \le C_2 \frac{|i - j|}{N} ,$$

holds with (ξ, ν)-high probability, for all $\lambda \in \mathcal{D}_{\lambda_0}$. If we assume further that $|i - j| \le (\varphi_N)^{c\xi} N^{1/2}$, for some constant $c > 0$, then there exists a constant K such that

$$\left| |\mu_i - \mu_j| - \frac{|i - j|}{N \rho_{fc}(\mu_i)} \right| \le (\varphi_N)^{K\xi} \frac{1}{N} , \tag{2.31}$$

with (ξ, ν)-high probability, for all $\lambda \in \mathcal{D}_{\lambda_0}$.

Remark 2.19. The estimate in (2.31) can be extended to $|i - j| \le (\varphi_N)^{-c\xi} N^{3/4}$ in the following sense: There exists a constant K such that, for some $\mu_i' \in [\mu_i, \mu_j]$,

$$\left| |\mu_i - \mu_j| - \frac{|i - j|}{N \rho_{fc}(\mu_i')} \right| \le (\varphi_N)^{K\xi} \frac{1}{N}, \tag{2.32}$$

with (ξ, ν)-high probability. The estimate (2.32) easily follows from the proof of Theorem 2.18.

5. Integrated density of states and rigidity of eigenvalues

Define the *integrated density of states* by

$$\mathfrak{n}(E) := \frac{1}{N} |\{\alpha \, : \, \mu_\alpha \le E\}| \, .$$

Similarly, we set

$$n_{fc}(E) := \int_{-\infty}^{E} \rho_{fc}(x) \, \mathrm{d}x \, ,$$

where ρ_{fc} denotes the density of the free convolution measure μ_{fc}. Then we have the following result:

Theorem 2.20. *Let $H = \lambda V + W$, where W satisfies the assumptions in Definition 2.1 and λV satisfies Assumption 2.3 or 2.8. Let ξ satisfy (2.17). Then there are constants $\nu > 0$ and c, depending on the constants θ and C_0 in (2.4), λ_0 in (2.15), E_0 in (2.16), A_0 in (2.17), and the measure μ, such* *that the event*

$$\bigcap_{\substack{E\in[-E_0,E_0]\\ \lambda\in\mathcal{D}_{\lambda_0}}} \left\{ |\mathfrak{n}(E) - n_{fc}(E)| \le (\varphi_N)^{c\xi}\left(\frac{1}{N} + \frac{\lambda^{3/2}}{N^{3/4}} + \frac{\lambda}{N^{5/6}} + \frac{\lambda\sqrt{\kappa_E}}{\sqrt{N}}\right)\right\},$$

has (ξ,ν)*-high probability.*

Our last result concerns the rigidity of the eigenvalue location. We define the "classical" location of the eigenvalue μ_α of H, γ_α, by

$$\int_{-\infty}^{\gamma_\alpha} \rho_{fc}(x)\mathrm{d}x = \frac{\alpha}{N}, \qquad \alpha\in\{1,\ldots,N\}, \tag{2.33}$$

where ρ_{fc} is the density of the free convolution measure μ_{fc}.

Theorem 2.21. *Let* $H = \lambda V + W$*, where* W *satisfies the assumptions in Definition 2.1 and* λV *satisfies Assumption 2.3 or 2.8. Let* ξ *satisfy* (2.17)*. Then there are constants* $\nu > 0$ *and* c*, depending on the constants* θ *and* C_0 *in* (2.4), λ_0 *in* (2.15), E_0 *in* (2.16), A_0 *in* (2.17)*, and the measure* μ*, such that*

$$|\mu_\alpha - \gamma_\alpha| \le (\varphi_N)^{c\xi}\left(N^{-2/3}\left[\widehat{\alpha}^{-1/3} + \mathbb{1}\left(\widehat{\alpha}\le(\varphi_N)^{c\xi}(1+\lambda^{3/2}N^{1/4})\right)\right] + \lambda^2 N^{-1/3}\widehat{\alpha}^{-2/3} + \lambda N^{-1/2}\right), \tag{2.34}$$

with (ξ,ν)*-high probability, for all* $\lambda\in\mathcal{D}_{\lambda_0}$*, where we have abbreviated* $\widehat{\alpha} := \min\{\alpha, N-\alpha\}$.

Remark 2.22. Let us compare this rigidity result with the corresponding rigidity result for Wigner matrices ($\lambda = 0$): In the bulk of the spectrum, where $\alpha\sim N$, we obtain from (2.34),

$$|\mu_\alpha - \gamma_\alpha| \le (\varphi_N)^{c\xi}\left(\frac{1+C\lambda}{N} + \frac{\lambda}{\sqrt{N}}\right), \tag{2.35}$$

with (ξ,ν)-high probability. Thus, for $\lambda\neq 0$, the leading corrections in the rigidity estimate arise from fluctuations in the diagonal matrix V, and the eigenvalues do not satisfy as strong a rigidity estimate for their locations as in the Wigner case; see, e.g., Refs. 16, 17, 19, and 20. However, the eigenvalues satisfy a strong rigidity estimate on intermediate scales for their relative position or their spacing; see Theorem 2.18 above. For $\lambda = 0$, the rigidity of eigenvalue spacing is an immediate consequence of the rigidity of eigenvalue location.

Remark 2.23. For $\lambda = 0$, the model is known to exhibit bulk universality, (see, e.g., Refs. 19–22, and 23), and it is easy to see that bulk universality holds true for the choice $\lambda = CN^{-\delta}$, with $\delta\ge 1/2$. In case W is a GUE matrix, bulk universality has been proved to hold for λ order one; see Ref. 35.

A corner stone in the proof of bulk universality for Wigner matrices in, e.g., Ref. 24, is the rigidity estimate

$$\frac{1}{N}\sum_{\alpha=1}^{N}\mathbb{E}|\mu_\alpha-\gamma_\alpha|^2 \le CN^{-1-2\mathfrak{a}}, \qquad (\lambda = 0), \tag{2.36}$$

for some $\mathfrak{a} > 0$, where (μ_α) are the eigenvalues of W and (γ_α) are the classical locations of the eigenvalues (with respect to the standard semicircle law). For $\lambda\neq 0$, one can show that (2.34) implies

$$\frac{1}{N}\sum_{\alpha=1}^{N}\mathbb{E}|\mu_\alpha-\gamma_\alpha|^2 \le C(\varphi_N)^{c\xi}\lambda^2 N^{-1} + CN^{-1-2\mathfrak{a}}, \tag{2.37}$$

for some constants C, c, and $\mathfrak{a} > 0$, where now (γ_α) denote the classical locations with respect to the deformed semicircle law. For $\lambda = CN^{-\delta}$, with $\delta > 0$, it is thus conceivable that one can prove bulk universality following the lines of, e.g., Ref. 24, using local ergodicity of Dyson Brownian motion and Green function comparison. Note that for $\lambda \ll 1$, the limiting eigenvalue distribution of H is the semicircle law. For λ order one, the proof of Ref. 24 seems not to be applicable without greater modifications.

Remark 2.24. When V is a deterministic instead of a random diagonal matrix, one can still prove the results in this paper, provided that the limiting density of the eigenvalues of V satisfies the required assumptions. Moreover, when V is symmetric and W is GUE or GOE, it is possible to prove the same results after diagonalizing $\lambda V + W$, due to the invariance of GUE (GOE) under unitary (orthogonal) conjugation; some knowledge on the convergence of empirical eigenvalue distribution is required, e.g., the analogue statement to (3.35) below. The extension to the more general case where V is non-diagonal and W is a general Wigner matrix requires a further investigation; the usual Lindeberg replacement strategy using moment matching conditions (e.g., Ref. 37) may not be sufficient, because the statements in this paper are stronger than those one gets from the moment matching conditions in the sense that the former hold with high probability while the latter hold after taking expectation.

D. Outline of proofs

In this subsection, we briefly outline the proofs of the main results.

In Sec. III, we derive a *weak local deformed semicircle law*, Theorem 3.1. Following the lines of Ref. 12 we derive a *(weak) self-consistent equation* for $m - m_{fc}$. The main differences with the Wigner case are: (1) The limiting eigenvalue distribution follows the deformed semicircle law instead of the semicircle law. For Wigner matrices, the stability of this self-consistent equation is obtained by elementary calculus using the exact form of m_{sc}. For the deformed model, the stability of the self-consistent equation (3.33) follows from the (μ- and λ-dependent) stability estimate (3.5). (2) When taking the normalized trace of the Green function, we average over the random variables (v_i); see (3.34). Eventually, we replace this average by its expected value. This replacement results in an error term that is, according to the CLT, of order $\lambda N^{-1/2}$ (up to logarithmic corrections); see (3.35). These fluctuations should be compared with the other error terms. Among those error terms the dominating one, the Z_i defined in (3.15), is of order $(N\eta)^{-1/2}$. Since $z \in \mathcal{D}_L$, we can combine the error terms and our estimates in Theorem 3.1 are the same as the corresponding estimates in Ref. 12. However, due to the order one diagonal entries (v_i), G_{ii} is not self-averaging. In Sec. III, we also prove Theorem 2.15: Delocalization of eigenvectors can be obtained from Ref. 22 as a corollary of the local deformed semicircle law, Theorem 2.10.

In Sec. IV, we prove the fluctuation average bound $|\frac{1}{N}\sum_i Z_i| \lesssim (N\eta)^{-1}$; see Lemma 4.1. The proof of this lemma is inspired by Erdős *et al.*,[14] where more general fluctuation averages are considered for Wigner and random band matrices. Our treatment is in so far different as fluctuations of (diagonal elements of) Green functions are not self-averaging. Having established an optimal error bound on the average of (Z_i), we have to keep track of the (η-independent) $\lambda N^{-1/2}$ fluctuation from the CLT alluded to above ($(N\eta)^{-1}$ cannot be compared with $\lambda N^{-1/2}$ on $\mathcal{D}_L$). As in Ref. 22, we obtain a *(strong) self-consistent equation*, Eq. (4.35), whose stability analysis yields a proof of the strong local deformed semicircle law.

In Sec. V, we prove Theorem 2.12. In (5.2), we define a random variable ζ_0, depending only on (v_i), such that $|m(z) - m_{fc}(z) - \zeta_0(z)|$ is minimized. This can be achieved by defining ζ_0 as the solution to the strong self-consistent equation with all error terms but the ones depending only on (v_i) discarded. Defining ζ_0 in this way yields an optimal bound on $|m(z) - m_{fc} - \zeta_0(z)|$ away from the spectral edges, or more precisely, for energies E satisfying $\kappa_E \gtrsim N^{-1/4}$.

In Sec. VI A, we establish, using the Helffer-Sjöstrand formula in the argument in Ref. 15, estimates on the density of states and the rigidity of eigenvalues. Using the results obtained in Sec. V on ζ_0, we obtain in Sec. VI B estimates on the rigidity of the eigenvalue spacing on intermediate scales in the bulk of the spectrum. More precisely, in the bulk of the spectrum we approximate $\zeta_0(z)$ by $\widetilde{\zeta}_0(z)$; see Remark 2.14 above for a definition. In Lemma 6.4, we show that the z-dependent random variables $\widetilde{\zeta}_0(E + \mathrm{i}\eta)$ is, for fixed η, a slowly varying function of E in the bulk of the spectrum. This in turn can be used to get more precise estimates on the density of states in the bulk of the spectrum; see (2.30). In Sec. VI C, we prove Theorem 2.18, based on results on the local density of states. Using once more the random variables $\widetilde{\zeta}_0$, one can obtain stronger estimates on the eigenvalue spacing in the bulk of the spectrum; see (2.32). Following Ref. 12, we prove Theorems 2.20 and 2.21 in Sec. VI D.

In the Appendix, we discuss properties of the deformed semicircular law, μ_{fc}, in particular, we prove Lemma 2.4 ($\lambda \leq 1$), Lemma 2.7 ($\lambda > 1$). In a slightly different setting Lemma 2.4 has been proven in Ref. 35, but we recall parts the proof, since it is used in the proof of Lemma 2.7. The first part of Lemma 2.7, can be proved in a similar way, but we have to impose some stronger assumptions on μ; cf. (A6). The second part of Lemma 2.7 shows that the deformed semicircle law may not have a square root behaviour at the edge. Our proof is based on elementary estimates, but we expect that the statement can be proven using methods of complex analysis.

III. WEAK DEFORMED SEMICIRCLE LAW

In this section, we prove a weaker form of the deformed semicircle law. This *weak deformed semicircle law* will be used to prove the *strong deformed law* in Theorem 2.10. Moreover, complete delocalization of eigenvectors is a direct consequence of the weak law stated in the next theorem.

Theorem 3.1. [Weak deformed semicircle law] *Let $H = \lambda V + W$ satisfy the assumptions in Definition 2.1 and Assumption 2.3 or 2.8. Then there are constants C, $\nu > 0$, depending on the constants θ and C_0 in (2.4), λ_0 in (2.15), E_0 in (2.16), A_0 in (2.17), and the measure μ, such that, for*

$$a_0 \leq \xi \leq A_0 \log\log N \,, \qquad L \geq 12\xi,$$

the event

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ \max_{i \neq j} |G_{ij}(z)| \leq C \frac{(\varphi_N)^\xi}{\sqrt{N\eta}} \right\}, \tag{3.1}$$

has (ξ, ν)-high probability.

Denote by $\mathbb{E}^{v_i}$, the expectation with respect to the random variable v_i, $i \in \{1, \ldots, N\}$. Then the event

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ \max_{1 \leq i \leq N} |\mathbb{E}^{v_i} G_{ii}(z) - m(z)| \leq C(\varphi_N)^\xi \left(\frac{\lambda}{\sqrt{N}} + \frac{1}{(N\eta)^{1/3}} \right) \right\}, \tag{3.2}$$

has (ξ, ν)-high probability.

Moreover, we have the weak local deformed semicircle law: The event

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ |m(z) - m_{fc}(z)| \leq C \frac{(\varphi_N)^\xi}{(N\eta)^{1/3}} \right\}, \tag{3.3}$$

has (ξ, ν)-high probability.

The rest of the section is devoted to the proof of Theorems 3.1 and 2.15. The proof follows closely the proof for Wigner matrices; see Refs. 12 and 24. We will always assume that W satisfies the assumptions in Definition 2.1 and that λV satisfies Assumption 2.3 or 2.8.

A. Preliminaries

1. Some properties of μ_{fc} and m_{fc}

The next lemma collects some useful properties of m_{fc} under Assumptions 2.3 or 2.8.

Lemma 3.2. There exist $L_1 < L_2$ such that the free convolution measure μ_{fc} has support $[L_1, L_2]$. For all $z = E + \mathrm{i}\eta \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, the Stieltjes transform, m_{fc}, of μ_{fc} has the following properties:

i. *Let $\kappa := \min\{|E - L_1|, |E - L_2|\}$, then*

$$\operatorname{Im} m_{fc}(z) \sim \begin{cases} \sqrt{\kappa + \eta}\,, & E \in [L_1, L_2]\,, \\ \frac{\eta}{\sqrt{\kappa+\eta}}\,, & E \in [L_1, L_2]^c\,. \end{cases} \tag{3.4}$$

ii. *There exist constants C, $c > 0$, depending on μ, E_0 and λ_0, such that*

$$c \le |\lambda - z - m_{fc}(z)| \le C. \tag{3.5}$$

We refer to (3.5) as "stability bound" and remark that a similar condition has already been used in Ref. 36. The proof of Lemma 3.2 is given in the Appendix.

2. Minors

Definition 3.3. Let $\mathbb{T} \subset \{1, \ldots, N\}$. Then we define $H^{(\mathbb{T})}$ as the $(N - |\mathbb{T}|) \times (N - |\mathbb{T}|)$ minor of H obtained by removing all columns and rows of H indexed by $i \in \mathbb{T}$. Note that we do not change the names of the indices of H when defining $H^{(\mathbb{T})}$. More specifically, we define an operation π_i, $i \in \{1, \ldots, N\}$, on the probability space by

$$(\pi_i(H))_{kl} := \mathbb{1}(k \neq i)\mathbb{1}(l \neq i)h_{kl}.$$

Then, for $\mathbb{T} \subset \{1, \ldots, N\}$, we set $\pi_{\mathbb{T}} := \prod_{i\in\mathbb{T}} \pi_i$ and define

$$H^{(\mathbb{T})} := ((\pi_{\mathbb{T}}(H)_{ij})_{i,j\notin\mathbb{T}}.$$

The Green functions $G^{(\mathbb{T})}$, are defined in an obvious way using $H^{(\mathbb{T})}$. Moreover, we use the shorthand notation

$$\sum_i^{(\mathbb{T})} := \sum_{\substack{i=1 \\ i\notin\mathbb{T}}}^{N}\,,$$

and abbreviate $(i) = (\{i\})$ and, similarly, $(\mathbb{T}i) = (\mathbb{T} \cup \{i\})$. Finally, we set

$$m^{(\mathbb{T})} := \frac{1}{N}\sum_i^{(\mathbb{T})} G_{ii}^{(\mathbb{T})}.$$

Here, we use the normalization N^{-1}, instead $(N - |\mathbb{T}|)^{-1}$, since it is more convenient for our computations.

3. Resolvent identities

The next lemma collects the main identities between resolvent matrix elements of H and $H^{(\mathbb{T})}$.

Lemma 3.4. Let H be an $N \times N$ matrix. Consider the Green function $G(z) \equiv G := (H - z)^{-1}$, $z \in \mathbb{C}^+$. Then, for i, j, $k \in \{1, \ldots, N\}$, the following identities hold:

- *Schur complement/Feshbach formula: For any i,*

$$G_{ii} = \frac{1}{h_{ii} - z - \sum_{m,n}^{(i)} h_{im}G_{mn}^{(i)}h_{ni}}. \tag{3.6}$$

- *For $i \neq j$,*

$$G_{ij} = -G_{ii}G_{jj}^{(i)}(h_{ij} - \sum_{m,n}^{(ij)} h_{im}G_{mn}^{(ij)}h_{nj}). \tag{3.7}$$

- *For* $i, j \neq k$,

$$G_{ij} = G_{ij}^{(k)} + \frac{G_{ik}G_{kj}}{G_{kk}}. \tag{3.8}$$

- *Ward identity: For any* i,

$$\sum_{n=1}^{N} |G_{in}|^2 = \frac{1}{\eta} \operatorname{Im} G_{ii}, \tag{3.9}$$

where $\eta = \operatorname{Im} z$.

For a proof we refer to, e.g., Ref. 12.

4. Large deviation estimates

We collect here some useful large deviation estimates for random variables with slowly decaying moments.

Lemma 3.5. Let (a_i) *and* (b_i) *be centered and independent complex random variables with variance* σ^2 *and having subexponential decay*

$$\mathbb{P}\left(|a_i| \geq x\sigma\right) \leq C_0\, \mathrm{e}^{-x^{1/\theta}}, \qquad \mathbb{P}\left(|b_i| \geq x\sigma\right) \leq C_0\, \mathrm{e}^{-x^{1/\theta}},$$

for some positive constant C_0 *and* $\theta > 1$. *Let* $A_i \in \mathbb{C}$ *and* $B_{ij} \in \mathbb{C}$. *Then there exist constants* $a_0 > 1$, $A_0 \geq 10$ *and* $C \geq 1$, *depending on* θ *and* C_0, *such that for* $a_0 \leq \xi \leq A_0 \log\log N$, *and* $\varphi_N = (\log N)^C$,

$$\mathbb{P}\left(\left| \sum_{i=1}^{N} A_i a_i \right| \geq (\varphi_N)^{\xi} \sigma \left(\sum_{i=1}^{N} |A_i|^2 \right)^{1/2} \right) \leq \mathrm{e}^{-(\log N)^{\xi}}, \tag{3.10}$$

$$\mathbb{P}\left(\left| \sum_{i=1}^{N} \overline{a}_i B_{ii} a_i - \sum_{i=1}^{N} \sigma^2 B_{ii} \right| \geq (\varphi_N)^{\xi} \sigma^2 \left(\sum_{i=1}^{N} |B_{ii}|^2 \right)^{1/2} \right) \leq \mathrm{e}^{-(\log N)^{\xi}}, \tag{3.11}$$

$$\mathbb{P}\left(\left| \sum_{i\neq j}^{N} \overline{a}_i B_{ij} a_j \right| \geq (\varphi_N)^{\xi} \sigma^2 \left(\sum_{i\neq j} |B_{ij}|^2 \right)^{1/2} \right) \leq \mathrm{e}^{-(\log N)^{\xi}}, \tag{3.12}$$

$$\mathbb{P}\left(\left| \sum_{i,j=1}^{N} \overline{a}_i B_{ij} b_j \right| \geq (\varphi_N)^{\xi} \sigma^2 \left(\sum_{i\neq j} |B_{ij}|^2 \right)^{1/2} \right) \leq \mathrm{e}^{-(\log N)^{\xi}}, \tag{3.13}$$

for N *sufficiently large*.

We refer to Ref. 22 for a proof.

5. Schur complement formula

The proof of Theorem 3.1 starts with Schur's formula

$$G_{ii} = \frac{1}{h_{ii} - z - \sum_{k,l}^{(i)} h_{ik} G_{kl}^{(i)} h_{li}}, \qquad z \in \mathcal{D}_L, \tag{3.14}$$

where, for brevity, $G_{ij} \equiv G_{ij}(z)$. Define $\mathbb{E}_i$ to be the partial expectation with respect to the ith-column/row of W and set

$$Z_i := (\mathbb{1} - \mathbb{E}_i) \sum_{k,l}^{(i)} h_{ik} G_{kl}^{(i)} h_{li} = \sum_{k,l}^{(i)} (h_{ik} G_{kl}^{(i)} h_{li} - \frac{1}{N} \delta_{kl} G_{kl}^{(i)})$$
$$= \sum_{k}^{(i)} (|w_{ik}|^2 - \frac{1}{N}) G_{kk}^{(i)} + \sum_{k \neq l}^{(i)} w_{ik} G_{kl}^{(i)} w_{li}, \tag{3.15}$$

here we used $h_{ik} = w_{ik} + \lambda \delta_{ik} v_i$. For a family of random variables $(F_1, \ldots, F_N)$ we introduce the notation

$$[F] := \frac{1}{N} \sum_{i=1}^{N} F_i. \tag{3.16}$$

Recalling the definition $m^{(i)} = \frac{1}{N} \operatorname{Tr} G^{(i)} = \frac{1}{N} \sum_k^{(i)} G_{kk}^{(i)}$, we obtain from Eqs. (3.14) and (3.15)

$$G_{ii} = \frac{1}{\lambda v_i + w_{ii} - z - m^{(i)} - Z_i}$$
$$= \frac{1}{\lambda v_i - z - m_{fc} - ([\mathrm{v}] - \mathcal{Y}_i)}, \tag{3.17}$$

where

$$\mathrm{v}_i := G_{ii} - m_{fc}, \quad \mathcal{Y}_i := w_{ii} - Z_i - (m^{(i)} - m). \tag{3.18}$$

Note the difference between v_i and w_{ii}: Since we assumed that the (rescaled) entries of W have subexponential decay, we have

$$|w_{ij}| \leq C \frac{(\varphi_N)^{\xi}}{\sqrt{N}}, \tag{3.19}$$

with (ξ, ν)-high probability, whereas $v_i = \mathcal{O}(1)$, almost surely.

Lemma 3.6. There is a constant C such that, for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$ and $1 \leq i \leq N$, we have

$$|m(z) - m^{(i)}(z)| \leq \frac{C}{N\eta}. \tag{3.20}$$

Proof. The claim follows from Cauchy's interlacing property of eigenvalues of H and its minor $H^{(i)}$. For a detailed proof we refer to Ref. 11. □

B. *A priori* estimates on the domain $\Omega(z)$

Define the z-dependent control quantities

$$\Lambda_o := \max_{i \neq j} |G_{ij}|, \qquad \Lambda_d := \max_{i} |G_{ii}|, \qquad \Lambda := |m - m_{fc}|. \tag{3.21}$$

Note that these quantities also depend on λ, but we do not display this dependence, since, as we shall see, uniformity in λ can always be achieved on the domain $\mathcal{D}_{\lambda_0}$ using the stability bound (3.5).

For $z \in \mathcal{D}_L$, we define an event $\Omega(z)$ by

$$\Omega(z) := \{\Lambda_o \leq (\varphi_N)^{-2\xi}\} \cap \{\Lambda \leq (\varphi_N)^{-2\xi}\}. \tag{3.22}$$

First, we check that we can bound the matrix elements of the Green function of the minor $H^{(i)}$ in terms of the matrix elements of the Green function of H.

Lemma 3.7. Let $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. Then there are constants C, $c > 0$ such that, for any $\mathbb{T} \subset \{1, \ldots, N\}$ with $|\mathbb{T}| \leq 10$, the following statements hold with high probability on $\Omega(z)$:

i. For any $i \notin \mathbb{T}$,

$$c \le |G_{ii}^{(\mathbb{T})}| \le C. \tag{3.23}$$

ii. For any $i, j \notin \mathbb{T}$, $i \neq j$,

$$c\Lambda_o \le |G_{ij}^{(\mathbb{T})}| \le C\Lambda_o. \tag{3.24}$$

iii.

$$|m - m^{(\mathbb{T})}| \le C\Lambda_o^2. \tag{3.25}$$

Moreover, the constants C and c can be chosen uniformly in $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$.

Proof. Let $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. We will successively use (3.8), i.e.,

$$G_{ij} - G_{ij}^{(k)} = \frac{G_{ik}G_{kj}}{G_{kk}}. \tag{3.26}$$

Since we are working on $\Omega(z)$ we have $|G_{ij}| \le \Lambda_o \le (\varphi_N)^{-2\xi}$, for $i \neq j$. Next, Eq. (3.17) yields

$$\left|\frac{1}{G_{ii}}\right| = |z + m^{(i)} - \lambda v_i - w_{ii} - Z_i|.$$

By the large deviation estimates (3.11) and (3.12), the Ward identity (3.9) and Inequality (3.20) we have

$$\begin{aligned}|Z_i| \le C(\varphi_N)^{\xi}\left(\frac{1}{N^2}\sum_{k,l}^{(i)}|G_{kl}^{(i)}|^2\right)^{1/2} &\le C(\varphi_N)^{\xi}\sqrt{\frac{\operatorname{Im} m^{(i)}}{N\eta}}\\ &\le C(\varphi_N)^{\xi}\sqrt{\frac{\Lambda + \operatorname{Im} m_{fc}}{N\eta}} + C(\varphi_N)^{\xi}\frac{1}{N\eta},\end{aligned} \tag{3.27}$$

with high probability on $\Omega(z)$. Since $\Lambda \le (\varphi_N)^{-2\xi}$ on $\Omega(z)$ and since $N\eta \ge (\varphi_N)^{12\xi}$, we conclude that $|Z_i| = o(1)$, with high probability on $\Omega(z)$, for $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Finally, since

$$\left|m^{(i)}\right| = \left|m_{fc}\right| + \mathcal{O}\left(\frac{1}{N\eta} + (\varphi_N)^{-2\xi}\right),$$

on $\Omega(z)$, by (3.20), we find

$$\left|\frac{1}{G_{ii}}\right| = |\lambda v_i - z - m_{fc}| + o(1),$$

with high probability on $\Omega(z)$. This, together with the stability bound (3.5), proves the lower and upper bound on G_{ii}. Note that by (3.5), we can choose the upper and lower bound on G_{ii} to be uniform in $\lambda \in \mathcal{D}_{\lambda_0}$, $z \in \mathcal{D}_L$.

Statements *i–iii* now follow by iterating (3.26). □

Next, we define the control parameter $\Psi(z)$, for $z \in \mathcal{D}_L$, by

$$\Psi(z) := (\varphi_N)^{\xi}\sqrt{\frac{\Lambda + \operatorname{Im} m_{fc}}{N\eta}}, \tag{3.28}$$

where $\Lambda = |m - m_{fc}|$. Again, we suppress the λ-dependence of $\Psi(z)$ from our notation. We will use $\Psi \equiv \Psi(z)$ to bound various quantities in the following:

Lemma 3.8. Let $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. Then there is a constant C such that we have with (ξ, ν)-high probability on $\Omega(z)$:

$$\Lambda_o \le C\Psi, \tag{3.29}$$

$$\max_i |Z_i| \le C\Psi, \tag{3.30}$$

$$\max_i |\mathcal{Y}_i| \le C\Psi. \tag{3.31}$$

The constant C can be chosen uniformly in $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$.

Proof. We prove (3.29). Let $i \neq j$, then by Eq. (3.7), the large deviation estimates of Lemma 3.5 and Inequality (3.19),

$$|G_{ij}| \le C(|w_{ij}| + \sum_{k,l}^{(ij)} |w_{ik} G_{kl}^{(ij)} w_{lj}|) \le C(\varphi_N)^{\xi} \left(\frac{1}{\sqrt{N}} + \sqrt{\frac{1}{N^2} \sum_{k,l}^{(ij)} |G_{kl}^{(ij)}|^2} \right)$$
$$= C(\varphi_N)^{\xi} \left(\frac{1}{\sqrt{N}} + \sqrt{\frac{\operatorname{Im} m^{(ij)}}{N\eta}} \right),$$

with high probability, where we used in the last step the Ward identity (3.9). Since $|m^{(ij)} - m| \le C\Lambda_o^2$, by Lemma 3.7, we get

$$|G_{ij}| \le C\left(\frac{(\varphi_N)^{\xi}}{\sqrt{N}} + \Psi(z) \right) + C\frac{(\varphi_N)^{\xi}}{\sqrt{N\eta}} \Lambda_o,$$

with high probability. Since $\operatorname{Im} m_{fc}(z) \ge C\eta$, by (3.4), we can absorb the term $(\varphi_N)^{\xi} N^{-1/2}$ into the term $\Psi(z)$. Taking the maximum over $i \neq j$, inequality (3.29) follows. The proofs for Z_i and $\mathcal{Y}_i$ are similar. □

C. Derivation of the weak self-consistent equation

We now put Eq. (3.17) into a form which admits an analysis of the average of the diagonal resolvent entries. For $n \in \mathbb{N}$, define

$$R_n(z) := \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^n}, \qquad z \in \mathcal{D}_L, \quad \lambda \in \mathcal{D}_{\lambda_0}. \tag{3.32}$$

For any n, R_n is bounded uniformly in z and λ. This follows from the stability bound (3.5). Note the special case $R_1 = m_{fc}$. Recall the definitions $[\mathrm{v}] = \frac{1}{N}\sum_i G_{ii} - m_{fc}$ and $|\Lambda| = |m - m_{fc}|$.

Lemma 3.9. [Weak self-consistent equation] *There is a constant C such that, for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, we have on $\Omega(z)$ with (ξ, ν)-high probability*

$$\left|(1 - R_2)[\mathrm{v}] - R_3[\mathrm{v}]^2\right| \le C\Psi + C\frac{\Lambda^2}{\log N}. \tag{3.33}$$

Proof. Since $|\lambda v_i - z - m_{fc}|$ is bounded below by (3.5), we can expand Eq. (3.17) to second order in $([\mathrm{v}] - \mathcal{Y}_i)$,

$$\frac{1}{N}\sum_{i=1}^N G_{ii} = \frac{1}{N}\sum_{i=1}^N \frac{1}{\lambda v_i - z - m_{fc}} + \frac{1}{N}\sum_{i=1}^N \frac{1}{(\lambda v_i - z - m_{fc})^2}([\mathrm{v}] - \mathcal{Y}_i)$$
$$+ \frac{1}{N}\sum_{i=1}^N \frac{1}{(\lambda v_i - z - m_{fc})^3}([\mathrm{v}] - \mathcal{Y}_i)^2 + \mathcal{O}(\Lambda^3) + \mathcal{O}(\max_i |\mathcal{Y}_i|^3), \tag{3.34}$$

where $\mathcal{Y}_i = w_{ii} - Z_i - (m^{(i)} - m)$; see Eq. (3.18).

Next, we use the "law of large numbers" to replace the averages in the first two terms on the right side of (3.34) by their expectation: It follows from the stability bound in (3.5) that the family of functions $g_i : \mathcal{D}_{\lambda_0} \times \mathcal{D}_L \to \mathbb{C}^+$, $(\lambda, z) \mapsto (\lambda v_i - z - m_{fc}(z))^{-1}$ are jointly Lipschitz continuous with a constant depending only on E_0, λ_0, and μ. Since the (v_i) are i.i.d. random variables, McDiarmid's inequality implies that, for $\lambda \in \mathcal{D}_{\lambda_0}$, $z \in \mathcal{D}_L$, $n = 1, 2, 3$, there is a constant C',

$$\left| \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(\lambda v_i - z - m_{fc})^n} - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc})^n} \right| \le C' \frac{\lambda(\varphi_N)^{\xi}}{\sqrt{N}}, \tag{3.35}$$

with (ξ, ν)-high probability. Uniformity in λ, z, and ν can be established by a lattice argument: Choose a lattice $\mathcal{L} \in \mathcal{D}_{\lambda_0} \times \mathcal{D}_L$, with $|\mathcal{L}| \le CN^4$, such that for any $(\lambda, z) \in \mathcal{D}_{\lambda_0} \times \mathcal{D}_L$ there is $(\lambda', z') \in \mathcal{L}$, with $|z - z'| \le N^{-2}$ and $|\lambda - \lambda'| \le N^{-2}$. Then (3.35) holds for all $(\lambda, z) \in \mathcal{L}$ for some sufficiently large C' and some sufficiently small $\nu > 0$. Using the joint Lipschitz continuity of (g_i), we conclude that there is a constant $C \ge C'$ such that the event,

$$\bigcap_{n=1,2,3} \bigcap_{(\lambda, z) \in \mathcal{D}_{\lambda_0} \times \mathcal{D}_L} \left\{ \left| \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(\lambda v_i - z - m_{fc})^n} - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc})^n} \right| \le C \frac{\lambda(\varphi_N)^{\xi}}{\sqrt{N}} \right\}, \tag{3.36}$$

has (ξ, ν)-high probability, for some $\nu > 0$, depending on E_0, λ_0 and the distribution μ.

Hence, we obtain from (3.34),

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^{N} G_{ii} = \int \frac{\mathrm{d}\mu(v)}{\lambda v - z - m_{fc}} + R_2[\mathrm{v}] + R_3[\mathrm{v}]^2 + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(\lambda v_i - z - m_{fc})^2} \mathcal{Y}_i \\ + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(\lambda v_i - z - m_{fc})^3} (\mathcal{Y}_i^2 - 2[\mathrm{v}]\mathcal{Y}_i) + \mathcal{O}(\Lambda^3) + \mathcal{O}(\max_i |\mathcal{Y}_i|^3) + \mathcal{O}\left(\frac{\lambda(\varphi_N)^{\xi}}{\sqrt{N}} \right), \end{aligned}$$

with high probability on $\Omega(z)$, for $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Recalling the functional equation (2.9) for m_{fc}, we obtain

$$\begin{aligned} (1 - R_2)[\mathrm{v}] = R_3[\mathrm{v}]^2 + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(\lambda v_i - z - m_{fc})^2} \mathcal{Y}_i + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(\lambda v_i - z - m_{fc})^3} (\mathcal{Y}_i^2 - 2[\mathrm{v}]\mathcal{Y}_i) \\ + \mathcal{O}(\Lambda^3) + \mathcal{O}(\max_i |\mathcal{Y}_i|^3) + \mathcal{O}\left(\frac{\lambda(\varphi_N)^{\xi}}{\sqrt{N}} \right), \end{aligned} \tag{3.37}$$

with high probability on $\Omega(z)$. Recalling that $|[\mathrm{v}]| = \Lambda$, we obtain

$$|2[\mathrm{v}]\mathcal{Y}_i| \le \left(\frac{\Lambda^2}{\log N} + (\log N) \max_i |\mathcal{Y}_i|^2 \right),$$

(the added factor $\log N$ will be useful below). Using the estimates in (3.30) and (3.31), Eq. (3.34) thus becomes

$$(1 - R_2)[\mathrm{v}] = R_3[\mathrm{v}]^2 + \mathcal{O}\left(\frac{\Lambda^2}{\log N} \right) + \mathcal{O}\left(\frac{\lambda(\varphi_N)^{\xi}}{\sqrt{N}} + \Psi \right),$$

which holds with high probability on $\Omega(z)$, $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Next, observe that, since $\operatorname{Im} m_{fc}(z) \ge C\eta$, we can absorb the third term on the right side of the above equation into the forth term. Finally, we note that we can choose the constants uniform in z and λ. □

To conclude the proof of Theorems 3.1 we reason as follows. Assume, for simplicity, that $1 - R_2(z)$, $(z = E + \mathrm{i}\eta)$, is bounded below (this holds true for E in the bulk of the spectrum).

Recalling that $|[\mathrm{v}]| = \Lambda$ and the definition of $\Psi(z)$, we are going to show that (3.33) implies

$$\Lambda \le C\Lambda^2 + \mathcal{O}\left(\frac{(\varphi_N)^{\xi}}{(N\eta)^{1/3}}\right),$$

with high probability on $\Omega(z)$. Hence, we obtain the following dichotomy: Either

$$\Lambda \le C\frac{(\varphi_N)^{\xi}}{(N\eta)^{1/3}}\,, \qquad \text{or} \qquad \Lambda \ge c\,, \tag{3.38}$$

for some N-independent constant $c > 0$, with high probability on $\Omega(z)$, $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. Using the self-consistent equation (3.33), we establish in Sec. III D, that, for large η, i.e., $\eta \ge 2$, $\Lambda + \Lambda_o \le (\varphi_N)^{-2\xi}$, with high probability. In other words, $\Omega(z)$ holds with high probability, for $\operatorname{Im} z \ge 2$. But then the first inequality in (3.38) must hold, for sufficiently large N, and we can reject the second inequality in (3.38) for such η. To extend this conclusion to all $\eta \ge (\varphi_N)^L N^{-1}$, we make use of the Lipschitz continuity of the resolvent mapping $z \mapsto G(z)$, which not only allows us to establish that $\Omega(z)$ holds with high probability for η small, but also shows that (3.38) holds for small η. This continuity, or bootstrapping, argument is outlined in Sec. III E. This argument applies in a straightforward way in the bulk of the spectrum where we have $|1 - R_2(z)| \ge c > 0$. For z close to the spectral edge, $|1 - R_2(z)|$ can become very small and a slightly modified version of the above dichotomy has to be applied (see Lemma 3.12), but the bootstrapping method still applies.

D. Initial estimates for large η

To get the bootstrapping started, we need estimates on Λ_o and Λ, for $\eta \sim 1$.

Lemma 3.10. Let $\eta \ge 2$. Then, for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_L$, we have

$$\Lambda_o + \Lambda \le \frac{(\varphi_N)^{2\xi}}{\sqrt{N}}, \tag{3.39}$$

with (ξ, ν)-high probability.

Proof. Let $\lambda \in \mathcal{D}_L$. We fix $z \in \mathcal{D}_L$, with $\eta \ge 2$. Then we have the following trivial estimates

$$|G_{ij}^{(\mathbb{T})}| \le \frac{1}{\eta}\,, \quad |m^{(\mathbb{T})}| \le \frac{1}{\eta}\,, \quad |m_{fc}| \le \frac{1}{\eta}\,, \quad |R_n| \le \left(\frac{1}{\eta}\right)^n, \tag{3.40}$$

for any $\mathbb{T} \subset \{1, \ldots, N\}$.

We start with estimating Λ_o: From Eq. (3.7) we obtain using the large deviation estimates in Lemma 3.5, that

$$|G_{ij}| \le C\left(\frac{(\varphi_N)^{\xi}}{\sqrt{N}} + (\varphi_N)^{\xi}\sqrt{\frac{m^{(ij)}}{N\eta}}\right) \le C\frac{(\varphi_N)^{\xi}}{\sqrt{N}}, \tag{3.41}$$

with high probability.

To bound Λ, we note that

$$|\mathcal{Y}_i| \le |Z_i| + |m^{(i)} - m| + |w_{ii}| \le C\frac{(\varphi_N)^{\xi}}{\sqrt{N}},$$

with high probability. The self-consistent equation (3.17) can be written as

$$\begin{aligned}[\mathrm{v}] =& \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{\lambda v_i - z - m_{fc} - ([\mathrm{v}] - \mathcal{Y}_i)} - \frac{1}{\lambda v_i - z - m_{fc}}\right] \\ &+ \frac{1}{N}\sum_{i=1}^{N}\int \mathrm{d}\mu(v)\left[\frac{1}{\lambda v_i - z - m_{fc}} - \frac{1}{\lambda v - z - m_{fc}}\right]. \end{aligned} \tag{3.42}$$

The second term on the right side of the above equation is bounded by $C\frac{(\varphi_N)^\xi}{\sqrt{N}}$ with high probability, as follows from (3.36). To bound the other term, we rewrite it as

$$\frac{1}{N}\sum_{i=1}^{N}\frac{([\mathrm{v}]-\mathcal{Y}_i)}{(\lambda v_i - z - m_{fc} - ([\mathrm{v}]-\mathcal{Y}_i))(\lambda v_i - z - m_{fc})}\,.$$

Taking the imaginary part, we see that the denominators of the summands are with high probability larger in absolute value than

$$\left(2-1+\mathcal{O}\left(\frac{(\varphi_N)^\xi}{\sqrt{N}}\right)\right)2 \geq \frac{3}{2},$$

for $\eta \geq 2$. Thus, taking the maximum over i, we can bound the right side of (3.42) as

$$\Lambda = \|[\mathrm{v}]\| \leq \frac{\|[\mathrm{v}]\| + \mathcal{O}(\frac{(\varphi_N)^\xi}{\sqrt{N}})}{3/2} + \mathcal{O}\left(\frac{(\varphi_N)^\xi}{\sqrt{N}}\right),$$

with high probability. This completes the estimate of Λ and hence the proof. □

E. Proof of Theorem 3.1

We introduce the control parameters

$$\alpha(z) \equiv \alpha := |1 - R_2|\,, \qquad \beta(z) \equiv \beta := \frac{(\varphi_N)^{2\xi/3}}{(N\eta)^{1/3}}\,. \tag{3.43}$$

Note that for any $z \in \mathcal{D}_L$, we have $\beta \ll (\varphi_N)^{-3\xi}$. Also note that we have chosen β to be independent of λ.

Lemma 3.11. For R_2 and R_3, we have the following estimates:

i. There exists a constant $K > 1$, depending only on E_0, λ_0 and μ, such that,

$$\frac{1}{K}\sqrt{\kappa+\eta} \leq \alpha(z) \leq K\sqrt{\kappa+\eta}\,, \qquad z \in \mathcal{D}_L\,, \quad \lambda \in \mathcal{D}_{\lambda_0}\,. \tag{3.44}$$

In particular, we have $\operatorname{Im} m_{fc}(z) \leq C_2\alpha(z)$, for some $C_2 \geq 1$.

ii. There exists a constant C_3 such that $|R_3(z)| \leq C_3$ uniformly in $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Moreover, there exist constants c and ϵ_0 such that $|R_3(z)| \geq c$ whenever $z \in \mathcal{D}_L$ satisfies $|z - L_i| < \epsilon_0$, $i = 1, 2$.

The proof of this lemma is stated in the Appendix; see Lemma A.6.

Next, we fix E and vary η from 2 down to $(\varphi_N)^L N^{-1}$. Since $\sqrt{\kappa+\eta}$ is increasing and $\beta(E + \mathrm{i}\eta)$ is decreasing in η, we conclude that the equation

$$\sqrt{\kappa+\eta} = 2U^2 K\beta(E+\mathrm{i}\eta) \tag{3.45}$$

has a unique solution $\eta = \tilde{\eta}(U, E)$, for any $U > 1$. Note that $\tilde{\eta}(U, E) \ll 1$.

Lemma 3.12. There exists a constant U_0 such that, for any fixed $U \geq U_0$, there exists a constant $C_1(U)$, depending only on U, such that the following estimates hold for any $z \in \mathcal{D}_L$:

$$\Lambda(z) \leq U\beta(z) \quad or \quad \Lambda(z) \geq \frac{\alpha(z)}{U}\,, \qquad\qquad if\, \eta \geq \tilde{\eta}(U, E)\,, \tag{3.46}$$

$$\Lambda(z) \leq C_1(U)\beta(z)\,, \qquad\qquad if\, \eta < \tilde{\eta}(U, E), \tag{3.47}$$

on $\Omega(z)$, with (ξ, ν)-high probability.

Proof. Fix $z \in \mathcal{D}_L$. Since

$$\Psi^2 = (\varphi_N)^{2\xi} \frac{\Lambda + \operatorname{Im} m_{fc}}{N\eta} = \mathcal{O}(\beta^3\Lambda + \beta^3\alpha),$$

we can write the weak self-consistent equation (3.33) as

$$(1 - R_2)[\mathrm{v}] = R_3[\mathrm{v}]^2 + \mathcal{O}\left(\frac{\Lambda^2}{\log N}\right) + \mathcal{O}\left(\frac{\lambda(\varphi_N)^\xi}{\sqrt{N}} + \sqrt{\beta^3\Lambda + \beta^3\alpha}\right). \tag{3.48}$$

Since $\sqrt{\beta^3\Lambda + \beta^3\alpha} \le \beta\sqrt{\beta\Lambda} + \beta\sqrt{\alpha\beta} \le C(\beta^2 + \beta\alpha + \beta\Lambda)$ by Young's inequality, we obtain from (3.48)

$$\left|(1 - R_2)[\mathrm{v}] - R_3[\mathrm{v}]^2\right| \le \mathcal{O}\left(\frac{\Lambda^2}{\log N}\right) + C^*(\beta\Lambda + \alpha\beta + \beta^2)\,, \tag{3.49}$$

with high probability on $\Omega(z)$, for some $C^* \ge 1$. We set $U_0 := 9(C^* + C_3 + 1)$, where C_3 is the constant in Lemma 3.11. Depending on the size of β relative to α, we estimate either [v] or $[\mathrm{v}]^2$ using the above inequality. We have to consider two cases:

Case 1: $\eta \ge \tilde{\eta}(U, E)$ ("Bulk estimate") From (3.45) we find $\sqrt{\kappa + \eta} \ge 2U^2 K\beta(z)$ and hence, using (3.44) and the definition of C^*,

$$\beta \le \frac{\alpha}{2U^2} \le \frac{\alpha}{2C^*} \le \alpha.$$

Thus we find from (3.49) with high probability on $\Omega(z)$ that

$$\alpha\Lambda \le (|R_3| + 1)\Lambda^2 + C^*(\beta\Lambda + \alpha\beta + \beta^2) \le (C_3 + 1)\Lambda^2 + \frac{\alpha\Lambda}{2} + 2C^*\alpha\beta\,.$$

Hence, $\alpha\Lambda \le 2(C_3 + 1)\Lambda^2 + 4C^*\alpha\beta$. Thus, we either have $\alpha\Lambda/2 \le 2(C_3 + 1)\Lambda^2$ implying $\Lambda \ge \alpha/[4(C_3 + 1)] \ge \alpha/U$ (recall that $U \ge U_0 = 9(C^* + C_3 + 1)$), or $\alpha\Lambda/2 \le 4C^*\alpha\beta$ implying $\Lambda \le 8C^*\beta \le U\beta$. This proves (3.46).

Case 2: $\eta \le \tilde{\eta}(U, E)$ ("Edge estimate"). Note that, when $\kappa \sim 1$, the left side of (3.45) is of order 1, while the right side $2U^2K\beta(E + \mathrm{i}\eta) = o(1)$. (Recall that $\eta \ge (\varphi_N)^N N^{-1}$.) Thus, if $\eta \le \tilde{\eta}(U, E)$, then $\kappa < \epsilon_0$, where ϵ_0 is the constant in Lemma 3.11. In particular, $|R_3| > c$ in this case.

From (3.44) and (3.45) we find $\alpha \le 2U^2K^2\beta$. Thus from (3.49), we find

$$c\Lambda^2 \le 2\alpha\Lambda + 2C^*(\beta\Lambda + \alpha\beta + \beta^2) \le C'\beta\Lambda + C'\beta^2\,,$$

for some constant C' depending on U. Inequality (3.47) follows. □

With Lemmas 3.11 and 3.12 at hand, we are prepared to start the continuity argument: We choose a decreasing sequence (η_k), $k = 1, \ldots, k_0$ satisfying $k_0 \le CN^8$, $|\eta_k - \eta_{k+1}| \le N^{-8}$, $\eta_1 = 2$ and $\eta_{k_0} = (\varphi_N)^L N^{-1}$. For fixed $E \in [-E_0, E_0]$ we set $z_k = E + \mathrm{i}\eta_k$. Recall Lemma 3.12. We fix a $U \ge U_0$ throughout the remainder of this section.

One easily sees that, for large enough N, $\eta_1 \ge \eta(U, E)$, for any $E \in [-E_0, E_0]$. Therefore, Lemma 3.10 implies that $\Omega(z_1)$ holds with high probability. This is the starting point of the continuity argument. The next lemma extends this result to all $k \le k_0$.

Lemma 3.13. Define the event

$$\Omega_k := \Omega(z_k) \cap \{\Lambda(z_k) \le C^{(k)}(U)\beta(z_k)\}, \tag{3.50}$$

where

$$C^{(k)}(U) := \begin{cases} U & if\ \eta_k \ge \tilde{\eta}(U, E)\,, \\ C_1(U) & if\ \eta_k < \tilde{\eta}(U, E)\,. \end{cases}$$

Then, there exists $\nu > 0$, *such that for any* ξ, $1 \le k \le k_0$,

$$\mathbb{P}(\Omega_k^c) \le 3k\,\mathrm{e}^{-\nu(\log N)^\xi}\,. \tag{3.51}$$

Note that the estimates in this lemma are uniform in $\lambda \in \mathcal{D}_{\lambda_0}$.

Proof. We proceed by induction on k. The case $k = 1$ has just been proven. Hence, assume that (3.51) holds for some $k \geq 2$. Then

$$\mathbb{P}(\Omega_{k+1}^c) \leq \mathbb{P}(\Omega_k \cap \Omega(z_{k+1}) \cap \Omega_{k+1}^c) + \mathbb{P}(\Omega_k \cap (\Omega(z_{k+1}))^c) + \mathbb{P}(\Omega_k^c) =: B + A + \mathbb{P}(\Omega_k^c)\,,$$

where we set

$$A := \mathbb{P}\big(\{\Omega_k \cap \{\Lambda > (\varphi_N)^{-2\xi}\}\} \cup \{\Omega_k \cap \{\Lambda_o > (\varphi_N)^{-2\xi}\}\}\big),$$
$$B := \mathbb{P}\big(\Omega_k \cap \Omega(z_{k+1}) \cap \{\Lambda(z_{k+1}) > C^{(k+1)}(U)\beta(z_{k+1})\}\big).$$

We start by estimating A. Using the Lipschitz continuity of the resolvent map $z \mapsto G(z)$, $z \in \mathbb{C}^+$, we obtain

$$|G_{ij}(z_{k+1}) - G_{ij}(z_k)| \leq |z_{k+1} - z_k| \sup_{z \in \mathcal{D}_L} |G'_{ij}(z)| \leq N^{-8} \sup_{z \in \mathcal{D}_L} \frac{1}{(\operatorname{Im} z)^2} \leq N^{-6}.$$

Thus $\Lambda(z_{k+1}) \leq \Lambda(z_k) + N^{-6} \leq C\beta(z_k) \ll (\varphi_N)^{-2\xi}$ and

$$\Lambda_o(z_{k+1}) \leq \Lambda_o(z_k) + N^{-6} \leq C\Psi(z_k) \ll (\varphi_N)^{-2\xi},$$

with high probability on $\Omega(z_k)$, where we used Lemma 3.8. Thus $A \leq 2\,\mathrm{e}^{-\nu(\log N)^{\xi}}$.

To bound B, suppose first that $\eta_k \geq \tilde{\eta}(U, E)$. Then, using the Lipschitz continuity of the resolvent map we find $|\Lambda(z_{k+1}) - \Lambda(z_k)| \leq N^{-6}$. Thus we find on Ω_k with high probability

$$\Lambda(z_{k+1}) \leq \Lambda(z_k) + N^{-6} \leq U\beta(z_k) + N^{-6} \leq \frac{3}{2}U\beta(z_{k+1}),$$

where we used that β is a deterministic decreasing function of η.

Suppose next that $\eta_k > \eta_{k+1} \geq \tilde{\eta}(U, E)$. Then since $\frac{3}{2}U\beta < \alpha U^{-1}$, by Eq. (3.45), we find, in this case, $\Lambda(z_{k+1}) < \alpha U^{-1}$. But the dichotomy of Eq. (3.46) then implies on $\Omega_k \cap \Omega(z_{k+1})$ with high probability that $\Lambda(z_{k+1}) \leq U\beta(z_{k+1})$. If $\eta_{k+1} < \tilde{\eta}(U, E)$, the dichotomy immediately yields $\Lambda(z_{k+1}) \leq U\beta(z_{k+1})$. This shows that $B \leq \mathrm{e}^{-(\log N)^{\xi}}$ if $\eta_k \geq \tilde{\eta}(U, E)$.

If $\eta_k < \tilde{\eta}(U, E)$, then also $\eta_{k+1} < \tilde{\eta}(U, E)$ and hence Eq. (3.47) gives $\Lambda(z_{k+1}) \leq C_1(U)\beta(z_{k+1})$.

Thus, we have proven that, for all $k \leq k_0$, $\mathbb{P}(\Omega_{k+1}^c) \leq 3\mathrm{e}^{-\nu(\log N)^{\xi}} + \mathbb{P}(\Omega_k^c)$. This concludes the proof of the lemma. □

To complete the proof of Theorem 3.1, we need to extend the conclusion of the previous lemma to all $z \in \mathcal{D}_L$. To accomplish this we use a simple lattice argument using the regularity of the Green function.

Corollary 3.14. There exists constants C and $\nu > 0$, *such that, for* ξ *satisfying* (2.14),

$$\mathbb{P}\left[\bigcup_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \Omega(z)^c\right] + \mathbb{P}\left[\bigcup_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \{\Lambda(z) > C\beta(z)\}\right] \leq \mathrm{e}^{-\nu(\log N)^{\xi}}\,. \tag{3.52}$$

Proof. We choose a lattice $\mathcal{L} \subset \mathcal{D}_L$ with $|\mathcal{L}| \leq CN^6$ such that for any $z \in \mathcal{D}_L$ there is a $z' \in \mathcal{L}$ satisfying $|z - z'| \leq N^{-3}$. Using the regularity of the Green function we have for $z, z' \in \mathcal{D}_L$,

$$|G_{ij}(z) - G_{ij}(z')| \leq \eta^{-2}|z - z'| \leq \frac{1}{N}\,. \tag{3.53}$$

Lemma 3.13 yields

$$\mathbb{P}\left[\bigcap_{\substack{z' \in \mathcal{L} \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{\Lambda(z') \le \frac{C}{2}\beta(z')\right\}\right] \ge 1 - \mathrm{e}^{-\nu(\log N)^{\xi}}\,, \tag{3.54}$$

for some constants C and ν. Hence, combining (3.53), (3.54), and $N^{-1} \le \beta(z)$, we get

$$\mathbb{P}\left[\bigcup_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \{\Lambda(z) > C\beta(z)\}\right] \ge 1 - \mathrm{e}^{-\nu(\log N)^{\xi}}.$$

The first term of (3.52) is estimated in a similar way. □

This proves (3.3) of Theorem 3.1. To prove (3.1), we observe that (3.29), (3.43), and (3.52) imply that

$$\Lambda_o \le C\frac{(\varphi_N)^{\xi}}{\sqrt{N\eta}}\,,$$

with high probability on $\Omega(z)$. Then (3.52) and a similar lattice argument as above yields (3.1). To prove (3.2), we note that (3.17) yields

$$|\mathbb{E}^{v_i} G_{ii} - m| = \left|\int \frac{\mathrm{d}\mu(v)}{\lambda v - z - m_{fc}(z)} - \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\lambda v_i - z - m_{fc}}\right| + \mathcal{O}([\mathrm{v}] + \max_i |\mathcal{Y}_i|),$$

with (ξ, ν)-high probability on $\Omega(z)$, $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. From the large deviation estimate in (3.36) we find

$$|\mathbb{E}^{v_i} G_{ii} - m| \le C\left(\frac{\lambda(\varphi_N)^{\xi}}{\sqrt{N}} + \frac{(\varphi_N)^{\xi}}{(N\eta)^{1/3}}\right),$$

with high probability on $\Omega(z)$, and we can conclude the proof of (3.2) as above. This finishes the proof of Theorem 3.1.

F. Delocalization of eigenvectors

Next, we show that the eigenvectors of H are completely delocalized. We denote by $\boldsymbol{u}_\alpha$ the normalized eigenvector to the eigenvalue μ_α of $H = \lambda V + W$, i.e.,

$$(\lambda V + W)\boldsymbol{u}_\alpha = \mu_\alpha \boldsymbol{u}_\alpha,$$

such that $\|\boldsymbol{u}_\alpha\|_2^2 = \sum_i |u_\alpha(i)|^2 = 1$, where $(u_\alpha(i))$ are the components of $\boldsymbol{u}_\alpha$.

Proof of Lemma 2.15. We follow Ref. 12. For $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_L$, we have

$$|G_{ii}(z)| \le \frac{1}{|\lambda v_i - z - m_{fc} + ([\mathrm{v}] - \mathcal{Y}_i)|}\,.$$

From the weak deformed semicircle law, Theorems 3.1, we conclude that $|[\mathrm{v}] - \mathcal{Y}_i| = o(1)$, with high probability. Since $|v_i - z - m_{fc}(z)| \ge c > 0$ is bounded below uniformly in $\lambda \in \mathcal{D}_{\lambda_0}$ and $z \in \mathcal{D}_L$, by (3.5), we have

$$\max_i |G_{ii}(z)| \le C\,,$$

with (ξ, ν)-high probability, uniformly in $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Set $\eta := (\varphi_N)^L N^{-1}$, $L := 12\xi$. Then, by the spectral decomposition of H,

$$C \geq \operatorname{Im} G_{ii}(\mu_\alpha + \mathrm{i}\eta) = \sum_{\beta=1}^{N} \frac{\eta |u_\beta(i)|^2}{(\mu_\alpha - \mu_\beta)^2 + \eta^2} \geq \frac{|u_\alpha(i)|^2}{\eta},$$

with (ξ, ν)-high probability. This concludes the proof. □

IV. FLUCTUATION LEMMA AND STRONG DEFORMED SEMICIRCLE LAW

In this section, we prove a fluctuation lemma (see Lemma 4.1 below) that, when combined with the weak local deformed law yields a proof of the strong local deformed law, i.e., Theorem 2.10.

Recall that we denote by $\mathbb{E}_i$ the partial expectation with respect to the ith-column/row of the matrix W. Set $Q_i := \mathbb{1} - \mathbb{E}_i$. Roughly speaking, the main result of Subsection IV A asserts, assuming the conclusions of Theorem 3.1, that we have

$$\frac{1}{N} \sum_{i=1}^{N} Q_i \left(\frac{1}{G_{ii}} \right) \lesssim \frac{1}{N\eta}, \tag{4.1}$$

with high probability, up to logarithmic corrections. For a detailed study of fluctuation averages (for generalized Wigner- and band matrices) similar to (4.1) we refer to Ref. 14, see also Ref. 13, whose arguments we follow. The situation for the deformed ensembles considered here is in so far different as $Q_i(G_{ii})$ is of order λ, whereas $Q_i(G_{ii}) \ll 1$ in the Wigner ensemble. Note, however, that $Q_i(G_{ii}^{-1}) \lesssim (N\eta)^{-1/2}$ for the deformed model studied here as well; see below.

Using the result of Subsection IV A, we derive in Subsection IV B a "strong" self-consistent equation for $m - m_{fc}$. In Subsection IV C, we prove, following the arguments of Ref. 12, Theorem 2.10.

A. Fluctuation lemma

Recall the notation $\Lambda = |m - m_{fc}|$. We set $Q_i := \mathbb{1} - \mathbb{E}_i$, where $\mathbb{E}_i$ denotes the partial expectation with respect to the ith-column/row of the matrix W.

Lemma 4.1. Suppose ξ satisfies (2.14) and let $L \geq 12\xi$. Let Ξ be an event defined by requiring that the following holds on it: There are constants $C, c > 0$ such that,

i. *for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$,*

$$\Lambda(z) \leq \gamma(z), \tag{4.2}$$

where γ is a deterministic function satisfying $\gamma(z) \leq (\varphi_N)^{-2\xi}$;

ii. *for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$,*

$$\Lambda_o(z) \leq C\Psi(z) \leq C\Phi(z), \tag{4.3}$$

where

$$\Phi(z)^2 := (\varphi_N)^{2\xi} \frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}$$

is a deterministic control parameter;

iii. *for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$ and any $i \in \{1, \ldots, N\}$, $c \leq |G_{ii}(z)| \leq C$, and*

$$\left| Q_i \left(\frac{1}{G_{ii}(z)} \right) \right| \leq C \left(\frac{(\varphi_N)^\xi}{\sqrt{N}} + \Psi(z) \right) \leq C\Phi(z). \tag{4.4}$$

Assume that Ξ *holds with* (ξ, ν)*-high probability, then there exist constants C, c, independent of* λ *and z, such that, for* $p \in \mathbb{N}$*, even and satisfying* $p \leq \nu(\log N)^{\xi - 3/2}$,

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} Q_i\left(\frac{1}{G_{ii}(z)}\right)\right|^p \leq (Cp)^{5p}\,(\Phi(z))^{2p}\ , \tag{4.5}$$

for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$.

For the proof of this lemma, we need the following two auxiliary results:

Lemma 4.2. Let the event Ξ *be defined as in Lemma 4.1. Let* ξ *satisfy* (2.14) *and let* $L > 12\xi$*. Then there exists a constant C such that, for* $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$*, the following holds: For any* $\mathbb{T} \subset \{1, \ldots, N\}$*, with* $|\mathbb{T}| \leq (\log N)^{\xi-1}$,

$$\max_{i\notin\mathbb{T}} |G_{ii}^{(\mathbb{T})}(z) - G_{ii}(z)| \leq C|\mathbb{T}|(\Lambda_o(z))^2\,, \qquad \max_{\substack{i\neq j\\ i,j\notin\mathbb{T}}} |G_{ij}^{(\mathbb{T})}(z)| \leq C\Lambda_o(z)\,,$$

on Ξ*. In particular, we have that* $|G_{ii}^{(\mathbb{T})}(z)| \geq c$*, for some* $c > 0$*, uniformly in* $\mathbb{T}$ *and* $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$.

Proof. For simplicity we drop the z-dependence from the notation. For $l \in \mathbb{N}$, we set

$$\Gamma_l := \max\left\{\left|G_{ij}^{(\mathbb{T}')}\right| : i, j \notin \mathbb{T}', i \neq j, |\mathbb{T}'| = l\right\}, \quad \widetilde{\Gamma}_l := \max\left\{\left|G_{ii}^{(\mathbb{T}')} - G_{ii}\right| : i \notin \mathbb{T}', |\mathbb{T}'| = l\right\}.$$

Equation (3.8), i.e., $G_{ij} = G_{ij}^{(k)} + G_{ik}G_{kj}/G_{kk}$, implies that we have on Ξ

$$\Gamma_1 \leq \Lambda_o + C\Lambda_o^2 \ll (\varphi_N)^{-2\xi}\,, \qquad \widetilde{\Gamma}_1 \leq C\Lambda_o^2 \leq C(\Gamma_1)^2 \leq \Gamma_1 \ll (\varphi_N)^{-2\xi}\,.$$

In particular, we have on Ξ that $|G_{ii}^{(k)}| \geq |G_{ii}| - \widetilde{\Gamma}_1 \geq |G_{ii}| - 2\Gamma_1 > 0$, for any $k \neq i$ and $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. Assume that there is a constant C_0 such that $|G_{ii}^{(\mathbb{T}')}| \geq |G_{ii}| - 2\Gamma_1 \geq C_0^{-1}$ for any $\mathbb{T}'$ with $|\mathbb{T}'| \leq l$, $i \notin \mathbb{T}'$, and $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. Then Eq. (3.8) implies

$$\Gamma_{l+1} \leq \Gamma_l + C_0\Gamma_l^2\,, \qquad \widetilde{\Gamma}_{l+1} \leq \widetilde{\Gamma}_l + C_0\Gamma_l^2,$$

thence

$$\Gamma_{l+1} \leq \Gamma_1 + C_0\sum_{n=1}^{l}\Gamma_n^2\,, \qquad \widetilde{\Gamma}_{l+1} \leq \widetilde{\Gamma}_1 + C_0\sum_{n=1}^{l}\Gamma_n^2.$$

Thus, as long as $C_0 l\Gamma_1 \leq 1/4$, we obtain by induction that

$$\Gamma_{l+1} \leq 2\Gamma_1, \qquad \widetilde{\Gamma}_{l+1} \leq \widetilde{\Gamma}_1 + 4C_0 l(\Gamma_1)^2 \leq 2\Gamma_1,$$

and $|G_{ii}^{(\mathbb{T}')}| \geq C_0^{-1}$, for any $i \notin \mathbb{T}'$, $|\mathbb{T}'| = l+1$, $l \leq (\log N)^{\xi - 1}$. By induction on l, this proves the desired lemma. □

Lemma 4.3. Let the event Ξ *be defined as in Lemma 4.1. Let* ξ *satisfy* (2.14) *and let* $L \geq 12\xi$*. Assume that* Ξ *has* (ξ, ν)*-high probability. Then there is a constant C such that for any* $p, l \in \mathbb{N}$*, with* $p, l \leq (\log N)^{\xi - 3/2}$*, and for any* $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$*, we have*

$$\mathbb{E}\left|\frac{1}{G_{ii}^{(\mathbb{T})}(z)}\right|^p \leq C^p\,, \tag{4.6}$$

where $\mathbb{T} \subset \{1, \ldots, N\}$*, with* $|\mathbb{T}| \leq l$*, and* $i \notin \mathbb{T}$.

Proof. For simplicity we drop the z-dependence from our notation. By Lemma 4.2 we have $|G_{ii}^{(\mathbb{T})}| \geq c$ on Ξ, for any $\mathbb{T} \not\ni i$ with $|\mathbb{T}| \leq (\log N)^{\xi-1}$. On the complementary event Ξ^c, we use

Schur's complement formula (3.6),

$$\frac{1}{G_{ii}^{(\mathbb{T})}} = \lambda v_i + w_{ii} - z - \sum_{k,l}^{(i\mathbb{T})} w_{ik} G_{kl}^{(i\mathbb{T})} w_{li}\,, \qquad i \notin \mathbb{T}\,.$$

Then by Cauchy-Schwarz, the trivial bounds $|G_{ii}^{(\mathbb{T})}| \le \eta^{-1} \le N$, $\mathbb{E}|h_{ij}|^p \le N^p$ and $\mathbb{E}|\lambda v_i|^p \le \lambda_0^p$, and the boundedness of $\mathcal{D}_L$, we find

$$\mathbb{E}\left|\frac{1}{G_{ii}^{(\mathbb{T})}}\right|^p \mathbb{1}(\Xi^c) \le \left[\mathbb{E}\left|\frac{1}{G_{ii}^{(\mathbb{T})}}\right|^{2p} \mathbb{1}(\Xi^c)\right]^{1/2} \mathbb{P}(\Xi^c)^{1/2} \le (C + CN + CN^3)^p \mathbb{P}(\Xi^c)^{1/2} \le C^p\,,$$

where we used that Ξ has (ξ, ν)-high probability and that $p \le (\log N)^{\xi - 3/2}$. □

Proof of Lemma 4.1. For simplicity we drop the z-dependence from our notation. We illustrate the idea of the proof for the simple case $p = 2$:

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^N Q_i\left(\frac{1}{G_{ii}}\right)\right|^2 = \frac{1}{N^2}\sum_{i=1}^N \mathbb{E}\left|Q_i\left(\frac{1}{G_{ii}}\right)\right|^2 + \frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\, Q_i\overline{\left(\frac{1}{G_{ii}}\right)} Q_j\left(\frac{1}{G_{jj}}\right). \tag{4.7}$$

The first term on the right side is bounded by

$$\frac{1}{N^2}\sum_{i=1}^N \mathbb{E}\left|Q_i\left(\frac{1}{G_{ii}}\right)\right|^2 \mathbb{1}(\Xi) + \frac{1}{N^2}\sum_{i=1}^N \mathbb{E}\left|Q_i\left(\frac{1}{G_{ii}}\right)\right|^2 \mathbb{1}(\Xi^c) \le \frac{C}{N}\Phi^2 + \frac{o(1)}{N^2} \le C\Phi^4, \tag{4.8}$$

where we used that Ξ has (ξ, ν)-high probability and that $N^{-1/2} \le C\Phi(z)$, since $\operatorname{Im} m_{fc}(z) \ge C\eta$, $z \in \mathcal{D}_L$.

To handle the second term on the right side of (4.7), we use Eq. (3.8) to write

$$Q_j\left(\frac{1}{G_{jj}}\right) = Q_j\left(\frac{1}{G_{jj}^{(i)}} - \frac{G_{ij}G_{ji}}{G_{jj}G_{jj}^{(i)}G_{ii}}\right), \tag{4.9}$$

for $i \neq j$. Hence,

$$\mathbb{E}\, Q_i\overline{\left(\frac{1}{G_{ii}}\right)} Q_j\left(\frac{1}{G_{jj}}\right) = \mathbb{E}\, Q_i\overline{\left(\frac{1}{G_{ii}}\right)} Q_j\left(\frac{1}{G_{jj}^{(i)}} - \frac{G_{ij}G_{ji}}{G_{jj}G_{jj}^{(i)}G_{ii}}\right) = -\,\mathbb{E}\, Q_i\overline{\left(\frac{1}{G_{ii}}\right)} Q_j \frac{G_{ij}G_{ji}}{G_{jj}G_{jj}^{(i)}G_{ii}},$$

where we used that $G_{jj}^{(i)}$ is independent of the entries in the ith-column/row of W, and that, for general random variables $A = A(W)$ and $B = B(W)$, $\mathbb{E}[(Q_i A)B] = \mathbb{E}[B\mathbb{E}_i Q_i A] = 0$ if B is independent of the variables in the ith-column/row of W. Using Eq. (4.9) once more and applying the same reasoning we obtain

$$\mathbb{E}\, Q_i\overline{\left(\frac{1}{G_{ii}}\right)} Q_j\left(\frac{G_{ij}G_{ji}}{G_{jj}G_{jj}^{(i)}G_{ii}}\right) = \mathbb{E}\, Q_i\overline{\left(\frac{G_{ji}G_{ij}}{G_{ii}G_{ii}^{(j)}G_{jj}}\right)} Q_j\left(\frac{G_{ij}G_{ji}}{G_{jj}G_{jj}^{(i)}G_{ii}}\right).$$

Hence,

$$\left|\frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\, Q_i\overline{\left(\frac{1}{G_{ii}}\right)} Q_j\left(\frac{1}{G_{jj}}\right)\right| \le \sup_{i\neq j}\mathbb{E}\left|Q_i\overline{\left(\frac{G_{ji}G_{ij}}{G_{ii}G_{ii}^{(j)}G_{jj}}\right)} Q_j\left(\frac{G_{ij}G_{ji}}{G_{jj}G_{jj}^{(i)}G_{ii}}\right)\right|. \tag{4.10}$$

Using that, for a general random variable $A = A(W)$, $q \in \mathbb{N}$,

$$\mathbb{E}|Q_i A|^q \le 2^{q-1}(\mathbb{E}|A|^q + \mathbb{E}|\mathbb{E}_i A|^q) \le 2^q\mathbb{E}|A|^q\,, \tag{4.11}$$

where we used Jensen's inequality for partial expectations, we obtain

$$\left|\frac{1}{N^2}\sum_{i\neq j}\mathbb{E}\,Q_i\overline{\left(\frac{1}{G_{ii}}\right)}Q_j\left(\frac{1}{G_{jj}}\right)\right|\leq\sup_{i\neq j}C\left(\mathbb{E}\left|\frac{G_{ji}G_{ij}}{G_{ii}G_{ii}^{(j)}G_{jj}}\right|^2\right)^{1/2}\left(\mathbb{E}\left|\frac{G_{ij}G_{ji}}{G_{jj}G_{jj}^{(i)}G_{ii}}\right|^2\right)^{1/2}. \tag{4.12}$$

Using that $|G_{ij}^{(\mathbb{T})}|\leq C\Phi$, $(i\neq j)$, $|G_{ii}^{(\mathbb{T})}|>c$, on Ξ, we have, for $i\neq j$,

$$\mathbb{E}\left|\frac{G_{ji}G_{ij}}{G_{ii}G_{ii}^{(j)}G_{jj}}\right|^2\mathbb{1}(\Xi)\leq C\Phi^4\,. \tag{4.13}$$

Using Lemma 4.3 and $|G_{ij}|\leq\eta^{-1}\leq N$, Hölder's inequality yields

$$\mathbb{E}\left|\frac{G_{ji}G_{ij}}{G_{ii}G_{ii}^{(j)}G_{jj}}\right|^2\mathbb{1}(\Xi^c)\leq C\,\mathbb{P}(\Xi^c)^{1/2}N^4\leq C\Phi^4\,, \tag{4.14}$$

where we used that Ξ has (ξ,ν)-high probability, and the fact that $N^{-1/2}\leq C\Phi(z)$, $z\in\mathcal{D}_L$.

Combining the estimates (4.8), (4.12), (4.13), and (4.14), Inequality (4.5) follows for $p=2$.

Next, let $4\leq p\leq\nu(\log N)^{\xi-3/2}$ be even. Writing $p=2r$, we have

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N}Q_i\left(\frac{1}{G_{ii}}\right)\right|^{2r}=\frac{1}{N^{2r}}\sum_{i_1,\dots,i_{2r}}\mathbb{E}\prod_{j=1}^{r}Q_{i_j}\overline{\left(\frac{1}{G_{i_ji_j}}\right)}\prod_{j'=r+1}^{2r}Q_{i_{j'}}\left(\frac{1}{G_{i_{j'}i_{j'}}}\right). \tag{4.15}$$

For simplicity, we first assume that we can replace the sum over the indices $\underline{i}\equiv(i_1,\dots,i_{2r})$ by a truncated sum, where all indices are distinct, i.e., we consider

$$\frac{1}{N^{2r}}\sum_{\substack{i_1,\dots,i_{2r}\\ \text{all distinct}}}\mathbb{E}\prod_{k=1}^{r}Q_{i_k}\overline{\left(\frac{1}{G_{i_ki_k}}\right)}\prod_{k'=r+1}^{2r}Q_{i_{k'}}\left(\frac{1}{G_{i_{k'}i_{k'}}}\right). \tag{4.16}$$

As in the $p=2$ case, we make each factor of G_{ii} in the above expression independent as of many summation indices as possible by an expansion procedure that uses the identities

$$G_{ij}^{(\mathbb{T})}=G_{ij}^{(\mathbb{T}k)}+\frac{G_{ik}^{(\mathbb{T})}G_{kj}^{(\mathbb{T})}}{G_{kk}^{(\mathbb{T})}}\,, \tag{4.17}$$

for $i,j,k\notin\mathbb{T}$, $k\neq i,j$, and

$$\frac{1}{G_{ii}^{(\mathbb{T})}}=\frac{1}{G_{ii}^{(\mathbb{T}k)}}-\frac{G_{ik}^{(\mathbb{T})}G_{ki}^{(\mathbb{T})}}{G_{ii}^{(\mathbb{T})}G_{ii}^{(\mathbb{T}k)}G_{kk}^{(\mathbb{T})}}\,, \tag{4.18}$$

for $k\notin\mathbb{T}$, $k\neq i$.

The expansion procedure goes as follows: We start with expanding $F_{i_1}:=(G_{i_1i_1})^{-1}$ in (4.16). Using formula (4.18), where the choice of $k\in\{i_1,\dots,i_{2r}\}\backslash\{i_1\}$ is immaterial, we can add to $(G_{i_1i_1})^{-1}$ one upper index k. This results in two terms, $(F_{i_1})_1:=(G_{i_1i_1}^{(k)})^{-1}$ and $(F_{i_1})_0:=-G_{i_1k}G_{ki_1}/G_{i_1i_1}G_{i_1i_1}^{(k)}G_{kk}$. Using formula (4.18) we can further expand $(F_{i_1})_1$ as $(F_{i_1})_{11}+(F_{i_1})_{10}$, where $(F_{i_1})_{11}=(G_{i_1i_1}^{(kl)})^{-1}$, for $l\in\{i_1,\dots,i_{2r}\}\backslash\{i_1,k\}$ (again the choice of l is immaterial), and $(F_{i_1})_{10}$ is a fraction with two off-diagonal resolvent entries in the numerator and three diagonal resolvent entries in the denominator. Similarly, we can split the term $(F_{i_1})_0=(F_{i_1})_{00}+(F_{i_1})_{01}$, where we applied (4.17) or (4.18) to one resolvent entry of $(F_{i_1})_0$, with an index $l\neq i_1,k$. There is some arbitrariness in the choice of the resolvent entry used for the splitting that can, if desirable, be removed by choosing an ordering on the set of all resolvent entries $G_{ij}^{(\mathbb{T})}$. We continue the splitting of the terms $(F_{i_1})_\sigma$, hereby generating terms indexed by sequences σ of zeros and ones.

The precise procedure is the following. Let $\mathcal{G}$ denotes the set of monomials of resolvent entries of the form $G_{nm}^{(\mathbb{T})}$, with $n \neq m$, $\mathbb{T} \subset \{i_1, \ldots, i_{2r}\}\backslash\{n, m\}$, and $1/G_{nn}^{(\mathbb{T})}$, $\mathbb{T} \subset \{i_1, \ldots, i_{2r}\}\backslash\{n\}$. Given $F \in \mathcal{G}$, the formulas (4.17) and (4.18) define an operation, $F \mapsto F_1 \in \mathcal{G}$, by adding an upper index, e.g., $G_{nm}^{(\mathbb{T})} \mapsto G_{nm}^{(\mathbb{T}k)}$, and its complementary operation $F \mapsto F_0$, e.g., $G_{nm}^{(\mathbb{T})} \mapsto G_{nk}^{(\mathbb{T})} G_{km}^{(\mathbb{T})}/G_{kk}^{(\mathbb{T})}$, such that $F = (F)_0 + (F)_1$. Composing these operations we generate from $F \equiv (F)_\emptyset \in \mathcal{G}$, elements $(F)_\sigma \in \mathcal{G}$, labeled by binary sequences σ. For these operations we use the notation $\sigma \mapsto \sigma 0$ and $\sigma \mapsto \sigma 1$. Given $F \equiv (F)_\emptyset \in \mathcal{G}$, the recursive algorithm is as follows:

(A) *Stopping rules:*
  (1) If all terms in $(F)_\sigma$ are maximally expanded, i.e., each resolvent entry in $(F)_\sigma$ is of the form $G_{nm}^{(\mathbb{T})}$ with $n, m \notin \mathbb{T}$, $(\mathbb{T}nm) = \{i_1, \ldots, i_{2r}\}$;
  (2) else if $(F)_\sigma$ contains at least $2p$ off-diagonal resolvent entries in the numerator;

  we stop the expansion.

(B) Else, we choose an arbitrary resolvent entry $G_{nm}^{(\mathbb{T})}$ in $(F)_\sigma$. If $n = m$, we use (4.18), with some arbitrary $k \in \{i_1, \ldots, i_{2r}\}\backslash\{(\mathbb{T}n)\}$, to split $(F)_\sigma = (F)_{\sigma 0} + (F)_{\sigma 1}$. If $n \neq m$, we use (4.17), with some arbitrary $k \in \{i_1, \ldots, i_{2r}\}\backslash\{(\mathbb{T}nm)\}$, to split $(F)_\sigma = (F)_{\sigma 0} + (F)_{\sigma 1}$.

Below, we show that the stopping rules ensure that the recursive procedure is terminated after a finite number of steps. Choosing $F = (F_{i_1}) = (G_{i_1 i_1})^{-1}$, the above procedure yields

$$Q_{i_1}\overline{\left(\frac{1}{G_{i_1 i_1}}\right)} Q_{i_2}\overline{\left(\frac{1}{G_{i_2 i_2}}\right)} \ldots Q_{i_p}\left(\frac{1}{G_{i_p i_p}}\right) = \sum_\sigma Q_{i_1}\overline{(F_{i_1})_\sigma}\, Q_{i_2}\overline{\left(\frac{1}{G_{i_2 i_2}}\right)} \ldots Q_{i_p}\left(\frac{1}{G_{i_p i_p}}\right) + R_{i_1}, \tag{4.19}$$

where the summation index σ runs over a set of finite binary sequences (the number of terms in the sum is estimated below). The summands $(F_{i_1})_\sigma \in \mathcal{G}$ are fractions with off-diagonal entries of G in the numerator (except for the maximally expanded leading term $(G_{i_1 i_1}^{(\mathbb{T})})^{-1}$) and diagonal resolvent entries in the denominator. All these entries are maximally expanded in the summation indices. Each term in the rest term R_{i_1}, a fraction of resolvent entries, contains at least $2p$ off-diagonal resolvent entries in the numerator.

We claim that the total number of terms generated by the above recursive procedure is bounded by $(Cp)^{2p}$, for some p-independent constant C. Indeed, the procedure described above generates a finite rooted binary tree, whose vertices are labeled by binary sequences σ. By the stopping rules (1) and (2), each term on the right side of (4.19) corresponds to a leaf node of this tree. Thus to get an upper bound on the number of terms in (4.19), it is enough to estimate the depth of this tree.

To estimate the depth of the tree, we estimate the maximal length of a generated sequences σ. We first observe that the number of off-diagonal resolvent entries is raised by one or two under the operation $\sigma \mapsto \sigma 0$ (in case it is first applied to F_{i_1}, the number is raised by two). Hence, by stopping rule (2), the leaf nodes are labeled by sequences σ with less than $2p$ zeros in it. Also note that the operation $\sigma \mapsto \sigma 0$ increases the number of resolvent entries by at most 4, but the operation $\sigma \mapsto \sigma 1$ does not change this number. Thus the total number of resolvent entries in a term $(F_{i_1})_\sigma$ is bounded by $8p + 1$. Hence, a bound on the number of upper indices for a vertex is $(8p + 1)p$. In other words, a sequence labeling a vertex has at most $(8p + 1)p$ ones in it. Thus a sequence labeling a leaf node has a most $(8p + 1)p$ ones and $2p$ zeros, therefore has length at most $8p^2 + 3p$ and we conclude that the number of leaf nodes of the tree is bounded by

$$\sum_{q=0}^{2p} \binom{8p^2 + 3p}{q} \leq (2p + 1)\frac{(11p^2)^{2p+1}}{(2p)!} \leq (Cp)^{4p}(2p)^{-2p} \leq (Cp)^{2p},$$

for some constant C, independent of p.

It follows that the right side of (4.19) contains at most $(Cp)^{2p}$ terms. In particular, the remainder R_{i_1} contains at most $(Cp)^{2p}$ terms, each of which contains at least $2p$ off-diagonal matrix resolvent entries and less than $3p$ diagonal resolvent entries. By assumptions *ii*, Inequality (4.11), Lemmas 4.2 and 4.3, the rest term R_{i_1} satisfies

$$\mathbb{E}|R_{i_1}| \le (Cp)^{2p}\Phi(z)^{2p}\,,$$

for some sufficiently large C.

Next, we expand the term term $(G_{i_2i_2})^{-1}$ in (4.19). We apply the same procedure to each "leaf node term"

$$Q_{i_1}\overline{(F_{i_1})_\sigma}\,Q_{i_2}\overline{\left(\frac{1}{G_{i_2i_2}}\right)}\dots Q_{i_p}\left(\frac{1}{G_{i_pi_p}}\right)$$

in (4.19). Note that we do not expand the remainder term R_{i_1} any further nor start a new expansion separately for $(G_{i_2i_2})^{-1}$ (this would yield an expansion with too many terms for our purposes). We also modify the stopping rule (2) accordingly: We stop expanding a term in (4.19) whenever it contains at least $2p$ off-diagonal resolvent entries. Applying the algorithm (A)-(B) to (4.19) we find

$$Q_{i_1}\overline{\left(\frac{1}{G_{i_1i_1}}\right)}Q_{i_2}\overline{\left(\frac{1}{G_{i_2i_2}}\right)}\dots Q_{i_{2r}}\left(\frac{1}{G_{i_{2r}i_{2r}}}\right)=\sum_{\sigma_1,\sigma_2}Q_{i_1}\overline{(F_{i_1})_{\sigma_1}}\,Q_{i_2}\overline{(F_{i_2})_{\sigma_2}}\dots Q_{i_{2r}}\left(\frac{1}{G_{i_{2r}i_{2r}}}\right)+R_{i_1}+R_{i_2}, \tag{4.20}$$

where the remainder R_{i_2} satisfies the same bound as R_{i_1}. The effect of the modified stopping rule (2) is that the sequences σ_1 and σ_2 together contain in total at most $2p-1$ zeros.

Expanding the remaining $2r-2$ factors of $(G_{ii})^{-1}$ in (4.20), we find

$$\mathbb{E}\prod_{k=1}^{r}Q_{i_k}\overline{\left(\frac{1}{G_{i_ki_k}}\right)}\prod_{k'=r+1}^{2r}Q_{i_{k'}}\left(\frac{1}{G_{i_{k'}i_{k'}}}\right)=\sum_{\sigma_1,\dots,\sigma_{2r}}\mathbb{E}\left[Q_{i_1}\overline{(F_{i_1})_{\sigma_1}}\cdots Q_{i_{2r}}(F_{i_{2r}})_{\sigma_{2r}}\right]+\mathbb{E}\mathcal{R}, \tag{4.21}$$

where the remainder $\mathcal{R}=\sum_{q=1}^{p}R_{i_q}$ satisfies

$$\mathbb{E}|\mathcal{R}|\le(Cp)^{2p}\Phi^{2p}\,, \tag{4.22}$$

for some sufficiently large C. It therefore suffices to consider only the first term on the right side of (4.21), in which all monomials $(F_{i_k})_{\sigma_k}$ are maximally expanded and the summation runs over $2r$ binary sequences of finite length. Note that the total number of zeros in the array of sequences $\underline{\sigma}=(\sigma_1,\dots,\sigma_{2r})$ is, by the modified stopping rule (2), at most $2p-1$. It follows that the total number of terms in (4.20) is less than $(Cp)^{3p}$. Indeed, this can be checked in the same way as is done above: A term in (4.20) corresponds to a leaf node on a rooted binary tree, whose vertices are labeled by $\underline{\sigma}$. The total number of zeros in $\underline{\sigma}$ indexing a leaf node is bounded by $2p$ and the number of ones is less than $(8p+1)p^2$. It follows that the total number of terms in the expansion of (4.20) is bounded by $(Cp)^{3p}$ and we find

$$\left|\sum_{\sigma_1,\dots,\sigma_{2r}}\mathbb{E}\,Q_{i_1}\overline{(F_{i_1})_{\sigma_1}}\cdots Q_{i_{2r}}(F_{i_{2r}})_{\sigma_{2r}}\right|\le(Cp)^{3p}\Phi^p\,. \tag{4.23}$$

Recall that, due to our simplification assumption all indices $(i_1,\dots,i_p)$ are distinct. As in the case $p=2$ we now use the presence of the Q's: First, we claim that, for any label $a\in\{1,\dots,2r\}$,

$$\left|(F_{i_a})_{\sigma_a}\right|\mathbb{1}(\Xi)\le(C\Phi)^{1+\mathbf{0}(\sigma_a)}\,, \tag{4.24}$$

where $\mathbf{0}(\sigma_a)$ denotes the number of zeros in the sequence σ_a. For $\mathbf{0}(\sigma_a)=0$, this follows from hypothesis *iii*. If $\mathbf{0}(\sigma_a)\ge1$, the successive application of the operation $\sigma\mapsto\sigma0$ has generated at least $\mathbf{0}(\sigma_a)+1$ off-diagonal resolvent entries and at most $3\,\mathbf{0}(\sigma_a)+1$ diagonal resolvent entries.

Next, choose $(i_1,\dots,i_{2r})$ and $(\sigma_1,\dots,\sigma_{2r})$ in (4.23) such that

$$\mathbb{E}\;Q_{i_1}\overline{(F_{i_1})_{\sigma_1}}\cdots Q_{i_{2r}}(F_{i_{2r}})_{\sigma_{2r}}\neq0\,. \tag{4.25}$$

The key observation is the following:

(C) Let $a\in\{1,\dots,2r\}$, then there is a label $b\in\{1,\dots,2r\}\backslash\{a\}$, such that the monomial $(F_{i_b})_{\sigma_b}$ contains an off-diagonal resolvent entry with i_a as a lower index. We use the notation $b=\mathbf{l}(a)$, if b is linked to a in this sense.

Indeed, assuming the contrary, we conclude that all monomials $(F_{i_c})_{\sigma_c}$ in (4.25), but $(F_{i_a})_{\sigma_a}$, are independent of the random variables indexed by i_a. But due to the presence of the Q_{i_a} this term has vanishing expectation. Note that this argument relies on the assumptions that all indices $(i_1, \ldots, i_{2r})$ are distinct.

Next, let $a \in \{1, \ldots, 2r\}$ and denote by $\mathbf{l}_a := |\mathbf{l}^{-1}(\{a\})|$, the number of times the label a is linked to some label b in the sense of (C). Then,

$$|(F_{i_a})_{\sigma_a}|\mathbb{1}(\Xi) \le C^p \Phi^{1+\mathbf{l}_a}\,. \tag{4.26}$$

Indeed, for each label $c \in \mathbf{l}^{-1}(\{a\})$ we had to use at least once the operation $\sigma \mapsto \sigma 0$ to get the lower index i_a. Hence, $\mathbf{0}(\sigma_b)$, the number of zeros in σ_b, is at least $\mathbf{l}_a$. Inequality (4.26) follows from (4.24). Finally, noting that $\sum_a \mathbf{l}_a \ge p$ by (C), we find that, for terms as in (4.25),

$$\left|\mathbb{E}\, Q_{i_1}\overline{(F_{i_1})_{\sigma_1}} \cdots Q_{i_{2r}}(F_{i_{2r}})_{\sigma_{2r}}\right| \le (C\Phi)^{2p}\,. \tag{4.27}$$

Combination with the bound (4.22) on the remainder term, the estimate on the number of terms in (4.21), we thus obtain

$$\left|\frac{1}{N^{2r}}\sum_{\substack{i_1,\ldots,i_{2r}\\ \text{all distinct}}} \mathbb{E}\prod_{k=1}^{r} Q_{i_k}\overline{\left(\frac{1}{G_{i_k i_k}}\right)}\prod_{k'=r+1}^{2r} Q_{i_{k'}}\left(\frac{1}{G_{i_{k'} i_{k'}}}\right)\right| \le (Cp)^{cp}\Phi^{2p}\,,$$

for any $p \le \nu(\log N)^{\xi\,-\,3/2}$, under the simplifying assumption that all indices are distinct in the sum.

To deal with the general case, we go back to (4.15). Abbreviate $\underline{i} = (i_1, \ldots, i_{2r})$. Denote by $\mathcal{P}_{2r}$ the set of partitions of $\{1, \ldots, 2r\}$. Let $\Gamma(\underline{i})$ be the element of $\mathcal{P}_{2r}$ defined by the equivalence relation $a \sim b$, if and only if $i_a = i_b$. Then we can write

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} Q_i\left(\frac{1}{G_{ii}}\right)\right|^{2r} = \frac{1}{N^{2r}}\sum_{\Gamma\in\mathcal{P}_{2r}}\sum_{i_1,\ldots,i_{2r}}\mathbb{1}(\Gamma = \Gamma(\underline{i}))\,\mathbb{E}\, Q_{i_1}\overline{\left(\frac{1}{G_{i_1 i_1}}\right)}\ldots Q_{i_{2r}}\left(\frac{1}{G_{i_r i_r}}\right). \tag{4.28}$$

Fix now $\underline{i}$, and denote by $\Gamma := \Gamma(\underline{i})$, the partition induced by the equivalence relation $\sim$. For a label $a \in \{1, \ldots, 2r\}$, we denote by $[a]$ the block of a in Γ. Let $S(\Gamma) := \{a : |[a]| = 1\} \subset \{1, \ldots, 2r\}$ denote the set of single labels and abbreviate by $s := |S(\Gamma)|$ its cardinality. We denote by $\underline{i}_{S(\Gamma)} := (i_a)_{a\in S}$, the summation indices associated with single labels. Notice that if a is a single label (for some Γ), then there is exactly one Q_{i_a} on the right side of (4.28). However, if a is not a single label (for some Γ), Q_{i_a} appears more than once on the right side of (4.28).

Next, we expand the summands on the right side of (4.28), using the recursive procedure (A)-(B), but we only expand in the single labels. More precisely, the recursive procedure is now defined as follows:

(A) *Stopping rules:*
  - $(1')$ If all terms in $(F)_\sigma$ are maximally expanded in the single labels; a resolvent entry $G_{nm}^{(\mathbb{T})}$, is maximally expanded in the single labels if $\underline{i}_S \subset (\mathbb{T}nm)$, $n, m \notin \mathbb{T}$;
  - (2) else if $(F)_\sigma$ contains at least $2p$ off-diagonal resolvent entries in the numerator;

  we stop the expansion.

(B′) Else, we choose an arbitrary resolvent entry $G_{nm}^{(\mathbb{T})}$ in $(F)_\sigma$. If $n = m$, we use (4.17), with some arbitrary index $k \in \{\underline{i}_S\}\backslash\{(\mathbb{T}n)\}$, to split $(F)_\sigma = (F)_{\sigma 0} + (F)_{\sigma 1}$. If $n \ne m$, we use (4.18), with some arbitrary $k \in \{\underline{i}_S\}\backslash\{(\mathbb{T}nm)\}$, to split $(F)_\sigma = (F)_{\sigma 0} + (F)_{\sigma 1}$.

Applying this procedure to $(G_{i_1 i_1})^{-1}$, we obtain a similar expansion as in (4.19). Expanding the remaining factors of $(G_{i_j i_j})^{-1}$ as before (using only single labels), we obtain the analogue expression to (4.21). The remainder terms can be estimated in the same way as before, simply by using the fact that each term in the remainder contains at least $2p$ off-diagonal resolvent entries. Also note that the bounds on the number of terms in the expansion still apply. It therefore suffices to bound the summands in the first term on the right side of (4.21), (now some of the indices may coincide).

Recall that s denotes the number of single labels in the fixed configuration $\underline{i}$. We claim that

$$\left|\mathbb{E}\, Q_{i_1}\overline{(F_{i_1})}_{\sigma_1}\cdots Q_{i_{2r}}(F_{i_{2r}})_{\sigma_{2r}}\right| \le C^{2p}\Phi^{p+s}\,. \tag{4.29}$$

This follows in a similar way as above, using the following observation:

(C′) Let $a \in S(\Gamma)$, then there is an label $b \in \{1, \dots, 2r\}\backslash\{a\}$, such that the monomial $(F_{i_b})_{\sigma_b}$ contains an off-diagonal resolvent entry with i_a as a lower index.

The bound (4.29) now follows in the same way as above, by only considering single labels.

We now return to the sum in (4.28). We perform the summation by first fixing a partition $\Gamma \in \mathcal{P}_{2r}$. Then

$$\frac{1}{N^{2r}}\sum_{\underline{i}}\mathbb{1}(\Gamma = \Gamma(\underline{i})) \le \left(\frac{1}{N}\right)^{2r-|\Gamma|} \le \left(\frac{1}{\sqrt{N}}\right)^{2r-s}\,. \tag{4.30}$$

Here we used that any block in the partition Γ that is not associated to a single label, consists of at least two elements. Thus $|\Gamma| \le (2r + s)/2 = r + s/2$. Now, using $N^{-1/2} \le C\Phi$, we find, combining (4.30), (4.29), and (4.22),

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} Q_i\left(\frac{1}{G_{ii}}\right)\right|^{2r} \le (Cp)^{3p}\sum_{\Gamma\in\mathcal{P}_{2r}}(C\Phi)^{2p}\,.$$

Finally, we recall that the number of partitions of p elements is bounded by $(Cp)^{2p}$, thus

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} Q_i\left(\frac{1}{G_{ii}}\right)\right|^{2r} \le (Cp)^{5p}\Phi^{2p}\,.$$

This proves the desired lemma. □

We will use the fluctuation Lemma 4.1 in a slightly generalized setting. Abbreviate

$$g_i(z) := \frac{1}{\lambda v_i - z - m_{fc}(z)}\,, \qquad z \in \mathcal{D}_L\,, \quad \lambda \in \mathcal{D}_{\lambda_0}\,, \quad i \in \{1, \dots, N\}\,, \tag{4.31}$$

and also recall that the random variables (g_i) are bounded uniformly in λ and z as follows form the stability bound (3.5). We will use the following corollary of the fluctuation lemma:

Corollary 4.4. Suppose ξ satisfies (2.14) and let $L \ge 12\xi$. Let Ξ be the event defined in Lemma 4.1 and assume it has (ξ, ν)-high probability. Then there exists a constant C, independent of λ and z, such that, for $p \in \mathbb{N}$, even and satisfying $p \le \nu(\log N)^{\xi-3/2}$, and $n = 1, 2, 3$,

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N} Q_i\left(g_i^n\frac{1}{G_{ii}}\right)\right|^{p} \le (Cp)^{5p}\,(\Phi(z))^{2p}\,, \tag{4.32}$$

for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$.

Proof. In the proof of Lemma 4.1, we used the following two properties of (Q_i):

i. For general random variables $A = A(W)$ and $B = B(W)$, $\mathbb{E}[(Q_iA)B] = \mathbb{E}[B\mathbb{E}_iQ_iA] = 0$, if B is independent of the variables in the ith-column/row of W.

ii. For a general random variable $A = A(W)$, $q \in \mathbb{N}$, $\mathbb{E}|Q_iA|^q \le 2^q\mathbb{E}|A|^q$; see (4.11).

Fix n and define $\widetilde{Q}_i := Q_ig_i^n$. Since the random variables (v_i) are independent of the random variables (w_{ij}), property i holds true with Q_i replaced by $\widetilde{Q}_i$. (Here $\mathbb{E}$ stands for the expectation with respect the (w_{ij}) and the (v_i) random variables, but, since the random variables (g_i) are uniformly bounded, one could replace $\mathbb{E}$ by the conditional expectation with respect the (w_{ij})). Since the family of random variables (g_i) is uniformly bounded and independent of W, property ii holds now with $\mathbb{E}|\widetilde{Q}_iA|^q \le 2^q\mathbb{E}[|g_i|^{qn}]\mathbb{E}[|A|^q] \le C^q\mathbb{E}|A|^q$, for some constant C, for any random variables

$A = A(W)$ depending only on W. Thus the proof of Lemma 4.1 also applies to left side of (4.32): It suffices to multiply the bounds with C^p. □

B. Strong self-consistent equation

With Corollary 4.4 at hand, it is easy to derive a stronger self-consistent equation for $m - m_{fc}$ than the one obtained in Lemma 3.9. Recall the notation $[Z] = \frac{1}{N}\sum_{i=1}^{N} Z_i$.

Lemma 4.5. Suppose ξ satisfies (2.14). Assume that there exists a deterministic function $\gamma(z)$ with $\gamma(z) \le (\varphi_N)^{-2\xi}$ such that, for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$,

$$\Lambda(z) \le \gamma(z)\,,$$

with (ξ, ν)-high probability. Then we have with $(\xi - 2, \nu)$-high probability

$$|[Z]| \le C(\varphi_N)^{10\xi}\left(\frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}\right), \tag{4.33}$$

and, for $n = 1, 2, 3$,

$$\left|\frac{1}{N}\sum_{i=1}^{N} Q_i\left(g_i^n \frac{1}{G_{ii}}\right)\right| \le C(\varphi_N)^{10\xi}\left(\frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}\right), \tag{4.34}$$

where the constant C can be chosen uniformly in $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$.

Moreover, the strong self-consistent equation

$$\left|(1 - R_2)[\mathrm{v}] - R_3[\mathrm{v}]^2\right| \le \mathcal{O}\left(\frac{\Lambda^2}{\log N}\right) + \mathcal{O}\left((\varphi_N)^{10\xi}\left(\frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}\right)\right) + \mathcal{O}\left(\frac{\lambda(\varphi_N)^{\xi}}{\sqrt{N}}\right) \tag{4.35}$$

holds with $(\xi - 2, \nu)$-high probability, uniformly in $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$.

Note that in the above lemma we have not changed the value of the parameter ν, but replaced the N-dependent parameter ξ by $\xi - 2$. This is necessary at this point, since in the iteration procedure below we apply this lemma $\log\log N$ times.

Proof. We begin by proving (4.33). From Schur's complement formula we obtain

$$\begin{aligned} Q_i\left(\frac{1}{G_{ii}}\right) &= Q_i\left(\lambda v_i + w_{ii} - z - \sum_{k,l}^{(i)} h_{ik} G_{kl}^{(i)} h_{li}\right) \\ &= w_{ii} - Q_i\left(\sum_{k,l}^{(i)} h_{ik} G_{kl}^{(i)} h_{li}\right) \\ &= w_{ii} - Z_i\,. \end{aligned} \tag{4.36}$$

Since $|w_{ii}| \le (\varphi_N)^{\xi} N^{-1/2}$, with high probability, we obtain from the large deviation estimate (3.10),

$$\left|\frac{1}{N}\sum_{i=1}^{N} w_{ii}\right| \le \frac{(\varphi_N)^{2\xi}}{N}\,,$$

with (ξ, ν)-high probability. Hence it suffices to bound the average of the left side of (4.36) to get (4.33).

Theorem 3.1 and Lemma 3.8 imply that assumptions *i* and *ii* of Lemma 4.1 hold with high probability. By Lemma 3.7, we have $c \le |G_{ii}| \le C$, with high-probability. Finally, from the estimate on Z_i in (3.30) and the bound on w_{ii}, we conclude that assumption *iii* of Lemma 4.1, i.e., Inequality (4.4), holds with high probability. Hence the event Ξ, as defined in Lemma 4.1, being the intersection of several (ξ, ν)-high probability events, has $(\xi - 1/2, \nu)$-high probability. Thus we can apply

Lemma 4.1: Choosing p in (4.5) as the largest even integer smaller than $\nu(\log N)^{\xi-2}$, Markov's inequality yields (4.33). Note that we have not changed the parameter ν here, but have replaced ξ by $\xi - 2$.

Similarly, (4.34) follows from Corollary 4.4 and a high-moment Markov estimate.

To derive the self-consistent equation (4.35), we return to (3.37), i.e.,

$$(1-R_2)[\mathrm{v}] = R_3[\mathrm{v}]^2 + \frac{1}{N}\sum_{i=1}^N \frac{1}{(\lambda v_i - z - m_{fc})^2}\mathcal{Y}_i + \frac{1}{N}\sum_{i=1}^N \frac{1}{(\lambda v_i - z - m_{fc})^3}(\mathcal{Y}_i^2 - 2[\mathrm{v}]\mathcal{Y}_i)$$
$$+ \mathcal{O}(\Lambda^3) + \mathcal{O}(\max_i |\mathcal{Y}_i|^3) + \mathcal{O}\left(\frac{\lambda(\varphi_N)^\xi}{\sqrt{N}}\right), \tag{4.37}$$

which holds with (ξ, ν)-high probability. Recall that, on the event Ξ,

$$\mathcal{Y}_i = w_{ii} - Z_i - (m^{(i)} - m) = w_{ii} - Z_i + \mathcal{O}\left(\Phi^2\right) = Q_i\left(\frac{1}{G_{ii}}\right) + \mathcal{O}\left(\Phi^2\right). \tag{4.38}$$

Using (4.34) to control the second term on the right side of (4.37) and the *a priori* bound (3.31) on $|\mathcal{Y}_i|$ to control the terms $\mathcal{O}(\mathcal{Y}_i^2)$ in (4.37), we obtain

$$(1-R_2)[\mathrm{v}] = R_3[\mathrm{v}]^2 - 2\frac{1}{N}\sum_{i=1}^N \frac{1}{(\lambda v_i - z - m_{fc})^3}[\mathrm{v}]\mathcal{Y}_i + \mathcal{O}\left((\varphi_N)^{10\xi}\frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}\right)$$
$$+ \mathcal{O}(\Lambda^3) + \mathcal{O}\left(\frac{\lambda(\varphi_N)^\xi}{\sqrt{N}}\right),$$

with $(\xi - 2, \nu)$-high probability. Arguing as in the proof of Lemma 3.9, we obtain (4.35). □

C. Proof of the strong deformed semicircle law

The proof of Theorem 2.10 is based on an iteration using the weak semicircle law, i.e., Theorem 3.1, and Lemma 4.5. We start with an entirely deterministic lemma:

Lemma 4.6. Assume that $1 \le \xi_1 \le \xi_2$. Let $0 < \tau < 1$ and $L > 40\xi_2$. Suppose that there is a function $\gamma(z)$ satisfying

$$\gamma(z) \le (\varphi_N)^{11\xi_2}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right)^{1-\tau}, \tag{4.39}$$

such that $\Lambda(z) \le \gamma(z)$, for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. We also assume that, for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$,

$$\left|(1-R_2)[\mathrm{v}] - R_3[\mathrm{v}]^2\right| = \mathcal{O}\left(\frac{\Lambda^2}{\log N}\right) + \mathcal{O}\left(\frac{\lambda(\varphi_N)^{\xi_1}}{\sqrt{N}} + (\varphi_N)^{10\xi_1}\frac{\alpha(z)+\gamma(z)}{N\eta}\right), \tag{4.40}$$

where $\alpha = |1 - R_2|$ was defined in (3.43). Moreover, we assume that $\Lambda \ll 1$, if $\eta \sim 1$. Then

$$\Lambda(z) \le (\varphi_N)^{11\xi_2}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right)^{1-\tau/2}, \tag{4.41}$$

for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$.

Proof. The proof is based on a dichotomy argument. We set

$$\alpha_0(z) := (\varphi_N)^{(10+3/4)\xi_2}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right)^{1-\tau/2}.$$

Note that $\alpha_0 \le \gamma$ and $\alpha_0 \ll 1$. Using (4.39) we find

$$C(\varphi_N)^{10\xi_1}\left(\frac{\lambda}{\sqrt{N}}+\frac{\gamma(z)}{N\eta}\right) \le \alpha_0^2 + (\varphi_N)^{21\xi_2}\frac{1}{N\eta}\left(\frac{\lambda^{1/2}}{N^{1/4}}+\frac{1}{N\eta}\right)^{1-\tau}$$
$$\le (\varphi_N)^{22\xi_2}\left(\frac{\lambda^{1/2}}{N^{1/4}}+\frac{1}{N\eta}\right)^{2-\tau}.$$

First, consider the case $\alpha \le \alpha_0$: In this case, $\kappa \ll 1$ and $|R_3| > c$ for some constant c. From Eq. (4.40) we find that

$$|[\mathrm{v}]^2| \le \left|[\mathrm{v}]^2 - \frac{\alpha[\mathrm{v}]}{c}\right| + \frac{\alpha|[\mathrm{v}]|}{c} \le o(1)|[\mathrm{v}]|^2 + (\varphi_N)^{22\xi_2}\left(\frac{\lambda^{1/2}}{N^{1/4}}+\frac{1}{N\eta}\right)^{2-\tau} + C(\varphi_N)^{10\xi_2}\frac{\alpha}{N\eta} + \frac{\alpha|[\mathrm{v}]|}{c},$$

and hence

$$(|[\mathrm{v}]| - \frac{\alpha}{c})^2 \le \frac{\alpha^2}{c^2} + (\varphi_N)^{22\xi_2}\left(\frac{\lambda^{1/2}}{N^{1/4}}+\frac{1}{N\eta}\right)^{2-\tau} + (\varphi_N)^{(10+3/4)\xi_2}\frac{\alpha}{N\eta}.$$

(Note that we used here φ_N to compensate for various constants, as we shall do below.) Thus taking the square root, recalling that $\alpha \le \alpha_0$ and using the definition of α_0, we find

$$|[\mathrm{v}]| \le C\alpha_0 + (\varphi_N)^{11\xi_2}\left(\frac{\lambda^{1/2}}{N^{1/4}}+\frac{1}{N\eta}\right)^{1-\tau/2},$$

and the claim follows for $\alpha \le \alpha_0$.

Next, consider $\alpha > \alpha_0$: Recall that $|R_3| < C_3$ for some constant $C_3 > 0$. Assume first that $\Lambda \le \alpha/(2C_3)$. Then in (4.40) we can absorb the terms $R_3[\mathrm{v}]^2$ and $\Lambda^2/\log N$ into the term $\alpha|[\mathrm{v}]|$ and we get

$$\Lambda \le C(\varphi_N)^{10\xi_1}\left(\frac{\lambda}{\alpha\sqrt{N}}+\frac{1}{N\eta}+\frac{\gamma}{\alpha N\eta}\right)$$
$$\le \frac{1}{(\varphi_N)^{\xi_2/4}}\left(\frac{\alpha_0^2}{\alpha}+\alpha_0+\frac{\alpha_0^2}{\alpha}\right), \tag{4.42}$$

where we used the definitions of γ and α_0. Since we assumed that $\alpha > \alpha_0$, we get $\Lambda \ll \alpha/(2C_3)$ if $\Lambda \le \alpha/(2C_3)$. Thus, if $\alpha \ge \alpha_0$, we either have $\Lambda > \alpha/(2C_3)$ or $\Lambda \ll \alpha/(2C_3)$. By the continuity of $\Lambda(z)$ in $\eta = \operatorname{Im} z$, we must have $\Lambda \ll \alpha$, since we assume that $\Lambda(z) \ll 1 = \mathcal{O}(\alpha)$, for $\eta \sim 1$. Thus, the claim follows from (4.42). □

Proof of Theorem 2.10. We prove (2.18). Let $\xi = \frac{A_0+o(1)}{2}\log\log N$ and set

$$\widetilde{\xi} := 2(\log\log N/\log 2) + \xi.$$

Note that $\widetilde{\xi} \le 3\xi/2 \le A_0 \log\log N$. Let $L \ge 40\widetilde{\xi}$. To prove (2.18) it suffices to prove

$$\bigcap_{\substack{z\in\mathcal{D}_L\\ \lambda\in\mathcal{D}_{\lambda_0}}}\left\{\left|m(z)-m_{fc}(z)\right| \le (\varphi_N)^{12\widetilde{\xi}}\left(\min\left\{\frac{(\varphi_N)^{12\widetilde{\xi}}}{\alpha}\frac{\lambda}{\sqrt{N}}, \frac{\lambda^{1/2}}{N^{1/4}}\right\}+\frac{1}{N\eta}\right)\right\}, \tag{4.43}$$

with (ξ, ν)-high probability.

The weak semicircle law, i.e., Theorem 3.1 with $\widetilde{\xi}$ replacing ξ, yields

$$\Lambda \le (\varphi_N)^{2\widetilde{\xi}}\left(\frac{1}{N\eta}\right)^{1/3} \le (\varphi_N)^{2\widetilde{\xi}}\left(\frac{\lambda^{1/2}}{N^{1/4}}+\frac{1}{N\eta}\right)^{1-2/3},$$

for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, with $(\widetilde{\xi}, \nu)$-high probability. Thus (4.2) holds with

$$\gamma(z) := (\varphi_N)^{11\widetilde{\xi}}\left(\frac{\lambda^{1/2}}{N^{1/4}}+\frac{1}{N\eta}\right)^{1-2/3}.$$

Since $L \geq 40\widetilde{\xi}$, we also have $\gamma(z) \leq (\varphi_N)^{-2\widetilde{\xi}}$. Hence, by Lemma 4.5 we have

$$\left|(1 - R_2)[\mathrm{v}] - R_3[\mathrm{v}]^2\right| \leq C\frac{\Lambda^2}{\log N} + C(\varphi_N)^{10\widetilde{\xi}}\left(\frac{\lambda}{\sqrt{N}} + \frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}\right),$$

for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, with $(\widetilde{\xi} - 2, \nu)$-high probability. Since $\operatorname{Im} m_{fc} \leq C\alpha$, by Lemma 3.11, this implies (4.40) with $\xi_1 = \widetilde{\xi}$. Also, γ satisfies (4.39) with $\xi_2 = \widetilde{\xi}$ and $\tau = 2/3$. Moreover, since $\Lambda \leq \gamma \leq (\varphi_N)^{-2\widetilde{\xi}}$, we have $\Lambda \ll 1$, if $\eta \sim 1$. Therefore, we can apply Lemma 4.6 with $\xi_1 = \xi_2 = \widetilde{\xi}$ to obtain

$$\Lambda \leq (\varphi_N)^{11\widetilde{\xi}}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right)^{1-1/3},$$

for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, with $(\widetilde{\xi} - 2, \nu)$-high probability. Iterating this process M times, we find that

$$\Lambda \leq (\varphi_N)^{11\widetilde{\xi}}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right)^{1-\frac{2}{3}\left(\frac{1}{2}\right)^M},$$

for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, holds with $(\widetilde{\xi} - 2M, \nu)$-high probability. We choose $M = \lfloor \log\log N/\log 2 \rfloor - 1$, here $\lfloor \cdot \rfloor$ denotes the integer part. Since $\lambda^{1/2}N^{-1/4} + N^{-1/2} + (N\eta)^{-1} \geq cN^{-1}$ on $\mathcal{D}_L$, we get

$$\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right)^{-\frac{2}{3}\left(\frac{1}{2}\right)^M} \leq C \leq (\varphi_N)^{\widetilde{\xi}}.$$

Thus

$$\Lambda \leq (\varphi_N)^{12\widetilde{\xi}}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right), \tag{4.44}$$

for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, with $(\xi + 2, \nu)$-high probability (the factor of 2 comes from the -1 in M). This proves (4.43) when

$$\frac{(\varphi_N)^{12\widetilde{\xi}}}{\alpha}\frac{\lambda}{\sqrt{N}} \geq \frac{\lambda^{1/2}}{N^{1/4}}.$$

In case

$$\frac{\lambda^{1/2}}{N^{1/4}} \leq (\varphi_N)^{-12\widetilde{\xi}}\alpha \leq \frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta},$$

we have

$$\min\left\{\frac{(\varphi_N)^{12\widetilde{\xi}}}{\alpha}\frac{\lambda}{\sqrt{N}}, \frac{\lambda^{1/2}}{N^{1/4}}\right\} + \frac{1}{N\eta} \geq \frac{\lambda}{\sqrt{N}}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right)^{-1} + \frac{1}{N\eta} \geq \frac{1}{2}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right),$$

and the proof for (4.43) is similar to the above case. Finally, when

$$(\varphi_N)^{-12\widetilde{\xi}}\alpha \geq \frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}, \tag{4.45}$$

set

$$\gamma(z) := (\varphi_N)^{12\widetilde{\xi}}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right).$$

Then by Lemma 4.5 we have

$$\alpha|[\mathrm{v}]| \leq C\Lambda^2 + C(\varphi_N)^{10\widetilde{\xi}}\left(\frac{\alpha + \gamma(z)}{N\eta}\right) + C\frac{\lambda(\varphi_N)^{\xi}}{\sqrt{N}},$$

with (ξ, ν)-high probability, where we used that $\operatorname{Im} m_{fc} \le C\alpha$. Assuming that (4.45) holds and using the definition of γ, we have $\gamma(z) \le \alpha(z)$, and we get, using (4.44),

$$|[\mathrm{v}]| \le C(\varphi_N)^{24\widetilde{\xi}} \frac{\lambda}{\alpha\sqrt{N}} + C(\varphi_N)^{12\widetilde{\xi}} \left(\frac{1}{N\eta} + \frac{\gamma}{\alpha N\eta}\right)$$
$$\le (\varphi_N)^{12\widetilde{\xi}} \left((\varphi_N)^{12\widetilde{\xi}} \frac{\lambda}{\alpha\sqrt{N}} + \frac{1}{N\eta}\right). \tag{4.46}$$

Hence, combining (4.44) and (4.46) we find, using a simple lattice argument, (4.43). Inequality (2.19) then follow from (4.43) combined with (3.29). □

V. IDENTIFYING THE LEADING CORRECTIONS IN THE BULK

In this section, we identify the leading correction terms to $m - m_{fc}$ stemming from the diagonal random matrix V. We define random variables $\zeta_0(z) \equiv \zeta_0^N(z)$, which only depends on the random variables (v_i), such that, in the bulk of the spectrum, the leading correction term in the estimate on $|m(z) - m_{fc}(z) - \zeta_0|$ is of order $(N\eta)^{-1}$. This estimate is then used to prove Theorem 2.12.

In this section, we fix $\xi = \frac{A_0+o(1)}{2} \log\log N$ and choose $L \ge 40\xi$.

A. Preliminaries

Recall the notation $\Lambda = |m - m_{fc}|$ and the definition of (R_n) in (3.32). In Lemma 3.11, we showed that $1 - R_2 \sim \sqrt{\kappa+\eta}$, $R_3 = \mathcal{O}(1)$, for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. We will need some more notation. For $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, $n \in \mathbb{N}$, set

$$r_n(z) \equiv r_n := \frac{1}{N}\sum_{i=1}^N \frac{1}{(\lambda v_i - z - m_{fc})^n} - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc})^n}.$$

Recall from (3.36) that $|r_n(z)| \le (\varphi_N)^{\xi}\frac{\lambda}{\sqrt{N}}$, with high probability, uniformly in $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Thus, combining the above observations, we obtain

$$C^{-1}\sqrt{\kappa+\eta} \le |1 - R_2 - r_2| \le C\sqrt{\kappa+\eta}\,, \qquad |R_3 + r_3| \le C\,, \tag{5.1}$$

for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, with (ξ, ν)-high probability, for some $C > 1$.

1. Definition of ζ_0

In order to define $\zeta_0 \equiv \zeta_0(z)$, it is convenient to introduce a high-probability event Ξ_0, by requiring that (5.1) holds on it. We define ζ_0 as the solution to the equation

$$(1 - R_2 - r_2)\zeta_0(z) = r_1(z) + (R_3 + r_3)\zeta_0(z)^2\,, \qquad z \in \mathbb{C}^+\,, \quad \lambda \in \mathcal{D}_{\lambda_0}, \tag{5.2}$$

such that $\zeta_0(z) \to 0$, as $\operatorname{Im} z \to \infty$.

First, note that $\zeta_0(z)$, $z \in \mathcal{D}_L$, is well-defined on Ξ_0. Second, note that ζ_0 only depends on (v_i), but is independent of the random entries (w_{ij}) of the Wigner matrix W. Third, from the discussion in Subsection V A, we infer:

Lemma 5.1. There is a constant $c > 0$, such that, for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$, we have on Ξ_0,

$$|\zeta_0(z)| \le (\varphi_N)^{c\xi} \min\left\{\frac{\lambda^{1/2}}{N^{1/4}}, \frac{\lambda}{\sqrt{\kappa+\eta}}\frac{1}{\sqrt{N}}\right\},$$

with (ξ, ν)-high probability.

We omit the proof and just remark that it is suffices to consider the cases $\kappa + \eta \sim |1 - R_2|^2 \ll |r_3|$ and $|r_3| \ll |1 - R_2|^2$.

Recall the (strong) self-consistent equation for $m(z) - m_{fc}(z)$ in (4.35). The definition of ζ_0 is natural in the sense that it embodies the leading correction to $m - m_{fc}$ stemming from the random matrix V: Subtracting the defining equation for ζ_0 from the self-consistent equation (4.35), we obtain, after some manipulations,

$$(1 - R_2 - r_2)(m - m_{fc} - \zeta_0) = (R_3 + r_3)(m - m_{fc})^2 - (R_3 + r_3)\zeta_0^2 + \mathcal{O}(\Lambda^3) + \mathcal{O}\left((\varphi_N)^{c\xi}\frac{\operatorname{Im} m_{fc} + \Lambda}{N\eta}\right),$$

on some high probability event Ξ. Theorem 2.12, now follows easily from analyzing the stability of this equation in the variable $\zeta(z) := m(z) - m_{fc}(z) - \zeta_0(z)$.

B. Proof of Theorem 2.12

Next, we carry out the details of the proof of Theorem 2.12.

Proof of Theorem 2.12. Recall the event Ξ_0 defined in (5.1). Let

$$\gamma(z) := (\varphi_N)^{c_1\xi}\left(\min\left\{\frac{\lambda^{1/2}}{N^{1/4}}, \frac{\lambda}{\sqrt{\kappa+\eta}}\frac{1}{\sqrt{N}}\right\} + \frac{1}{N\eta}\right), \qquad z \in \mathcal{D}_L\,. \tag{5.3}$$

Choosing c_1 sufficiently large in (5.3), we can achieve that $|\zeta_0| \le \gamma(z)$ on Ξ_0. Next, it follows from Theorem 2.10 and Lemmas 4.5 and 3.8, that there is an event Ξ_1, having (ξ, ν)-high probability, such that the following holds on it: There is a constant c_0 such that $|\Lambda(z)| \le \gamma(z)$,

$$\max_i |\mathcal{Y}_i(z)| \le (\varphi_N)^{c_0\xi/2}\sqrt{\frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}}\,, \tag{5.4}$$

and, recalling (4.31),

$$\left|\frac{1}{N}\sum_{i=1}^N g_i^n \mathcal{Y}_i(z)\right| \le (\varphi_N)^{c_0\xi}\left(\frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}\right), \tag{5.5}$$

for $n = 1, 2, 3$, where $\mathcal{Y}_i = w_{ii} - Z_i - (m^{(i)} - m)$; see (3.18). Since both events Ξ_0 and Ξ_1 have high probability, the event $\Xi := \Xi_0 \cap \Xi_1$ has (ξ, ν)-high probability, with a slightly smaller $\nu > 0$. Set $\zeta(z) := m(z) - m_{fc}(z) - \zeta_0(z)$. Subtracting the defining equation of ζ_0, from Eq. (3.34), we obtain, using the bounds in (5.4) and (5.5),

$$(1 - R_2 - r_2)\zeta = (R_3 + r_3)(m - m_{fc})^2 - (R_3 + r_3)\zeta_0^2 + \mathcal{O}(\Lambda^3) + \mathcal{O}\left((\varphi_N)^{c_0\xi}\frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}\right), \tag{5.6}$$

on Ξ, for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$. Let $c_2 > \max\{c_0, c_1\}$ and set

$$\alpha_0(z) := (\varphi_N)^{c_2\xi}\left(\min\left\{\frac{\lambda^{1/2}}{N^{1/4}}, \frac{\lambda}{\sqrt{\kappa+\eta}}\frac{1}{\sqrt{N}}\right\} + \frac{1}{N\eta}\right).$$

Thus, on Ξ, we have $\Lambda \ll \alpha_0$ and $|\zeta_0| \ll \alpha_0$, for all $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$.

Recall that we defined the domain

$$\mathcal{B}_L = \mathcal{D}_L \cap \{z = E + \mathrm{i}\eta \in \mathbb{C} \,:\, \sqrt{\kappa_E + \eta} \ge (\varphi_N)^{L\xi} N^{-1/4}\}.$$

Note that we have on the domain $\mathcal{B}_L$,

$$\min\left\{\frac{\lambda^{1/2}}{N^{1/4}}, \frac{\lambda}{\sqrt{\kappa+\eta}}\frac{1}{\sqrt{N}}\right\} = \frac{\lambda}{\sqrt{\kappa+\eta}}\frac{1}{\sqrt{N}}\,.$$

First, consider the case

$$\frac{\lambda}{\sqrt{\kappa+\eta}}\frac{1}{\sqrt{N}} \ge \frac{1}{N\eta}\,.$$

In this case, we can easily see that

$$\widetilde{\alpha} := \left| \frac{1 - R_2 - r_2}{R_3 + r_3} \right| \geq \alpha_0,$$

on Ξ. Then we obtain from (5.6),

$$\begin{aligned} |m - m_{fc} - \zeta_0| &\leq \left| \frac{m - m_{fc} + \zeta_0}{\alpha_0} \right| |m - m_{fc} - \zeta_0| + C \frac{\Lambda^3}{\widetilde{\alpha}} + C \frac{(\varphi_N)^{c_0 \xi}}{\widetilde{\alpha}} \frac{\operatorname{Im} m_{fc} + \gamma(z)}{N\eta} \\ &\leq o(1)|m - m_{fc} - \zeta_0| + C \frac{\gamma(z)^3}{\widetilde{\alpha}} + C \frac{(\varphi_N)^{c_0 \xi}}{\widetilde{\alpha}} \frac{\operatorname{Im} m_{fc} + \gamma(z)}{N\eta}, \end{aligned}$$

on Ξ. By Lemma 3.11, we have $K\widetilde{\alpha} \geq \sqrt{\kappa + \eta} \sim \operatorname{Im} m_{fc}$ on Ξ_0, for some constant K, and we obtain, for some $c > c_2$,

$$|m - m_{fc} - \zeta_0| \leq (\varphi_N)^{c\xi} \left(\frac{\lambda^3}{(\kappa + \eta)^2} \frac{1}{N^{3/2}} + \frac{1}{N\eta} \right),$$

on Ξ, for all $\lambda \in \mathcal{D}_{\lambda_0}$. Since the condition $\sqrt{\kappa + \eta} \geq (\varphi_N)^{L\xi} N^{-1/4}$ implies that

$$\frac{\lambda^3}{(\kappa + \eta)^2} \frac{1}{N^{3/2}} \ll \frac{1}{N(\kappa + \eta)} \leq \frac{1}{N\eta},$$

we conclude that

$$|m - m_{fc} - \zeta_0| \leq \frac{(\varphi_N)^{c\xi}}{N\eta},$$

on Ξ, for $\lambda \in \mathcal{D}_{\lambda_0}$. If

$$\frac{\lambda}{\sqrt{\kappa + \eta}} \frac{1}{\sqrt{N}} \leq \frac{1}{N\eta},$$

we may use the bound

$$|m - m_{fc} - \zeta_0| \leq \Lambda + |\zeta_0| \leq 2\gamma \leq \frac{(\varphi_N)^{c\xi}}{N\eta},$$

on Ξ, for $\lambda \in \mathcal{D}_{\lambda_0}$. This finishes the proof. □

VI. DENSITY OF STATES

In this section, we prove Theorems 2.17, 2.18, 2.20, and 2.21. Recall that we denote by (μ_α) the eigenvalues of $H = \lambda V + W$. Define the *normalized eigenvalue counting function* of H by

$$\rho(x) := \frac{1}{N} \sum_{\alpha=1}^{N} \delta(x - \mu_\alpha). \tag{6.1}$$

Then we can write

$$m(z) = \frac{1}{N} \sum_{i=1}^{N} G_{ii}(z) = \int_{\mathbb{R}} \frac{\rho(x)\mathrm{d}x}{x - z}, \qquad z \in \mathbb{C}^+.$$

For $E_1 < E_2$, we defined in (2.28) the counting functions

$$\mathfrak{n}(E_1, E_2) = \frac{1}{N} |\{\alpha : E_1 < \mu_\alpha \leq E_2\}|, \qquad \mathfrak{n}(E) = \frac{1}{N} |\{\alpha : \mu_\alpha \leq E\}|.$$

Similarly, we have denoted

$$n_{fc}(E_1, E_2) = \int_{E_1}^{E_2} \rho_{fc}(x)\,\mathrm{d}x\,, \qquad n_{fc}(E) = \int_{-\infty}^{E} \rho_{fc}(x)\,\mathrm{d}x,$$

where ρ_{fc} stands for the density of the free convolution measure μ_{fc}.

Throughout this section we fix $\xi = \frac{A_0+o(1)}{2}\log\log N$ and choose $L \geq 40\xi$.

A. Local density of states

Recall that $\kappa_E := \min\{|E - L_i|\,,\ i = 1, 2\}$. In the following, we set $\eta := N^{-1}$. The first part of Theorem 2.17, Inequality (2.29), is an immediate consequence of the next two lemmas. Their proofs follow closely the proof of Lemmas 8.1 and 8.2 in Ref. 12.

Lemma 6.1. Let $\eta := N^{-1}$. For any $E_1 < E_2$ in $[-E_0, E_0]$, we define $f(x) \equiv f_{E_1,E_2,\eta}(x)$ to be an indicator function of the interval $[E_1, E_2]$, smoothed out on a scale η, i.e., $f(x) = 1$, for $x \in [E_1, E_2]$, $f(x) = 0$, for x in $[E_1 - \eta, E_2 + \eta]^c$, $|f'(x)| \leq C\eta^{-1}$ and $|f''(x)| \leq C\eta^{-2}$. Assume that the event,

$$\bigcap_{\substack{z\in\mathcal{D}_L\\ \lambda\in\mathcal{D}_{\lambda_0}}} \left\{ |m(z) - m_{fc}(z)| \leq (\varphi_N)^{c_0\xi}\left(\min\left\{\frac{\lambda^{1/2}}{N^{1/4}}, \frac{\lambda}{\sqrt{\kappa_E+\eta}\sqrt{N}}\right\} + \frac{1}{N\eta}\right)\right\}, \tag{6.2}$$

holds with (ξ, ν)-high probability with $L := C_0\xi$, for some constant $C_0 > 0$. Abbreviate

$$\kappa := \min\{\kappa_{E_1}, \kappa_{E_2}\}\,, \qquad \mathcal{E} := \max\{E_2 - E_1,\ (\varphi_N)^L N^{-1}\}\,.$$

Then, we have

$$\left|\int_{\mathbb{R}} f(x)(\rho - \rho_{fc})(x)\mathrm{d}x\right| \leq (\varphi_N)^{c\xi}\left(\frac{1}{N} + \frac{\mathcal{E}\lambda}{\sqrt{\kappa+\mathcal{E}}}\frac{1}{\sqrt{N}}\right), \tag{6.3}$$

with (ξ, ν)-high probability, for some $c > c_0$.

Proof. For convenience denote

$$\rho^{\Delta} := \rho - \rho_{fc}\,, \qquad m^{\Delta} := m - m_{fc}\,.$$

We use the Helffer-Sjöstrand formula. We set $y_0 := (\varphi_N)^L N^{-1}$ and choose a smooth cut-off function χ such that

$$\chi(y) = 1\,, \quad \text{on} \quad [-\mathcal{E}, \mathcal{E}]\,; \qquad \chi(y) = 0\,, \quad \text{on} \quad [-2\mathcal{E}, 2\mathcal{E}]^c\,; \qquad |\chi'(y)| \leq \frac{C}{\mathcal{E}}\,. \tag{6.4}$$

Starting from the Helffer-Sjöstrand formula,

$$f(w) = \frac{1}{2\pi}\int_{\mathbb{R}^2} \frac{\mathrm{i}yf''(x)\chi(y) + \mathrm{i}(f(x) + \mathrm{i}yf'(x))\chi'(y)}{w - x - \mathrm{i}y}\,\mathrm{d}x\,\mathrm{d}y\,, \tag{6.5}$$

we obtain

$$\begin{aligned}\left|\int_{\mathbb{R}} f(w)\rho^{\Delta}(w)\mathrm{d}w\right| &\leq C\int \mathrm{d}x \int_0^{\infty} \mathrm{d}y(|f(x)| + |y||f'(x)|)|\chi'(y)||m^{\Delta}(x+\mathrm{i}y)| \\ &\quad + C\left|\int \mathrm{d}x\int_0^{\eta} \mathrm{d}y f''(x)\chi(y)y\,\mathrm{Im}\,m^{\Delta}(x+\mathrm{i}y)\right| + C\left|\int\mathrm{d}x\int_{\eta}^{\infty}\mathrm{d}y f''(x)\chi(y)y\,\mathrm{Im}\,m^{\Delta}(x+\mathrm{i}y)\right|.\end{aligned} \tag{6.6}$$

Using that χ' is supported on $[\mathcal{E}, 2\mathcal{E}]$, we can bound the first term on the right side of the above inequality by

$$\frac{(\varphi_N)^{c\xi}}{\mathcal{E}} \int \mathrm{d}x \int_{\mathcal{E}}^{2\mathcal{E}} \mathrm{d}y \left(|f(x)| + y|f'(x)|\right) \left(\frac{\lambda}{\sqrt{\kappa_x + \mathcal{E}}\sqrt{N}} + \frac{1}{\mathcal{E}N} \right) \le (\varphi_N)^{c\xi} \left(\frac{\mathcal{E}\lambda}{\sqrt{\kappa + \mathcal{E}}\sqrt{N}} + \frac{1}{N} \right), \tag{6.7}$$

with (ξ, ν)-high probability. In order to bound the two remaining terms in (6.6), we first bound the imaginary part of $m^{\Delta}(x + \mathrm{i}y)$. For $y \ge y_0$, we can use (6.2). So assume that $0 < y < y_0$. Using the spectral decomposition of $\lambda V + W$, it is easy to see that the function $y \mapsto y \operatorname{Im} m(x + \mathrm{i}y)$ is monotone increasing. Thus,

$$y \operatorname{Im} m(x + \mathrm{i}y) \le y_0 \operatorname{Im} m(x + \mathrm{i}y_0) \le y_0 \operatorname{Im} m_{fc}(x + \mathrm{i}y_0) + (\varphi_N)^{c\xi} y_0 \left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N y_0} \right), \quad (y \le y_0). \tag{6.8}$$

Recalling that, by Lemma 3.2, we have $\operatorname{Im} m_{fc}(x + \mathrm{i}y) \le C\sqrt{\kappa_x + y}$, we get

$$y \operatorname{Im} m(x + \mathrm{i}y) \le y_0 C\sqrt{\kappa_x + y} + (\varphi_N)^{c\xi} y_0 \left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N y_0} \right) \le \frac{(\varphi_N)^{c\xi}}{N}, \qquad (y \le y_0), \tag{6.9}$$

with (ξ, ν)-high probability. Using that $y \le y_0 = (\varphi_N)^L N^{-1}$, we can now easily bound

$$|y \operatorname{Im} m^{\Delta}(x + \mathrm{i}y)| \le \frac{(\varphi_N)^{c\xi}}{N}, \qquad (y \le y_0), \tag{6.10}$$

with (ξ, ν)-high probability. Since by assumption we have $\eta \le y_0$, we can bound the second term on the right side of (6.6) by

$$\frac{(\varphi_N)^{c\xi}}{N} \int \mathrm{d}x |f''(x)| \int_0^{\eta} \mathrm{d}y\, \chi(y) \le \frac{(\varphi_N)^{c\xi}}{N}, \tag{6.11}$$

with (ξ, ν)-high probability, where we used that the support of f'' has measure $\mathcal{O}(\eta)$. To bound the third term on the right side of (6.6), we integrate by parts, first in x then in y to find the bound

$$C \left| \int \mathrm{d}x f'(x) \eta \operatorname{Re} m^{\Delta}(x + \mathrm{i}\eta) \right| + C \left| \int \mathrm{d}x \int_{\eta}^{\infty} \mathrm{d}y f'(x) \chi'(y) y \operatorname{Re} m^{\Delta}(x + \mathrm{i}y) \right| + C \left| \int \mathrm{d}x \int_{\eta}^{\infty} \mathrm{d}y f'(x) \chi(y) \operatorname{Re} m^{\Delta}(x + \mathrm{i}y) \right|. \tag{6.12}$$

The second term in (6.12) can be bounded similarly to the first term of (6.6) and we obtain

$$\left| \int \mathrm{d}x \int_{\eta}^{\infty} \mathrm{d}y f'(x) \chi'(y) y \operatorname{Re} m^{\Delta}(x + \mathrm{i}y) \right| \le (\varphi_N)^{c\xi} \left(\frac{\mathcal{E}\lambda}{\sqrt{\kappa + \mathcal{E}}\sqrt{N}} + \frac{1}{N} \right), \tag{6.13}$$

with (ξ, ν)-high probability. To bound the first and the third term in (6.12), we write, for $y \le y_0$,

$$|m^{\Delta}(x + \mathrm{i}y)| \le |m^{\Delta}(x + \mathrm{i}y_0)| + \int_y^{y_0} \mathrm{d}u \left(|\partial_u m(x + \mathrm{i}u)| + |\partial_u m_{fc}(x + \mathrm{i}u)| \right). \tag{6.14}$$

The first term on the right side of (6.14) can be estimated using (6.2). For the others we observe that the Ward identity (3.9) implies, for $u \le y_0$,

$$|\partial_u m(x + \mathrm{i}u)| = \left| \frac{1}{N} \operatorname{Tr} G^2(x + \mathrm{i}u) \right| \le \frac{1}{N} \sum_{i,j=1}^{N} |G_{ij}(x + \mathrm{i}u)|^2 = \frac{1}{u} \operatorname{Im} m(x + \mathrm{i}u) \le \frac{1}{u^2} y_0 \operatorname{Im} m(x + \mathrm{i}y_0).$$

Similarly, we obtain

$$|\partial_u m_{fc}(x + \mathrm{i}u)| \le \int \frac{\rho_{fc}(t)\mathrm{d}t}{|t - x - \mathrm{i}u|^2} = \frac{1}{u} \operatorname{Im} m_{fc}(x + \mathrm{i}u) \le \frac{1}{u^2} y_0 \operatorname{Im} m_{fc}(x + \mathrm{i}y_0).$$

From (6.14) we hence obtain

$$|m^{\Delta}(x+\mathrm{i}y)| \le (\varphi_N)^{c\xi}\left(1+\int_y^{y_0}\mathrm{d}u\,\frac{y_0}{u^2}\right) \le (\varphi_N)^{c\xi}\,\frac{y_0}{y}\,, \qquad (y\le y_0), \tag{6.15}$$

with (ξ,ν)-high probability. Thus we can bound the first term on the right side of (6.12) by

$$\left|\int \mathrm{d}x f'(x)\eta\,\mathrm{Re}\,m^{\Delta}(x+\mathrm{i}\eta)\right| \le \frac{(\varphi_N)^{c\xi}}{N}\,, \tag{6.16}$$

with high probability. To bound the third term on the right side of (6.12), we split the integration in the y variable into the pieces $[\eta, y_0)$ and $[y_0,\infty)$. Using (6.15) we can bound the first piece by

$$\int \mathrm{d}x\,|f'(x)|\int_\eta^{y_0}\mathrm{d}y\,|m^{\Delta}(x+\mathrm{i}y)| \le \frac{(\varphi_N)^{c\xi}}{N}\,,$$

with high probability. For the second integration piece we find

$$\begin{aligned}\int \mathrm{d}x\,|f'(x)|\int_{y_0}^{2\mathcal{E}}\mathrm{d}y\,|m^{\Delta}(x+\mathrm{i}y)| &\le (\varphi_N)^{c\xi}\int \mathrm{d}x\,|f'(x)|\int_{y_0}^{2\mathcal{E}}\mathrm{d}y\,\left(\frac{\lambda}{\sqrt{\kappa_x+y}}\frac{1}{\sqrt{N}}+\frac{1}{Ny}\right)\\ &\le (\varphi_N)^{c\xi}\left(\frac{1}{N}+\frac{1}{\sqrt{N}}\int_{y_0}^{2\mathcal{E}}\mathrm{d}y\,\frac{\lambda}{\sqrt{\kappa+y}}\right)\\ &\le (\varphi_N)^{c\xi}\left(\frac{1}{N}+\frac{\mathcal{E}\lambda}{\sqrt{\kappa+\mathcal{E}}}\frac{1}{\sqrt{N}}\right)\,,\end{aligned}$$

with high probability. Adding all the contributions together, we have proven that

$$\left|\int_{\mathbb{R}} f(w)\rho^{\Delta}(w)\mathrm{d}w\right| \le (\varphi_N)^{c\xi}\left(\frac{1}{N}+\frac{\mathcal{E}\lambda}{\sqrt{\kappa+\mathcal{E}}}\frac{1}{\sqrt{N}}\right),$$

with (ξ,ν)-high probability. □

As a simple corollary, we obtain

Corollary 6.2. Under the assumptions of Lemma 6.1, there is $c>0$ *such that, for any* $-E_0\le E_1<E_2\le E_0$,

$$\left|\big(\mathfrak{n}(E_2)-\mathfrak{n}(E_1)\big)-\big(n_{fc}(E_2)-n_{fc}(E_1)\big)\right| \le (\varphi_N)^{c\xi}\left(\frac{1}{N}+\frac{\mathcal{E}\lambda}{\sqrt{\kappa+\mathcal{E}}\sqrt{N}}\right)\,, \tag{6.17}$$

with (ξ,ν)*-high probability.*

Proof. Observe that, for $\eta = N^{-1}$,

$$|\mathfrak{n}(x+\eta)-\mathfrak{n}(x-\eta)| \le C\eta\,\mathrm{Im}\,m(x+\mathrm{i}\eta) \le \frac{(\varphi_N)^{c\xi}}{N}\,,$$

with (ξ,ν)-high probability, where we used (6.8). Hence,

$$\left|\mathfrak{n}(E_1)-\mathfrak{n}(E_2)-\int_{\mathbb{R}} f(w)\rho(w)\mathrm{d}w\right| \le C\sum_{i=1,2}\big(\mathfrak{n}(E_i+\eta)-\mathfrak{n}(E_i-\eta)\big) \le \frac{(\varphi_N)^{c\xi}}{N}\,, \tag{6.18}$$

with (ξ,ν)-high probability. Moreover, since ρ_{fc} is a bounded function, we find

$$\left|n_{fc}(E_1)-n_{fc}(E_2)-\int_{\mathbb{R}} f(w)\rho_{fc}(w)\mathrm{d}w\right| \le C\eta = \frac{C}{N}\,.$$

Combination with the claims of Lemma 6.1 yields the statements. □

The first statement of Theorem 2.17, i.e., (2.29), now follows easily from the two preceding lemmas.

B. Bulk fluctuations

The aim of this section is to prove the second part of Theorem 2.17, i.e., Inequality (2.30). Recall the definition of the random variables ζ_0 in (5.2). Since we will restrict the discussion to the bulk of the spectrum, we may use slightly modified random variables, $\widetilde{\zeta}_0(z) \equiv \widetilde{\zeta}_0^N(z)$, approximating ζ_0 in the bulk that are easier to handle in computations.

1. Definition of $\widetilde{\zeta}_0$

We define a random variable $\widetilde{\zeta}_0(z) \equiv \widetilde{\zeta}_0^N(z)$ by

$$\widetilde{\zeta}_0(z) = (1 - R_2(z))^{-1} \left(\frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda v_i - z - m_{fc}(z)} - m_{fc}(z) \right), \tag{6.19}$$

for $z \in \mathbb{C}^+$, $\lambda \in \mathcal{D}_{\lambda_0}$, where (R_n) have been defined in (3.32). Recall that, $1 - R_2(z) \sim \sqrt{\kappa + \eta}$. Hence, by the large deviation estimate (3.36),

$$|\widetilde{\zeta}_0(z)| \le \frac{(\varphi_N)^{c\xi} \lambda}{\sqrt{\kappa_E + \eta}\sqrt{N}}, \qquad (z = E + \mathrm{i}\eta), \tag{6.20}$$

with (ξ, ν)-high probability, for some c, uniformly in $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Also note that $\widetilde{\zeta}_0$ approximates the random variables ζ_0 in the bulk: Since $1 - R_2 \sim 1$ away from the spectral edge, it is straightforward to show that $|\zeta_0(z) - \widetilde{\zeta}_0(z)| = \mathcal{O}(N^{-1})$, with high probability for such z.

Lemma 6.3. Under the assumptions of Theorem 2.17, there is $c > 0$ such that the event

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ |m(z) - m_{fc}(z) - \widetilde{\zeta}_0(z)| \le (\varphi_N)^{c\xi} \left(\min\left\{ \frac{\lambda}{\sqrt{\kappa_E + \eta}} \frac{1}{\sqrt{N}}, \frac{\lambda^2}{(\kappa_E + \eta)^{3/2}} \frac{1}{N} \right\} + \frac{1}{N\eta} \right) \right\}, \tag{6.21}$$

where $z = E + \mathrm{i}\eta$, has (ξ, ν)-high probability.

We omit the proof of this lemma since it is similar to the proof of Theorem 2.12. Note, however, that the estimate in (6.21), deteriorates at the spectral edge and we have to restrict the discussion below mostly to the bulk of the spectrum.

Next, we show that the random variable $\widetilde{\zeta}_0(z)$, $z = E + \mathrm{i}\eta$, is a slowly varying function of E (for fixed η), in the bulk of the spectrum.

Lemma 6.4. Under the assumptions of Theorem 2.17, the event

$$\bigcap_{\substack{z \in \mathcal{D}_L \\ \lambda \in \mathcal{D}_{\lambda_0}}} \left\{ \left| \frac{\partial \widetilde{\zeta}_o(E + \mathrm{i}\eta)}{\partial E} \right| \le \frac{(\varphi_N)^{2\xi} \lambda}{(\kappa_E + \eta)^{3/2}} \frac{1}{\sqrt{N}} \right\},$$

has (ξ, ν)-high-probability.

Proof. Recalling the definition of $\widetilde{\zeta}_0$ in (6.19), we compute, for $z \in \mathcal{D}_L$, $\lambda \in \mathcal{D}_{\lambda_0}$,

$$\begin{aligned} \frac{\partial \widetilde{\zeta}_0(E + \mathrm{i}\eta)}{\partial E} &= \left(\frac{\partial}{\partial E} \frac{1}{1 - R_2(z)} \right) \left(\frac{1}{N} \sum_{i=1}^{N} Q_{v_i} \frac{1}{\lambda v_i - z - m_{fc}(z)} \right) \\ &\quad + \frac{1}{1 - R_2(z)} \left(\frac{1}{N} \sum_{i=1}^{N} Q_{v_i} \frac{1}{(\lambda v_i - z - m_{fc}(z))^2} \right) (1 + m'_{fc}(E + \mathrm{i}\eta)), \end{aligned}$$

where we abbreviate $m'_{fc}(E + \mathrm{i}\eta) \equiv \frac{\partial m_{fc}(E + \mathrm{i}\eta)}{\partial E}$ and $Q_{v_i} := \mathbb{1} - \mathbb{E}_{v_i}$, where $\mathbb{E}_{v_i}$ denotes the partial expectation with respect the random variable v_i. Differentiating the functional equation (2.9) for

$m_{fc}(z)$, we get

$$m'_{fc}(E+\mathrm{i}\eta)=\int\frac{\mathrm{d}\mu(v)}{(\lambda v-E-\mathrm{i}\eta-m_{fc}(E+\mathrm{i}\eta))^2}\left(1+m'_{fc}(E+\mathrm{i}\eta)\right),$$

hence,

$$|1+m'_{fc}(E+\mathrm{i}\eta)|=\frac{1}{|1-R_2(E+\mathrm{i}\eta)|}\leq\frac{K}{\sqrt{\kappa_E+\eta}},$$

for some constant $K>1$, where we used Lemma 3.2. Similarly,

$$\begin{aligned}\frac{\partial}{\partial E}\frac{1}{1-R_2(E+\mathrm{i}\eta)}&=\frac{2(1+m'_{fc}(E+\mathrm{i}\eta))}{(1-R_2(E+\mathrm{i}\eta))^2}\int\frac{\mathrm{d}\mu(v)}{(\lambda v-z-m_{fc}(E+\mathrm{i}\eta))^3}\\&=\frac{2(1+m'_{fc}(E+\mathrm{i}\eta))R_3(E+\mathrm{i}\eta)}{(1-R_2(E+\mathrm{i}\eta))^2},\end{aligned}$$

and hence, by Lemma 3.2,

$$\left|\frac{\partial}{\partial E}\frac{1}{1-R_2(E+\mathrm{i}\eta)}\right|\leq\frac{C}{(\kappa_E+\eta)^{3/2}},$$

for some constant C. The terms involving the Q_{v_i} can be bounded by the large deviation estimates (3.36). Uniformity in λ and z follows from a lattice argument using the stability bound (3.5). □

Next, let $f(x)\equiv f_{E_1,E_2,\eta}(x)$ be an indicator function of the interval $[E_1,E_2]$, smoothed out on scale $\eta=N^{-1}$. Let $\chi(y)$ be a smooth cut-off function as defined in (6.4). We set $m^\Delta:=m-m_{fc}$. Appealing to the discussion in Sec. VI A, we define

$$\mathfrak{X}_0(E_1,E_2):=\frac{1}{2\pi}\int_{\mathbb{R}^2}\mathrm{d}x\,\mathrm{d}y\left(\mathrm{i}yf''(x)\chi(y)+\mathrm{i}(f(x)+\mathrm{i}yf'(x))\chi'(y)\right)\widetilde{\zeta}_0(x+\mathrm{i}y),\qquad(6.22)$$

where we extend $\widetilde{\zeta}_0$ to the lower half-plane as $\widetilde{\zeta}_0(\bar{z})=\overline{\widetilde{\zeta}_0(z)}$, $z\in\mathbb{C}^+$.

Lemma 6.5. There is a constant $c>0$ such that, for $E_1<E_2$ with $E_1,E_2\in[-E_0,E_0]$ and for $\lambda\in\mathcal{D}_{\lambda_0}$, we have

$$|\mathfrak{X}_0(E_1,E_2)|\leq(\varphi_N)^{c\xi}\frac{\mathcal{E}^2\lambda}{(\kappa+\mathcal{E})^{3/2}}\frac{1}{\sqrt{N}},\qquad(6.23)$$

with (ξ,ν)-high probability, where $\mathcal{E}=\max\{E_2-E_1,(\varphi_N)^LN^{-1}\}$ and $\kappa=\min\{|E_i-L_i|:i=1,2\}$.

Choosing the energies E_1,E_2, such that $\min\{\kappa_{E_1},\kappa_{E_2}\}\geq\varkappa$, for some fixed $\varkappa>0$, we obtain

$$|\mathfrak{X}_0(E_1,E_2)|\leq C_\varkappa(\varphi_N)^{c\xi}\frac{\mathcal{E}^2\lambda}{\sqrt{N}},\qquad(6.24)$$

with (ξ,ν)-high probability, for some constant $C_\varkappa$, depending on $\varkappa$.

Proof. Starting from the definition of $\widetilde{\zeta}_0$, we find

$$\begin{aligned}|\mathfrak{X}_0(E_1,E_2)|\leq C&\left|\int\mathrm{d}x\int_0^\infty\mathrm{d}y\,f(x)\chi'(y)\widetilde{\zeta}_0(x+\mathrm{i}y)\right|+\left|\int\mathrm{d}x\int_0^\infty\mathrm{d}y\,yf'(x)\chi'(y)\widetilde{\zeta}_0(x+\mathrm{i}y)\right|\\&+C\left|\int\mathrm{d}x\int_0^{2\mathcal{E}}\mathrm{d}y\,f''(x)\chi(y)y\,\mathrm{Im}\,\widetilde{\zeta}_0(x+\mathrm{i}y)\right|.\end{aligned}\qquad(6.25)$$

To bound the first term on the right side of (6.25) we integrate by part in the variable y to find, with (ξ, ν)-high probability,

$$\begin{aligned}\left|\int \mathrm{d}x\, f(x)\int_{\mathcal{E}}^{2\mathcal{E}}\mathrm{d}y\, \chi'(y)\widetilde{\zeta}_0(x+\mathrm{i}y)\right| &= \left|\int \mathrm{d}x\, f(x)\int_{\mathcal{E}}^{2\mathcal{E}}\mathrm{d}y\, \chi(y)\partial_y\widetilde{\zeta}_0(x+\mathrm{i}y)\right|\\ &= \left|\int \mathrm{d}x\, f(x)\int_{\mathcal{E}}^{2\mathcal{E}}\mathrm{d}y\, \chi(y)\partial_x\widetilde{\zeta}_0(x+\mathrm{i}y)\right|\\ &\le (\varphi_N)^{c\xi}\mathcal{E}\left|\int \mathrm{d}x\, f(x)\frac{\lambda}{(\kappa_x+\mathcal{E})^{3/2}}\frac{1}{\sqrt{N}}\right|\\ &\le (\varphi_N)^{c\xi}\frac{\mathcal{E}\lambda}{(\kappa+\mathcal{E})^{3/2}}\frac{1}{\sqrt{N}}\,, \end{aligned}\tag{6.26}$$

where we used in the second line that $\widetilde{\zeta}_0(z)$ in an analytic function in the upper half plane, and in the third line we used Lemma 6.4. In the fourth line we used that $\kappa_x \ge \kappa = \min\{\kappa_{E_1}, \kappa_{E_2}\}$. Finally, we used that f is supported on $[E_1 - \eta, E_2 + \eta]$, $\eta = N^{-1}$.

To bound the second term on the right side of (6.25), we integrate by part in the variable x and find, similarly to the computation above,

$$\begin{aligned}\left|\int_{E_1-\eta}^{E_2+\eta}\mathrm{d}x\int_0^{\infty}\mathrm{d}y\, yf'(x)\chi'(y)\widetilde{\zeta}_0(x+\mathrm{i}y)\right| &= \left|\int_{E_1-\eta}^{E_2+\eta}\mathrm{d}x\int_0^{\infty}\mathrm{d}y\, yf(x)\chi'(y)\partial_x\widetilde{\zeta}_0(x+\mathrm{i}y)\right|\\ &\le (\varphi_N)^{c\xi}\int_{E_1-\eta}^{E_2+\eta}\mathrm{d}x\int_{\mathcal{E}}^{2\mathcal{E}}\mathrm{d}y\, yf(x)|\chi'(y)|\frac{\lambda}{(\kappa_x+\mathcal{E})^{3/2}}\frac{1}{\sqrt{N}}\\ &\le (\varphi_N)^{c\xi}\frac{\lambda}{(\kappa+\mathcal{E})^{3/2}}\frac{1}{\sqrt{N}}\int_{E_1-\eta}^{E_2+\eta}\mathrm{d}x\, f(x)\int_{\mathcal{E}}^{2\mathcal{E}}\mathrm{d}y\,\frac{y}{\mathcal{E}}\\ &\le (\varphi_N)^{c\xi}\frac{\mathcal{E}^2\lambda}{(\kappa+\mathcal{E})^{3/2}}\frac{1}{\sqrt{N}}, \end{aligned}\tag{6.27}$$

with (ξ, ν)-high probability.

Finally, the third term in (6.25) can be bounded by integrating by parts in x to obtain

$$\begin{aligned}\left|\int \mathrm{d}x\int_0^{2\mathcal{E}}\mathrm{d}y\, f''(x)\chi(y)y\,\mathrm{Im}\,\widetilde{\zeta}_0(x+\mathrm{i}y)\right| &= \left|\int \mathrm{d}x\int_0^{2\mathcal{E}}\mathrm{d}y\, f'(x)\chi(y)y\,\mathrm{Im}\,\partial_x\widetilde{\zeta}_0(x+\mathrm{i}y)\right|\\ &\le (\varphi_N)^{c\xi}\left|\int_0^{2\mathcal{E}}\mathrm{d}y\frac{\lambda y}{(\kappa+y)^{3/2}}\frac{1}{\sqrt{N}}\right|\\ &\le (\varphi_N)^{c\xi}\frac{\mathcal{E}^2\lambda}{(\kappa+\mathcal{E})^{3/2}}\frac{1}{\sqrt{N}}\,, \end{aligned}\tag{6.28}$$

with (ξ, ν)-high probability. Adding up the estimates (6.26), (6.27), and (6.28) yields the claim. □

2. Local eigenvalue density in the bulk

In this subsection, we show that we can control the difference $\mathfrak{n}(E_2) - \mathfrak{n}(E_1)$ in terms of $n_{fc}(E_2) - n_{fc}(E_1)$ in the bulk of the spectrum up to an optimal error: Fix some $\varkappa > 0$. We consider energies $E_1 < E_2$, such that $\min\{\kappa_{E_1}, \kappa_{E_2}\} \ge \varkappa$, $L_1 < E_1 < E_2 < L_2$, and $E_2 - E_1 \ge (\varphi_N)^{L\xi}N^{-1}$. We denote with $C_\varkappa$ constants that only depend on $\varkappa$ (with $C_\varkappa \to \infty$, as $\varkappa \to 0$).

As above, let $f(x) \equiv f_{E_1,E_2,\eta}(x)$ be an indicator function of the interval $[E_1, E_2]$, smoothed out on scale $\eta = N^{-1}$. Let $\chi(y)$ be a smooth cut-off function as defined in (6.4) and let $m^{\Delta} := m - m_{fc}$. Define

$$\mathfrak{X}_1(E_1, E_2) := \frac{1}{2\pi}\int_{\mathbb{R}^2}\big(\mathrm{i}yf''(x)\chi(y) + \mathrm{i}(f(x) + \mathrm{i}yf'(x))\chi'(y)\big)m^{\Delta}(x+\mathrm{i}y)\,,\tag{6.29}$$

and recall the definition of $\mathfrak{X}_0$ in (6.22),

$$\mathfrak{X}_0(E_1, E_2) := \frac{1}{2\pi}\int_{\mathbb{R}^2}\big(\mathrm{i}yf''(x)\chi(y) + \mathrm{i}(f(x) + \mathrm{i}yf'(x))\chi'(y)\big)\widetilde{\zeta}_0(x + \mathrm{i}y). \tag{6.30}$$

Here we implicitly assume that the functions f and χ in both definitions agree.

Following the discussion given in Sec. VI A, one easily sees that

$$\big|\big(\mathfrak{n}(E_1, E_2) - n_{fc}(E_1, E_2)\big) - \mathfrak{X}_1(E_1, E_2)\big| \le \frac{(\varphi_N)^{c\xi}}{N}, \tag{6.31}$$

with (ξ, ν)-high probability. Recalling the estimate on $\mathfrak{X}_0$ in (6.23), we observe that it suffices to bound $\mathfrak{X}_1 - \mathfrak{X}_0$ in order to control the density of states.

Lemma 6.6. Let $L_1 < E_1 < E_2 < L_2$, with $\min\{\kappa_{E_1}, \kappa_{E_2}\} \ge \varkappa$ and $E_2 - E_1 \ge (\varphi_N)^{L\xi}N^{-1}$. Then

$$|\mathfrak{X}_1(E_1, E_2) - \mathfrak{X}_0(E_1, E_2)| \le C_\varkappa(\varphi_N)^{c\xi}\frac{1}{N}, \tag{6.32}$$

with (ξ, ν)-high probability. The constant $c > 0$ can be chosen independent of E_1, E_2 and $\lambda \in \mathcal{D}_{\lambda_0}$.

Proof. We set $y_0 := (\varphi_N)^L N^{-1}$ and abbreviate $\widetilde{\zeta}(z) = m(z) - m_{fc}(z) - \widetilde{\zeta}_0(z)$. Using the definition of $\widetilde{\zeta}_0$, we find

$$\begin{aligned}
&|(\mathfrak{X}_1 - \mathfrak{X}_0)(E_1, E_2)| \\
&\le C\int \mathrm{d}x\int_0^\infty(|f(x)| + |yf'(x)|)|\chi'(y)|\,|\widetilde{\zeta}(x + \mathrm{i}y)| + C\left|\int \mathrm{d}x\int_0^{y_0}\mathrm{d}y f''(x)\chi(y)y\,\mathrm{Im}\,\widetilde{\zeta}_0(x + \mathrm{i}y)\right| \\
&\quad + C\left|\int \mathrm{d}x\int_0^{y_0}\mathrm{d}y f''(x)\chi(y)y\,\mathrm{Im}\,m^{\Delta}(x + \mathrm{i}y)\right| + C\left|\int \mathrm{d}x\int_{y_0}^{\infty}\mathrm{d}y f''(x)\chi(y)y\,\mathrm{Im}\,\widetilde{\zeta}(x + \mathrm{i}y)\right|.
\end{aligned} \tag{6.33}$$

Using (6.21) we can bound the first term on the right side of (6.33) as

$$\begin{aligned}
\left|\int \mathrm{d}x\int_{\mathcal{E}}^{2\mathcal{E}}\big(|f(x)| + y|f'(x)|\big)|\chi'(y)|\,|\widetilde{\zeta}(x + \mathrm{i}y)|\right| &\le C_\varkappa\frac{(\varphi_N)^{c\xi}}{\mathcal{E}}\int \mathrm{d}x\int_{\mathcal{E}}^{2\mathcal{E}}\mathrm{d}y(|f(x)| + y|f'(x)|)\frac{1}{Ny} \\
&\le C_\varkappa(\varphi_N)^{c\xi}\frac{1}{N},
\end{aligned} \tag{6.34}$$

with (ξ, ν)-high probability.

The second term on the right side of (6.33) is, by (6.20), bounded by

$$\begin{aligned}
\left|\int \mathrm{d}x\int_0^{y_0}\mathrm{d}y f''(x)\chi(y)y\,\mathrm{Im}\,\widetilde{\zeta}_0(x + \mathrm{i}y)\right| &\le (\varphi_N)^{c\xi}\frac{\lambda}{\eta}\int_0^{y_0}\mathrm{d}y\;\chi(y)y\frac{1}{\sqrt{\kappa + y}}\frac{1}{\sqrt{N}} \\
&\le (\varphi_N)^{c\xi}\frac{1}{\eta}\int_0^{y_0}\mathrm{d}y\frac{\lambda\sqrt{y}}{\sqrt{N}} \\
&\le \frac{(\varphi_N)^{c\xi}}{N},
\end{aligned} \tag{6.35}$$

with (ξ, ν)-high probability.

To control the third term, we note that both functions $y \mapsto y\;\mathrm{Im}\,m(x + \mathrm{i}y)$, $y\;\mathrm{Im}\,m_{fc}(x + \mathrm{i}y)$, are monotone increasing. Thus we get from (6.2),

$$y\,\mathrm{Im}\,m(x + \mathrm{i}y) \le y_0\,\mathrm{Im}\,m(x + \mathrm{i}y_0) \le (\varphi_N)^{c\xi}y_0\left(\sqrt{\kappa_x + y_0} + \frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{Ny_0}\right) \le \frac{(\varphi_N)^{c\xi}}{N}, \quad (y \le y_0),$$

and

$$y\,\mathrm{Im}\,m_{fc}(x + \mathrm{i}y) \le y_0\,\mathrm{Im}\,m_{fc}(x + \mathrm{i}y_0) \le Cy_0\sqrt{\kappa_x + y_0}, \qquad (y \le y_0).$$

Since $y_0 = (\varphi_N)^L N^{-1}$, this yields

$$y\,|\operatorname{Im}\widetilde{\zeta}(x+\mathrm{i}y)| \le \frac{(\varphi_N)^{c\xi}}{N}\,, \qquad (y \le y_0)\,, \tag{6.36}$$

with (ξ,ν)-high probability. The third term on the right side of (6.33) is thus bounded as

$$\left|\int \mathrm{d}x \int_0^{y_0} \mathrm{d}y f''(x)\chi(y) y \operatorname{Im} m^{\Delta}(x+\mathrm{i}y)\right| \le \frac{(\varphi_N)^{c\xi}}{N}\int \mathrm{d}x |f''(x)| \int_0^{y_0}\mathrm{d}y\,\chi(y) \le \frac{(\varphi_N)^{c\xi}}{N}, \tag{6.37}$$

with (ξ,ν)-high probability.

To bound the fourth term in (6.33), we integrate first by parts in the variable x and then in y, to find the bound

$$\left|\int \mathrm{d}x \int_{y_0}^{2\mathcal{E}} \mathrm{d}y f'(x)\partial_y(\chi(y)y)\operatorname{Re}\widetilde{\zeta}(x+\mathrm{i}y)\right| + \left|\int \mathrm{d}x f'(x)\chi(y_0)y_0 \operatorname{Re}\widetilde{\zeta}(x+\mathrm{i}y_0)\right|. \tag{6.38}$$

Using the *a priori* high probability bounds

$$|m^{\Delta}(x+\mathrm{i}y_0)| \le (\varphi_N)^{c\xi}\left(\frac{\lambda^{1/2}}{N^{1/4}}+\frac{1}{Ny_0}\right) \le (\varphi_N)^{c\xi}, \quad |\widetilde{\zeta}_0(x+\mathrm{i}y_0)| \le (\varphi_N)^{c\xi}\frac{\lambda}{\sqrt{\kappa_x+y_0}}\frac{1}{\sqrt{N}} \le (\varphi_N)^{c\xi},$$

we bound the second term on the right side of (6.38) as

$$\left|\int \mathrm{d}x f'(x) y_0 \widetilde{\zeta}(x+\mathrm{i}\eta)\right| \le (\varphi_N)^{c\xi} y_0 \le \frac{(\varphi_N)^{c\xi}}{N},$$

with (ξ,ν)-high probability. It remains to bound the first term in (6.38),

$$\begin{aligned}\left|\int \mathrm{d}x \int_{y_0}^{2\mathcal{E}} \mathrm{d}y f'(x)\partial_y(\chi(y)y)\operatorname{Re}\widetilde{\zeta}(x+\mathrm{i}y)\right| &\le \left|\int \mathrm{d}x \int_{y_0}^{2\mathcal{E}} \mathrm{d}y f'(x)\chi'(y)y\operatorname{Re}\widetilde{\zeta}(x+\mathrm{i}y)\right| \\ &\quad + \left|\int \mathrm{d}x \int_{y_0}^{2\mathcal{E}} \mathrm{d}y f'(x)\chi(y)\operatorname{Re}\widetilde{\zeta}(x+\mathrm{i}y)\right|. \end{aligned} \tag{6.39}$$

For the first term on the right side, we use (6.21) to find

$$\left|\int \mathrm{d}x \int_{y_0}^{2\mathcal{E}} \mathrm{d}y f'(x)\chi'(y)y\operatorname{Re}\widetilde{\zeta}(x+\mathrm{i}y)\right| \le C_{\varkappa}(\varphi_N)^{c\xi}\frac{1}{N}\,,$$

with (ξ,ν)-high probability. Using once more (6.21), we bound the second term on the right side of (6.39) as

$$\left|\int \mathrm{d}x \int_{y_0}^{\infty} \mathrm{d}y f'(x)\chi(y)\operatorname{Re}\widetilde{\zeta}(x+\mathrm{i}y)\right| \le C_{\varkappa}(\varphi_N)^{c\xi}\int_{y_0}^{2\mathcal{E}}\mathrm{d}y\frac{1}{\mathcal{E}N} \le (\varphi_N)^{c\xi}\frac{1}{N}, \tag{6.40}$$

with (ξ,ν)-high probability. Adding up the different contributions, we find (6.32). □

To conclude this subsection, we prove (2.30) of Theorem 2.17:

Proof of (2.30). Let $E_1 < E_2$. Then we have from (6.31)

$$\begin{aligned}|\mathfrak{n}(E_1,E_2) - n_{fc}(E_1,E_2)| &\le |\mathfrak{X}_1(E_1,E_2)| + C(\varphi_N)^{c\xi}\frac{1}{N} \\ &\le |\mathfrak{X}_0(E_1,E_2)| + |\mathfrak{X}_1(E_1,E_2)-\mathfrak{X}_0(E_1,E_2)| + C(\varphi_N)^{c\xi}\frac{1}{N}\,,\end{aligned}$$

with (ξ,ν)-high probability. Using Lemma 6.5 and Lemma 6.6, we therefore get

$$|\mathfrak{n}(E_1,E_2) - n_{fc}(E_1,E_2)| \le C_{\varkappa}(\varphi_N)^{c\xi}\left(\frac{1}{N}+\frac{\lambda^2\mathcal{E}^2}{\sqrt{N}}\right),$$

with (ξ,ν)-high probability. Inequality (2.30) follows by choosing $E_2 - E_1 \ge (\varphi_N)^{L\xi}N^{-1}$. □

C. Eigenvalue spacing in the bulk

In this subsection, we prove Theorem 2.18.

Proof of Theorem 2.18. Let $\lambda \in \mathcal{D}_{\lambda_0}$. Starting from the identity

$$\frac{i-j}{N} = \mathfrak{n}(\mu_i) - \mathfrak{n}(\mu_j)\,,$$

we obtain from (2.30) that

$$n_{fc}(\mu_i) - n_{fc}(\mu_j) = \frac{i-j}{N} + \mathcal{O}\left((\varphi_N)^{c\xi}\frac{(\mu_i-\mu_j)^2}{\sqrt{N}}\right) + \mathcal{O}\left((\varphi_N)^{c\xi}\frac{1}{N}\right),$$

with (ξ,ν)-high probability, for some c large enough. Then, using $n_{fc}(\mu_i) - n_{fc}(\mu_j) = (\mu_i - \mu_j)n'_{fc}(\mu'_i)$, for some $\mu'_i \in [\mu_i, \mu_j]$,

$$\mu_i - \mu_j = \frac{i-j}{N\rho_{fc}(\mu'_i)} + \mathcal{O}\left((\varphi_N)^{c\xi}\frac{(\mu_i-\mu_j)^2}{\sqrt{N}}\right) + \mathcal{O}\left((\varphi_N)^{c\xi}\frac{1}{N}\right), \tag{6.41}$$

where we used that $n'_{fc}(\mu'_i) = \rho_{fc}(\mu'_i) > 0$ and $1/C' < \rho_{fc} < C'$ in the bulk for some constant $C' > 1$, depending on λ and μ. Since $|\mu_i - \mu_j| = \mathcal{O}(1)$, we have

$$(\varphi_N)^{c\xi}\frac{(\mu_i-\mu_j)^2}{\sqrt{N}} \ll |\mu_i - \mu_j|\,,$$

which shows that the second term in the right side of (6.41) can be absorbed into the left side. Similarly, the last term on the right side can be absorbed into the first term in the right side, as we can see from the condition $|i - j| \gg (\varphi_N)^{c\xi}$. Thus,

$$C_1\frac{|i-j|}{N} \le |\mu_i - \mu_j| \le C_2\frac{|i-j|}{N}\,, \tag{6.42}$$

with (ξ,ν)-high probability for some constants $C_1, C_2 > 0$. This proves the first part of the theorem.

If $|i - j| \le (\varphi_N)^{c\xi}N^{1/2}$, we find from (6.42) that $|\mu_i - \mu_j| \le C_2(\varphi_N)^{c\xi}N^{-1/2}$. In this case,

$$(\varphi_N)^{c\xi}\frac{(\mu_i-\mu_j)^2}{\sqrt{N}} \ll \frac{1}{N}\,,$$

with high probability. Furthermore, since $|\mu'_i - \mu_i| \le |\mu_i - \mu_j|$ and since ρ_{fc} is Lipschitz continuous inside supp μ_{fc} (see, e.g., Ref. 5), we get

$$|\rho_{fc}(\mu'_i) - \rho_{fc}(\mu_i)| \le (\varphi_N)^{K\xi}\frac{1}{\sqrt{N}}\,, \tag{6.43}$$

hence

$$\left|\frac{i-j}{N\rho_{fc}(\mu'_i)} - \frac{i-j}{N\rho_{fc}(\mu_i)}\right| \le C\frac{|i-j|}{N}\frac{|\rho_{fc}(\mu'_i)-\rho_{fc}(\mu_i)|}{\rho_{fc}(\mu'_i)\rho_{fc}(\mu_i)} \le (\varphi_N)^{K\xi}N^{-1}\,,$$

with (ξ,ν)-high probability, for some constant K. Thus, we obtain that

$$\left||\mu_i - \mu_j| - \frac{|i-j|}{N\rho_{fc}(\mu_i)}\right| \le (\varphi_N)^{K\xi}N^{-1}\,,$$

with (ξ,ν)-high probability, proving the second part the theorem. □

D. Integrated density of states and rigidity of eigenvalues

The goal of this subsection is to prove Theorems 2.20 and 2.21. The proofs follow closely Ref. 12.

1. Estimate on $\|H\|$

As a first step, we need an estimate on the operator norm of $H = \lambda V + W$. We have the following result:

Lemma 6.7. There is a constant $c_0 > 0$, such that for all $\lambda \in \mathcal{D}_{\lambda_0}$, we have

$$\|H\| \le \max\{|L_1|, L_2\} + (\varphi_N)^{c_0\xi}\left(\frac{\lambda}{\sqrt{N}} + \frac{1}{N^{2/3}}\right),$$

with (ξ, ν)-high probability.

Proof. We will only consider the largest eigenvalue μ_N. A bound on the lowest eigenvalue μ_1 is obtained in a similar way. From the strong local law (2.18), we get

$$\Lambda(z) \le (\varphi_N)^{c\xi}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right), \qquad z \in \mathcal{D}_L, \quad \lambda \in \mathcal{D}_{\lambda_0},$$

with $(\xi + 2, \nu)$-high probability. Then we can apply Lemma 4.6, with

$$\gamma(z) := (\varphi_N)^{c\xi}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right),$$

to get, for some sufficiently large constant c_1,

$$\left|(1-R_2)[\mathrm{v}] - R_3[\mathrm{v}]^2\right| \le C\frac{\Lambda^2}{\log N} + C(\varphi_N)^{c_1\xi}\left(\frac{\lambda}{\sqrt{N}} + \frac{\operatorname{Im} m_{fc}(z) + \gamma(z)}{N\eta}\right),$$

with (ξ, ν)-high probability, for any $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Now, if $E > L_2$ and $\kappa \ge \eta$, we have $\operatorname{Im} m_{fc}(z) \sim \eta/\sqrt{\kappa}$ and

$$\alpha := |1 - R_2| \sim \sqrt{\kappa},$$

by Lemmas 3.2 and 3.11. Thus, we obtain, upon using Young's inequality,

$$\left|(1-R_2)[\mathrm{v}] - R_3[\mathrm{v}]^2\right| \le C\frac{\Lambda^2}{\log N} + C(\varphi_N)^{c_1\xi}\left(\frac{\lambda}{\sqrt{N}} + \frac{1}{(N\eta)^2} + \frac{1}{N\sqrt{\kappa}}\right), \tag{6.44}$$

with (ξ, ν)-high probability, for some c_1 sufficiently large.

Given c_1, it is straightforward to check that there is a constant $c_2 > 2c_1$, such that, for any E satisfying

$$L_2 + (\varphi_N)^{c_2\xi}\left(\frac{\lambda}{\sqrt{N}} + \frac{1}{N^{2/3}}\right) \le E \le E_0, \tag{6.45}$$

we have

$$\min\{N^{-1/2}\kappa^{1/4}, N^{-1/2}\lambda^{-1}\kappa^{1/2}, \kappa\} \ge (\varphi_N)^{c_1\xi+2}\frac{1}{N\sqrt{\kappa}}. \tag{6.46}$$

We assume now that E satisfies (6.45) and set

$$\eta \equiv \eta_E := (\varphi_N)^{c_1\xi+1}\frac{1}{N\sqrt{\kappa}}.$$

Note that $z = E + \mathrm{i}\eta \in \mathcal{D}_L$. From (6.46), we have $\kappa \ge \eta$. Similarly, we have

$$\operatorname{Im} m_{fc}(E+\mathrm{i}\eta) \sim \frac{\eta}{\sqrt{\kappa}} \ll \frac{1}{N\eta}; \qquad \frac{\lambda}{\sqrt{\kappa}N} \ll \frac{1}{N\eta}. \tag{6.47}$$

Furthermore, since $\alpha \ge \sqrt{\kappa}/K$, for some $K > 1$, we must have

$$2|R_3|\Lambda(z) \le C(\varphi_N)^{c_1\xi+1}\left(\frac{\lambda^{1/2}}{N^{1/4}} + \frac{1}{N\eta}\right) \le \alpha, \tag{6.48}$$

with (ξ, ν)-high probability. Here, we used that $N\eta\sqrt{\kappa} = (\varphi_N)^{c_1\xi+1}$ and $\sqrt{\kappa} \geq (\varphi_N)^{c_2\xi/2}\lambda^{1/2}N^{-1/4}$. Since $\alpha \geq 2|R_3|\Lambda$, we get from (6.44),

$$\Lambda \leq C(\varphi_N)^{c_1\xi}\left(\frac{\lambda}{\alpha\sqrt{N}} + \frac{1}{\alpha(N\eta)^2} + \frac{1}{\alpha N\sqrt{\kappa}}\right), \tag{6.49}$$

with (ξ, ν)-high probability. Since $\alpha \geq \sqrt{\kappa}/K$, we obtain from (6.47),

$$(\varphi_N)^{c_1\xi+1}\frac{1}{\alpha N\sqrt{\kappa}} \leq C\frac{\eta}{\sqrt{\kappa}} \ll \frac{1}{N\eta}.$$

The second term on the right side of (6.49), can be bounded by using (6.48). The first term on the right side of (6.49) is estimated by using the second inequality in (6.47). Thus, for any E satisfying (6.45), $z \in \mathcal{D}_L$, and any $\lambda \in \mathcal{D}_{\lambda_0}$, we obtain that

$$\Lambda(z) \ll \frac{1}{N\eta},$$

with (ξ, ν)-high probability. Thus

$$\operatorname{Im} m(z) \leq \operatorname{Im} m_{fc}(z) + \Lambda(z) \ll \frac{1}{N\eta}, \tag{6.50}$$

with (ξ, ν)-high probability, for such E. By the spectral decomposition of H, we have

$$\operatorname{Im} m(z) = \frac{1}{N}\sum_{\alpha=1}^{N}\frac{\eta}{(\mu_\alpha - E)^2 + \eta^2},$$

and we conclude that

$$\operatorname{Im} m(z) \geq \frac{C}{N\eta}, \tag{6.51}$$

for some $C > 0$, if there is an eigenvalue in the interval $[E - \eta, E + \eta]$. Thus (6.50), implies, for any E, satisfying (6.45), that there is no eigenvalue in the interval $[E - \eta, E + \eta]$, with (ξ, ν)-high probability.

To cover energies $E \geq E_0$, we use the following result: For a Wigner matrix W satisfying the assumptions in Definition 2.1 we have

$$\|W\| \leq 2 + \frac{(\varphi_N)^\xi}{N^{1/4}}, \tag{6.52}$$

with (ξ, ν)-high probability. We refer, e.g., to Lemma 4.3. in Ref. 12. Spectral perturbation theory then implies $\|H\| \leq \|\lambda V\| + \|W\| \leq 2 + \frac{(\varphi_N)^\xi}{N^{1/4}} + \lambda$, with (ξ, ν)-high probability, covering the regime $E \geq E_0$. This concludes the proof. □

2. Integrated density of states

In this subsection, we prove Theorem 2.20. Given the results on $\mathfrak{n}(E_1, E_2)$ in Theorem 2.17 and the estimate on $\|H\|$ this is straightforward:

Proof of Theorem 2.20. We assume that E is such that $|E - L_1| \leq |E - L_2|$. The other case is dealt with in the same way. Set

$$E_1 = L_1 - (\varphi_N)^{c_1\xi}\left(\frac{\lambda}{\sqrt{N}} + \frac{1}{N^{2/3}}\right), \tag{6.53}$$

with some c_1 large enough, such that $n_{fc}(E_1) = 0$ and $\mathfrak{n}(E_1) = 0$ with (ξ, ν)-high probability; see Lemma 6.7.

Next, choose $E \geq E_1$, then from (2.29), we get, setting $E_2 = E$ and bounding $\mathcal{E} \leq E - E_1 + (\varphi_N)^{L\xi}N^{-1}$,

$$\left|\mathfrak{n}(E) - n_{fc}(E)\right| \leq (\varphi_N)^{c\xi}\left(\frac{1}{N} + \frac{\lambda}{\sqrt{N}}\sqrt{E - E_1 + (\varphi_N)^{L\xi}N^{-1}}\right),$$

with (ξ, ν)-high probability. Using our assumption on E and (6.53), we get

$$\left|\mathfrak{n}(E) - n_{fc}(E)\right| \le (\varphi_N)^{c\xi}\left(\frac{1}{N} + \frac{\lambda^{3/2}}{N^{3/4}} + \frac{\lambda}{N^{5/6}} + \frac{\lambda\sqrt{\kappa_E}}{\sqrt{N}}\right),$$

with (ξ, ν)-high probability, for some c_2 large enough. This estimate holds for any E and λ. Uniformity is obtained with a lattice argument, we omit the details. □

3. Rigidity of eigenvalues

In this subsection, we prove Theorem 2.21. Recall the definition of the classical location γ_α of the eigenvalue μ_α in (2.33).

Lemma 6.8. There exists a constant C, such that, for all $\lambda \in \mathcal{D}_{\lambda_0}$, the following statements hold with (ξ, ν)-high probability for some large enough c,

i. *if* $\max\{\kappa_{\gamma_\alpha}, \kappa_{\mu_\alpha}\} \le (\varphi_N)^{c\xi}(\frac{\lambda}{\sqrt{N}} + \frac{1}{N^{2/3}})$, *then*

$$|\mu_\alpha - \gamma_\alpha| \le (\varphi_N)^{C\xi}\left(\frac{\lambda}{\sqrt{N}} + \frac{1}{N^{2/3}}\right);$$

ii. *if* $\max\{\kappa_{\gamma_\alpha}, \kappa_{\mu_\alpha}\} \ge (\varphi_N)^{c\xi}(\frac{\lambda}{\sqrt{N}} + \frac{1}{N^{2/3}})$, *then*

$$|\mu_\alpha - \gamma_\alpha| \le (\varphi_N)^{C\xi}\left(\frac{\lambda}{\sqrt{N}} + \frac{1}{\widehat{\alpha}^{1/3}N^{2/3}} + \frac{\lambda^2}{N^{1/3}\widehat{\alpha}^{2/3}}\right),$$

where $\widehat{\alpha} := \min\{\alpha, N-\alpha\}$.

Proof. We will focus on the eigenvalues $\mu_1, \ldots, \mu_{N/2}$. The other eigenvalues can be treated in a similar way. Define an event Ξ as the intersection of the events on which the estimates

$$\|H\| \le \max\{|L_1|, L_2\} + (\varphi_N)^{C_0\xi}\left(\frac{\lambda}{\sqrt{N}} + \frac{1}{N^{2/3}}\right), \tag{6.54}$$

(see Lemma 6.7), and

$$|\mathfrak{n}(E) - n_{fc}(E)| \le (\varphi_N)^{C_0\xi}\left(\frac{1}{N} + \frac{\lambda^{3/2}}{N^{3/4}} + \frac{\lambda}{N^{5/6}} + \frac{\lambda\sqrt{\kappa_E}}{\sqrt{N}}\right),$$

(see Theorem 2.20), hold, for any $\lambda \le \lambda_0$ and $|E| \le E_0$. We note that on Ξ, we have $\mu_{N/2} \le K$, for some $K < L_2$.

Let $C_1 > C_0$. We use the dyadic decomposition

$$\{1, \ldots, N/2\} = \bigcup_{k=0}^{2\log N} U_k,$$

where

$$U_0 := \left\{\alpha \le N/2 \,:\, |L_1| + \max\{\mu_\alpha, \gamma_\alpha\} \le 2(\varphi_N)^{C_1\xi}\lambda N^{-1/2}\right\},$$
$$U_k := \left\{\alpha \le N/2 \,:\, 2^k\lambda(\varphi_N)^{C_1\xi}N^{-1/2} \le |L_1| + \max\{\mu_\alpha, \gamma_\alpha\} \le 2^{k+1}(\varphi_N)^{C_1\xi}\lambda N^{-1/2}\right\}, \quad (k \ge 1).$$

By the definition of U_0 and (6.54), we have

$$|\mu_\alpha - \gamma_\alpha| \le (\varphi_N)^{C\xi}\frac{\lambda}{N^{1/2}},$$

on Ξ, for $\alpha \in U_0$.

For $k \ge 1$, we find on Ξ that

$$\frac{\alpha}{N} = n_{fc}(\gamma_\alpha) = \mathfrak{n}(\mu_\alpha) = n_{fc}(\mu_\alpha) + (\varphi_N)^{C_0\xi}\mathcal{O}\left(\frac{1}{N} + \frac{\lambda^{3/2}}{N^{3/4}} + \frac{\lambda}{N^{5/6}} + \frac{\lambda\sqrt{\kappa_E}}{\sqrt{N}}\right). \tag{6.55}$$

On Ξ, and for $\alpha \in U_k$, we can bound the second term on the right side of the above equation as

$$(\varphi_N)^{C_0\xi}\mathcal{O}\left(\frac{1}{N}+\frac{\lambda^{3/2}}{N^{3/4}}+\frac{\lambda}{N^{5/6}}+\frac{\lambda\sqrt{\kappa_E}}{\sqrt{N}}\right)$$
$$\leq C(\varphi_N)^{C_0\xi}\left(\frac{1}{N}+\frac{\lambda^{3/2}}{N^{3/4}}+\frac{\lambda}{N^{5/6}}\right)+C2^{(k+1)/2}(\varphi_N)^{(C_0+C_1/2)\xi}\frac{\lambda^{3/2}}{N^{3/4}},$$

where we used $\kappa_{\mu_\alpha} \leq |L_1| + \mu_\alpha$. Furthermore, we have on Ξ, for $\alpha \in U_k$,

$$n_{fc}(\gamma_\alpha)+n_{fc}(\mu_\alpha)\geq c2^{3k/2}(\varphi_N)^{3C_1\xi/2}\lambda^{3/2}N^{-3/4},$$

where we used $n_{fc}(L_1 + x) \sim x^{3/2}$, for $0 \leq x \leq |L_1| + K$. Thus

$$(\varphi_N)^{C_0\xi}\mathcal{O}\left(\frac{1}{N}+\frac{\lambda^{3/2}}{N^{3/4}}+\frac{\lambda}{N^{5/6}}+\frac{\lambda\sqrt{\kappa_E}}{\sqrt{N}}\right)\ll n_{fc}(\gamma_\alpha)+n_{fc}(\mu_\alpha),$$

which implies by (6.55) that

$$n_{fc}(\mu_\alpha)=n_{fc}(\gamma_\alpha)\left(1+\mathcal{O}\left((\varphi_N)^{-(C_1-C_0)\xi}\right)\right),$$

on Ξ, for $\alpha \in U_k$. Using that $n'_{fc}(x) \sim (n_{fc}(x))^{1/3} \sim (|L_1| + x)^{1/2}$, for $L_1 \leq x \leq K$, we have $|L_1| + \gamma_\alpha \sim |L_1| + \mu_\alpha$. Hence

$$n'_{fc}(x)\sim n'_{fc}(\gamma_\alpha),$$

for any x between γ_α and μ_α. Recalling that the density $\rho_{fc}(x)$ is continuous, we conclude that, on Ξ, for $\alpha \in U_k$,

$$\begin{aligned}|\mu_\alpha-\gamma_\alpha| &\leq C\frac{|n_{fc}(\mu_\alpha)-n_{fc}(\gamma_\alpha)|}{n'_{fc}(\gamma_\alpha)}\\ &\leq \frac{C(\varphi_N)^{C_0\xi}}{(\alpha/N)^{1/3}}\left(\frac{1}{N}+\frac{\lambda^{3/2}}{N^{3/4}}+\frac{\lambda}{N^{5/6}}+\frac{\lambda\sqrt{\kappa_{\mu_\alpha}}}{\sqrt{N}}\right)\\ &\leq \frac{C(\varphi_N)^{C_0\xi}}{\alpha^{1/3}}\left(\frac{1}{N^{2/3}}+\frac{\lambda^{3/2}}{N^{5/12}}+\frac{\lambda\alpha^{1/3}}{\sqrt{N}}+\frac{\lambda\sqrt{|\mu_\alpha-\gamma_\alpha|}}{N^{1/6}}\right),\end{aligned} \tag{6.56}$$

where we used $\kappa_{\mu_\alpha} \leq \kappa_{\gamma_\alpha} + |\mu_\alpha - \gamma_\alpha|$ and $\kappa_{\gamma_\alpha} \sim (\alpha/N)^{2/3}$. Next, since $\alpha = Nn_{fc}(\gamma_\alpha) \sim N(|L_1| + \gamma_\alpha)^{3/2}$, we find for $\alpha \in U_k$, $(k \geq 1)$,

$$\alpha \geq cN\left(2^k(\varphi_N)^{C_1\xi}\frac{\lambda}{\sqrt{N}}\right)^{3/2}\gg N^{1/4},$$

hence $\alpha^{-1/3} \ll N^{-1/12}$. Using Young's inequality, we can absorb the last term on the right side of (6.56) into the left side and we obtain

$$|\mu_\alpha-\gamma_\alpha|\leq(\varphi_N)^{C\xi}\left(\frac{1}{\alpha^{1/3}N^{2/3}}+\frac{\lambda^2}{\alpha^{2/3}N^{1/3}}+\frac{\lambda}{\sqrt{N}}\right),$$

on Ξ, for $\alpha \in U_k$, some C sufficiently large. The proof is completed by noticing that the event Ξ has (ξ, ν)-high probability. □

We conclude this section with the proof of Theorem 2.21.

Proof of Theorem 2.21. We restrict the discussion to eigenvalues with $\alpha \leq N/2$, the other eigenvalues are dealt with in the same way. From $\alpha/N = n_{fc}(\gamma_\alpha) \sim (|L_1| + \gamma_\alpha)^{3/2}$, we find that

$$\alpha\leq(\varphi_N)^{C\xi}(1+\lambda^{3/2}N^{1/4}), \tag{6.57}$$

if α is as in item *i* of Lemma 6.8. Combing the conclusions of items *i* and *ii* of Lemma 6.8 with (6.57) completes the proof of the theorem. □

ACKNOWLEDGMENTS

We thank Horng-Tzer Yau for suggesting this problem to us and for numerous helpful discussions. We are grateful to Antti Knowles for discussions and many useful remarks. We also thank Alex Bloemendal, Paul Bourgade, László Erdős, and an anonymous referee for helpful comments.

Ji Oon Lee is partially supported by Basic Science Research Program through the National Research Foundation of Korea, Grant No. 2011-0013474.

The stay of Kevin Schnelli at IAS is supported by The Fund For Math. Most of the work has been accomplished while Kevin Schnelli stayed at the Department of Mathematics, Harvard University, Cambridge MA, USA.

APPENDIX: FREE CONVOLUTION MEASURE AND STABILITY BOUNDS

1. Introduction

In this appendix, we discuss some properties of the (rescaled) free convolution measure, μ_{fc}, defined through the functional equation

$$m_{fc}(z) = \int_{-1}^{1} \frac{\mathrm{d}\mu_{fc}(v)}{\lambda v - z - m_{fc}(z)}\,, \qquad z = E + \mathrm{i}\eta \in \mathbb{C}^+\,, \tag{A1}$$

such that $\operatorname{Im} m_{fc}(z) > 0$, for $\eta > 0$; cf. Eq. (2.9). Here $\lambda \geq 0$ and we assume that μ is an absolutely continuous measure, with bounded and continuous density $\mu(v)$ such that $\operatorname{supp}\mu = [-1, 1]$. For simplicity, we always assume that μ is centered, although this is not essential for our argument.

To see that Eq. (A1) has a unique solution such that $\operatorname{Im} m_{fc}(E + \mathrm{i}\eta) > 0$, for $\eta > 0$, one can choose $\eta > 2$ first. Then it is straightforward to check that the right side of (A1) is a contraction (in the sup-norm on the set of analytic function on the upper half plane with positive imaginary part). The fixed point equation (A1) thus has a unique solution for $\eta > 2$. By analytic continuation, the solution extends to the whole upper half plane. We leave the details aside and refer, e.g., to Ref. 34.

A deep study of Eq. (A1), with slightly different conventions, can be found in Ref. 5. One important result of Ref. 5 is the following: The measure μ_{fc} is absolutely continuous with respect to Lebesgue measure, in particular, we have $\pi\mu_{fc}(E) = \lim_{\eta\searrow 0} \operatorname{Im} m_{fc}(E + \mathrm{i}\eta)$. For general probability measures (of bounded support), the support of μ_{fc} may consist of several disjoint intervals, however, under our assumptions, the support of μ_{fc} is a single interval, i.e., $\operatorname{supp}\mu_{fc} = [L_1, L_2]$, with $L_1 < 0 < L_2$; see Lemma A.1. We refer to Ref. 5 for a discussion of the general case.

We are mainly interested in the behaviour of $\mu_{fc}(E)$ and $\operatorname{Im} m_{fc}(E + \mathrm{i}\eta)$, for $E \in \mathbb{R}$ close to L_1 and L_2, respectively. We distinguish the cases $\lambda \leq 1$ and $\lambda > 1$:

For the former case, it was already pointed out in Ref. 5 (see also Refs. 31 and 36) that μ_{fc} has a square root behaviour near L_2, i.e., $\mu_{fc}(L_2 - \kappa) \sim \sqrt{\kappa}$, $\kappa \geq 0$, and similar for L_1.

For $\lambda > 1$, we will restrict our attention to Jacobi measures, a special class of measures whose densities are of the form

$$\mu(v) = Z^{-1}(1 + v)^{\mathrm{a}}(1 - v)^{\mathrm{b}} d(v)\chi_{[-1,1]}(v),$$

where $\mathrm{a}, \mathrm{b} > -1$, $d \in C^1([-1, 1])$ with $d(v) > 0$, $v \in [-1, 1]$ and the normalization constant Z is appropriately chosen so that μ becomes a probability density. Again, for simplicity, we will always assume that μ is centered. Note that we also admit exponents a, b smaller than zero, thus $\mu(v) \to \infty$ as $v \to \pm 1$ is allowed. As it turns out, the square root behaviour at the endpoint of the support persists for $\lambda > 1$, in case we have $-1 < \mathrm{a}, \mathrm{b} \leq 1$, respectively. However, if $\mathrm{a}, \mathrm{b} > 1$, there exists $\lambda_0 > 1$, such that for any $\lambda > \lambda_0$, we have $\mu_{fc}(L_1 - \kappa) \sim \kappa^{\mathrm{b}}$; for a precise statement see Lemma A.4.

2. Case $\lambda \leq 1$

In this subsection, we choose $\lambda \leq 1$. Adopting the proof of Proposition 2 from Ref. 36, we have the following result:

Lemma A.1. Let μ be a centered probability measure supported on $[-1, 1]$. Assume that μ has a continuous, strictly positive, bounded density $\mu(v)$ on $(-1, 1)$. Suppose that $0 \leq \lambda \leq 1$. Then, *there exits $L_1, L_2 \in \mathbb{R}$, with $L_1 < 0 < L_2$, such that the (rescaled) free convolution of μ with the semicircle law, μ_{fc}, satisfies*

$$\operatorname{supp} \mu_{fc} = [L_1, L_2]\,.$$

Moreover, denoting by κ_E the distance to the endpoints of the support of μ_{fc}, i.e.,

$$\kappa_E := \min\{|E - L_1|, |E - L_2|\},$$

we have

$$C^{-1}\sqrt{\kappa_E} \le \mu_{fc}(E) \le C\sqrt{\kappa_E}\,, \qquad E \in [L_1, L_2], \tag{A2}$$

for some constant $C \ge 1$.

We briefly outline how the proof in Ref. 36 can be adopted to our setting: We denote by $m_{fc}(z)$, $z \in \mathbb{C}^+$, the Stieltjes transform of the free convolution measure μ_{fc}. Define $\tau := z + m_{fc}(z)$ and consider instead of (A1) the equation $F(\tau) = z$, where

$$F(\tau) := \tau - \int_{-1}^{1} \frac{\mathrm{d}\mu(v)}{\lambda v - \tau}\,, \quad \tau \in \mathbb{C}^+. \tag{A3}$$

Note that $\lim_{y\searrow 0} \operatorname{Im} F(x + \mathrm{i}y) = -\pi\mu(x) < 0$, for $x \in (-\lambda, \lambda)$, since we have assumed that the density of μ is bounded and continuous, and strictly positive in the interval $(-1, 1)$. Thus F extends to a function on $\mathbb{R}$, which is continuous and bounded, except possibly at the point $\{\pm\lambda\}$. As shown in Ref. 36, the endpoints, (L_i), of the support of μ_{fc} are characterized as the real valued solutions, τ_i, with $|\tau_i| \ge \lambda$, of the equation $F'(\tau) = 0$ (L_i are then obtained by solving $\tau_i = L_i + m_{fc}(L_i)$). Setting

$$H(\tau) := \int_{-1}^{1} \frac{\mathrm{d}\mu(v)}{(\lambda v - \tau)^2}\,, \qquad \tau \in \mathbb{C}^+\,, \tag{A4}$$

a point $E \in \mathbb{R}$ is an endpoint of the support of μ_{fc}, if $H(\tau) = 1$, $|\tau| \ge \lambda$, $\tau = E + m_{fc}(E) \in \mathbb{R}$. Since $\lambda \le 1$ and μ is centered, we have from Jensen's inequality

$$H(\lambda) = \int_{-1}^{1} \frac{\mathrm{d}\mu(v)}{(\lambda v - \lambda)^2} > \frac{1}{\lambda^2}\frac{1}{\left(\int \mathrm{d}\mu(v)(v-1)\right)^2} = \frac{1}{\lambda^2} \ge 1. \tag{A5}$$

Here, the first inequality is strict since μ is absolutely continuous. Since $H(\tau)$ is monotone decreasing (on $\mathbb{R}$) as $|\tau| \to \infty$, we conclude that there are only two real solutions τ_1, τ_2. One then checks that the endpoints of the support of μ_{fc}, L_1, and L_2 satisfy $L_1 < -2$ and $L_2 > 2$. The square root behaviour of μ_{fc} at L_i, i.e., (A2), follows as in Ref. 35: It suffices to observe that $F''(\tau_i) \neq 0$, thus by the inverse function theorem, we have, for $z \in \mathbb{C}$ in a neighborhood of L_i, $F^{-1}(z) = \tau_i + c_i\sqrt{z - L_i}(1 + A_i(\sqrt{z - L_i}))$ (such that $\operatorname{Im} F^{-1}(z) \ge 0$, for $z \in \mathbb{C}^+$), for real constants $c_i \neq 0$ and analytic functions A_i, with $|A_i| \le 1$ in a neighborhood of zero. This concludes our discussion on the proof of Lemma A.1.

As an important corollary of the proof of Lemma A.1, we have the following stability bound already pointed out in Ref. 36:

Corollary A.2. Under the assumptions of Lemma A.1 there exist constants C, c > 0, *such that*

$$c \le |\lambda v - z - m_{fc}(z)| \le C\,, \quad z = E + \mathrm{i}\eta,$$

for any $\lambda v \in (-\lambda, \lambda)$ and $|E| \le E_0$, $0 < \eta \le 3$.

Proof. For the upper bound, we note that $|m_{fc}(z)| \le 1$, as follows from considering the imaginary part of m_{fc} in (A1). For the lower bound, note that in a neighborhood of L_i, $\operatorname{Re}(z + m_{fc}(z)) = \operatorname{Re}\tau(z) = \operatorname{Re}\tau_i + \mathcal{O}(|z - L_i|^{1/2})$. Since $|\tau_i| > 1$, $|\operatorname{Re}(z + m_{fc}(z))| > 1$, for $|z - L_i| < \epsilon$ for a sufficiently small $\epsilon > 0$. For $|\operatorname{Re} z| \ge |L_i| + (\epsilon/2)$, the estimate is trivial. In the region not covered by the two preceding estimates, we must have $\operatorname{Im}\tau > c$, thus $\operatorname{Im} m_{fc} + \eta > c$. The claim follows. $\square$

3. Case $\lambda > 1$

In this subsection, we choose for simplicity μ as a Jacobi measure, i.e., μ is described in terms of its density

$$\mu(v) = Z^{-1}(1+v)^{\mathrm{a}}(1-v)^{\mathrm{b}} d(v)\chi_{[-1,1]}(v)\,, \tag{A6}$$

where a, b > -1, $d \in C^1([-a, b])$ such that $d(v) > 0$, $v \in [-a, b]$ and Z is an appropriately chosen normalization constant such that μ is a probability measure. Below, we will assume, for simplicity of the arguments, that μ is centered, but this condition can easily be relaxed.

Lemma A.3. Let μ be a centered Jacobi measure. Suppose that $\lambda > 1$. If $-1 <$ a, b ≤ 1, *the results in Lemma A.1 and Corollary A.2 hold true.*

Proof. We can apply the same argument as in the proof of Lemma A.1. The only thing we need to prove is that

$$H(\lambda+\epsilon) = \int_{-1}^{1} \frac{\mathrm{d}\mu(v)}{(\lambda v - \lambda - \epsilon)^2} > 1,$$

for any sufficiently small $\epsilon > 0$, and a similar estimate for $H(-\lambda - \epsilon)$.

From the assumptions, we find that there exist constants $C, C_0 > 0$ such that $\mu(v) \geq C(1-v)^{\mathrm{b}} \geq C_0(1-v)$ for any $v \in (0, 1)$. Let $n := e^{1+\lambda^2 C_0^{-1}}$ and choose $\epsilon < 1/n$. Then, we have

$$H(\lambda+\epsilon) \geq C_0 \int_{1-(n-1)\epsilon/\lambda}^{1} \frac{(1-v)\mathrm{d}v}{(\lambda v - \lambda - \epsilon)^2} = \frac{C_0}{\lambda^2} \int_{\epsilon/\lambda}^{n\epsilon/\lambda} \frac{t - (\epsilon/\lambda)}{t^2}\mathrm{d}t = \frac{C_0}{\lambda^2}(\log n - 1 + \frac{1}{n}) > 1.$$

From the continuity of H, we get the desired results. The same argument applies to $H(-\lambda - \epsilon)$. □

For a, b > 1, we have the following result:

Lemma A.4. Let μ be a centered Jacobi measure with a, b > 1. *Define*

$$\lambda_2 := \left(\int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{(1-v)^2}\right)^{1/2}, \qquad \tau_2 := \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{1-v}\,.$$

Then, there exist $L_1 < 0 < L_2$ such that the support of μ_{fc} is $[L_1, L_2]$. Moreover,

i. if $\lambda < \lambda_2$, then for $0 \leq \kappa \leq L_2$,

$$C^{-1}\sqrt{\kappa} \leq \mu_{fc}(L_2 - \kappa) \leq C\sqrt{\kappa}\,, \tag{A7}$$

for some $C \geq 1$.

ii. If $\lambda > \lambda_2$, then $L_2 = \lambda + (\tau_2/\lambda)$ and, for $0 \leq \kappa \leq L_2$,

$$C^{-1}\kappa^{\mathrm{b}} \leq \mu_{fc}(L_2 - \kappa) \leq C\kappa^{\mathrm{b}}\,, \tag{A8}$$

for some $C \geq 1$.

Moreover, for $0 \leq E \leq E_0$, $0 < \eta \leq 2$, $z = E + \mathrm{i}\eta$, $v \in [-1, 1]$,

$$|\lambda v - z - m_{fc}(z)|$$

remains bounded from below in case i uniformly in z and v, but in case ii, it can be arbitrarily small as $v \to 1$, $E = L_2$, and $\eta \to 0$.

Similar statements hold for the lower endpoint L_1 of the support of μ_{fc}, with τ_2 and λ_2 replaced by

$$\lambda_1 := \left(\int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{(1+v)^2}\right)^{1/2}, \qquad \tau_1 := \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{1+v}\,. \tag{A9}$$

Proof. We first note that $0 < \lambda_2, \tau_2 < \infty$, for $\mathrm{b} > 1$. Since $\mu(v) > 0$, for $v \in (-1, 1)$, μ_{fc} is supported on a single interval. Consider now

$$H(\lambda) = \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{(\lambda v - \lambda)^2} = \frac{1}{\lambda^2}\int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{(v-1)^2} = \left(\frac{\lambda_2}{\lambda}\right)^2.$$

When $\lambda < \lambda_2$, we may follow the proof of Lemma A.1 to prove the claims in *i*.

We now choose $\lambda > \lambda_2$. We claim that there exists a unique continuous bounded curve γ in $\mathbb{C}^+$ on which $\operatorname{Im} F(\tau) = 0$. For $z \in \mathbb{C}^+$,

$$\operatorname{Im} F(\tau) = \operatorname{Im} \tau \left(1 - \int \frac{\mathrm{d}\mu(v)}{|\lambda v - \tau|^2}\right).$$

We know that the non-negative continuous function

$$\widetilde{H}(\tau) := \int \frac{\mathrm{d}\mu(v)}{|\lambda v - \tau|^2} = \int \frac{\mathrm{d}\mu(v)}{(\lambda v - \operatorname{Re}\tau)^2 + (\operatorname{Im}\tau)^2}, \qquad \tau \in \mathbb{C}^+ \cup (\mathbb{R}\backslash[-\lambda, \lambda]),$$

is monotonically decreasing in $\operatorname{Im}\tau$. Let $\tau = x + \mathrm{i}y$. For $x \in (-\lambda, \lambda)$, the continuity of μ implies that, as $y \searrow 0$, $y\widetilde{H}(x + \mathrm{i}y) \to \pi\mu(x) > 0$, hence $\widetilde{H}(x + \mathrm{i}y) \to \infty$. Since $\widetilde{H}(x + \mathrm{i}y)$ is monotonically decreasing as y increases and $\widetilde{H}(x + \mathrm{i}y) \to 0$ as $y \to \infty$, the equation $y = y\widetilde{H}(x + \mathrm{i}y)$ has a unique solution $0 < y < \infty$. The analyticity of F in the upper half plane then implies that on the interval $(-\lambda, \lambda)$ there exists a single bounded curve such that the imaginary part of F vanishes on it.

The endpoints of the support of μ_{fc} are characterized as the points where the curve γ approaches to the real line. Since $\widetilde{H}(\lambda) = H(\lambda) < 1$, the curve γ does not connect with the real axis on $\mathbb{R}^+\backslash(0, \lambda)$. Since this curve cannot end at some point where F is analytic, we can conclude that the curve approaches to λ on $\mathbb{R}^+$. When $\tau = \lambda$, we have

$$z = \tau - m_{fc}(z) = \lambda - \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{\lambda v - \lambda} = \lambda + \frac{\tau_2}{\lambda},$$

which corresponds to the endpoint, L_2, of the support of μ_{fc} on $\mathbb{R}^+$.

To prove (A8), let

$$\tau = \lambda - \lambda k + \mathrm{i}\lambda y, \qquad z = \lambda + \frac{\tau_2}{\lambda} - \kappa + \mathrm{i}\eta.$$

Considering the imaginary part of m_{fc}, we obtain

$$\lambda y - \eta = \operatorname{Im} m_{fc}(z) = \operatorname{Im} \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{\lambda v - \tau} = \frac{y}{\lambda}\int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{(v - 1 + k)^2 + y^2}. \tag{A10}$$

We claim that in the limit $\eta \searrow 0$,

$$y \sim (k + y)^{\mathrm{b}}, \tag{A11}$$

for $\kappa, y \ll 1$.

For the upper bound, we consider first the case $y < k$: Let $\epsilon = \min\{1/2, (\lambda^2/\lambda_2^2) - 1\}$, then we have

$$y\int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{(v - 1 + k)^2 + y^2} = y\left(\int_{-1}^{1-8\epsilon^{-1}k} + \int_{1-8\epsilon^{-1}k}^{1-k-y} + \int_{1-k-y}^{1-k+y} + \int_{1-k+y}^{1}\right)\frac{\mu(v)\mathrm{d}v}{(1 - v - k)^2 + y^2}. \tag{A12}$$

The first term in (A12) can be estimated as

$$\begin{aligned} y\int_{-1}^{1-8\epsilon^{-1}k} \frac{\mu(v)\mathrm{d}v}{(1 - v - k)^2 + y^2} &\le y\int_{-1}^{1-8\epsilon^{-1}k} \frac{\mu(v)\mathrm{d}v}{(1 - v - k)^2} \le y\left(1 + \frac{\epsilon}{2}\right)\int_{-1}^{1-8\epsilon^{-1}k} \frac{\mu(v)\mathrm{d}v}{(1 - v)^2} \\ &\le \left(1 + \frac{\epsilon}{2}\right)\lambda_2^2 y. \end{aligned} \tag{A13}$$

Here, we used that $v \leq 1 - 8\epsilon^{-1}k$ implies that $1 - v - k \geq (1 - \epsilon/8)(1 - v)$, hence

$$\frac{1}{(1-v-k)^2} \leq \left(1 - \frac{\epsilon}{8}\right)^{-2} \frac{1}{(1-v)^2} \leq \left(1 + \frac{\epsilon}{2}\right) \frac{1}{(1-v)^2}\,.$$

The second term in (A12) can be estimated as

$$y \int_{1-8\epsilon^{-1}k}^{1-k-y} \frac{\mu(v)\mathrm{d}v}{(1-v-k)^2 + y^2} \leq Cy \int_{1-8\epsilon^{-1}k}^{1-k-y} \frac{k^{\mathrm{b}}\mathrm{d}v}{(1-v-k)^2} \leq Ck^{\mathrm{b}}\,.$$

The third term in (A12) can be estimated as

$$y \int_{1-k-y}^{1-k+y} \frac{\mu(v)\mathrm{d}v}{(v-1+k)^2 + y^2} \leq Cy \int_{1-k-y}^{1-k+y} \frac{(k+y)^{\mathrm{b}}\mathrm{d}v}{y^2} \leq C(y+k)^{\mathrm{b}}.$$

The last term in (A12) can be estimated as

$$y \int_{1-k+y}^{1} \frac{\mu(v)\mathrm{d}v}{(1-v-k)^2 + y^2} \leq y \int_{1-k+y}^{1} \frac{(k-y)^{\mathrm{b}}\mathrm{d}v}{(1-v-k)^2} = y \int_{y}^{k} \frac{(k-y)^{\mathrm{b}}\mathrm{d}w}{w^2} \leq Ck^{\mathrm{b}}.$$

Thus, as $\eta \searrow 0$, we have that

$$\lambda y \leq \frac{1}{\lambda}\left(1 + \frac{\epsilon}{2}\right)\lambda_2^2 y + C(k+y)^{\mathrm{b}}.$$

Since

$$\lambda - \frac{1}{\lambda}\left(1 + \frac{\epsilon}{2}\right)\lambda_2^2 = \frac{\lambda_2^2}{\lambda}\left(\frac{\lambda^2}{\lambda_2^2} - 1 - \frac{\epsilon}{2}\right) \geq \frac{\epsilon\lambda_2^2}{2\lambda},$$

we obtain

$$y \leq C(k+y)^{\mathrm{b}},$$

provided $y < k$.

When $y \geq k$, we decompose the integral in (A12) as

$$y \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{(v-1+k)^2 + y^2} = y\left(\int_{-1}^{1-8\epsilon^{-1}k} + \int_{1-8\epsilon^{-1}k}^{1}\right) \frac{\mu(v)\mathrm{d}v}{(1-v-k)^2 + y^2}.$$

The first term is again estimated as in (A13). The second term can be estimated as

$$y \int_{1-8\epsilon^{-1}k}^{1} \frac{\mu(v)\mathrm{d}v}{(1-v-k)^2 + y^2} \leq Cy \int_{1-8\epsilon^{-1}k}^{1} \frac{k^{\mathrm{b}}\mathrm{d}v}{y^2} \leq Ck^{\mathrm{b}+1}y^{-1} \leq Cy^{\mathrm{b}}.$$

Following the argument we used for the case $y < k$, we find the relation $y \leq Cy^{\mathrm{b}}$ in this case. For sufficiently small y, this is impossible, so this case does not happen.

To complete the proof of (A11), we need a lower bound: Observe that

$$y \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{(v-1+k)^2 + y^2} \geq y \int_{1-k-y}^{1-k} \frac{\mu(v)\mathrm{d}v}{(v-1+k)^2 + y^2} \geq Cy \int_{1-k-y}^{1-k} \frac{\mu(v)\mathrm{d}v}{y^2} \geq Ck^{\mathrm{b}}, \qquad \text{(A14)}$$

and (A11) follows from (A10) and $k \geq y$. When $y, k \ll 1$, (A11) implies that $k \gg y$ and since $y \to C\mu_{fc}(L_2 - \kappa)$ as $\eta \to 0$, we have $\mu_{fc}(L_2 - \kappa) \sim y \sim k^{\mathrm{b}}$.

To compare k and κ, we consider the real part of m_{fc} and get

$$\kappa - \lambda k - \frac{\tau_2}{\lambda} = \operatorname{Re} m_{fc}(z) = \operatorname{Re} \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{\lambda v - \tau} = \frac{1}{\lambda} \int_{-1}^{1} \frac{(v-1+k)\mu(v)\mathrm{d}v}{(v-1+k)^2 + y^2}.$$

From the definition of τ_2, we find that

$$\kappa - \lambda k = \frac{1}{\lambda} \int_{-1}^{1} \left(\frac{(v-1+k)\mu(v)\mathrm{d}v}{(v-1+k)^2 + y^2} + \frac{\mu(v)\mathrm{d}v}{1-v}\right) = \frac{1}{\lambda} \int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{1-v} \cdot \frac{k(v-1) + k^2 + y^2}{(v-1+k)^2 + y^2}.$$

We now separate the integral and estimate each term as in (A14) and (A12). We then get

$$\int_{-1}^{1} \frac{k\mu(v)\mathrm{d}v}{(v-1+k)^2+y^2} \sim \frac{k}{y}(k+y)^{\mathrm{b}}$$

and

$$\frac{1}{\lambda}\int_{-1}^{1} \frac{\mu(v)\mathrm{d}v}{1-v}\cdot\frac{k^2+y^2}{(v-1+k)^2+y^2} \sim \frac{k^2+y^2}{y}(k+y)^{\mathrm{b}-1}.$$

Recalling that $y \sim k^{\mathrm{b}}$, when $y, k \ll 1$, we find that $\kappa - \lambda k = O(k)$. Therefore, we get

$$\mu_{fc}(L_2-\kappa) \sim y \sim k^{\mathrm{b}} \sim \kappa^{\mathrm{b}},$$

as $\kappa \searrow 0$. Finally, it is easy to see that $|\lambda v - z - m_{fc}(z)|$ is not bounded from below: Choosing $z = L_2$, we have $\mathrm{Im}(m_{fc}(L_2)) = 0$, but $\mathrm{Re}(\lambda v - L_2 - m_{fc}(L_2)) = \lambda v - \lambda$. This proves the claims in claim *ii*. □

4. Square root behaviour of m_{fc} and further stability bounds

In this subsection, we prove that the Stieltjes transform m_{fc} inherits the square root behaviour from μ_{fc}:

Lemma A.5. Assume that μ_{fc} has support $[L_1, L_2]$ and satisfies

$$C^{-1}\sqrt{\kappa} \le \mu_{fc}(L_2-\kappa) \le C\sqrt{\kappa}\,, \tag{A15}$$

$0 \le \kappa \le L_2$, $C \ge 1$. *Then,*

i. *for* $z = L_2 - \kappa + \mathrm{i}\eta$, *with* $0 \le \kappa \le L_2$ *and* $0 < \eta \le 2$, *we have,* $C \ge 1$,

$$C^{-1}\sqrt{\kappa+\eta} \le \mathrm{Im}\, m_{fc}(z) \le C\sqrt{\kappa+\eta}\,;$$

ii. *for* $z = L_2 + \kappa + \mathrm{i}\eta$, *with* $0 \le \kappa \le 1$ *and* $0 < \eta \le 2$, *we have,* $C \ge 1$,

$$C^{-1}\frac{\eta}{\sqrt{\kappa+\eta}} \le \mathrm{Im}\, m_{fc}(z) \le C\frac{\eta}{\sqrt{\kappa+\eta}}.$$

The analogous statements hold for $z = L_1 \pm \kappa + \mathrm{i}\eta$.

Proof. We start with the claim *i*: Notice that

$$\mathrm{Im}\, m_{fc}(z) = \mathrm{Im}\int \frac{\mathrm{d}\mu_{fc}(x)}{x-z} = \int \frac{\eta\,\mathrm{d}\mu_{fc}(x)}{(x-L_2+\kappa)^2+\eta^2}\,.$$

To prove the lower bound, consider the following cases:

Case 1. When $\kappa, \eta < 1/2$, computing the integral from $x = L_2 - \kappa - 2\eta$ to $x = L_2 - \kappa - \eta$, we find from (A15) that

$$\mathrm{Im}\, m_{fc}(z) = \int \frac{\eta\,\mathrm{d}\mu_{fc}(x)}{(x-L_2+\kappa)^2+\eta^2} \ge C\int_{L_2-\kappa-2\eta}^{L_2-\kappa-\eta} \frac{\eta\sqrt{\kappa+\eta}}{\eta^2}\mathrm{d}x \ge C\sqrt{\kappa+\eta}\,.$$

Case 2. When $\kappa \ge 1/2$, $\eta < 1/2$, we obtain from (A15) that

$$\mathrm{Im}\, m_{fc}(z) \ge C\int_{L_2-\kappa+\eta/8}^{L_2-\kappa+\eta/4} \frac{\eta\,\mathrm{d}\mu_{fc}(x)}{(x-L_2+\kappa)^2+\eta^2} \ge C\sqrt{\kappa}\int_{L_2-\kappa+\eta/8}^{L_2-\kappa+\eta/4}\frac{\eta\,\mathrm{d}x}{\eta^2} \ge C\sqrt{\kappa} \ge C\sqrt{\kappa+\eta}\,.$$

Case 3. When $\eta \ge 1/2$, we have the bound

$$\mathrm{Im}\, m_{fc}(z) = \int \frac{\eta\,\mathrm{d}\mu_{fc}(x)}{(x-L_2-\kappa)^2+\eta^2} \ge C\int \frac{\eta\,\mathrm{d}\mu_{fc}(x)}{\eta^2} = \frac{C}{\eta} \ge C\sqrt{\kappa+\eta}\,.$$

This proves the lower bound. To prove the upper bound, we consider the following cases:

Case 1. When $\eta < \kappa < 1/2$, from (A15) we have

$$\begin{aligned}\operatorname{Im} m_{fc}(z) &= \int \mathrm{d}\mu_{fc}(x) \frac{\eta}{(x - L_2 + \kappa)^2 + \eta^2} \\ &\le C\eta \int_{-L_1}^{L_2-\kappa-\eta} \frac{\sqrt{L_2 - x}}{(x - L_2 + \kappa)^2} \mathrm{d}x + C\eta \int_{L_2-\kappa-\eta}^{L_2-\kappa+\eta} \frac{\sqrt{\kappa+\eta}}{\eta^2} \mathrm{d}x \\ &\quad + C\eta \int_{L_2-\kappa+\eta}^{L_2} \frac{\sqrt{\kappa}}{(x - L_2 + \kappa)^2} \mathrm{d}x \\ &\le C\eta \int_{\eta}^{L_1+L_2-\kappa} \frac{\sqrt{y+\kappa}}{y^2} \mathrm{d}y + C\sqrt{\kappa+\eta} + C\eta \int_{\eta}^{\kappa} \frac{\sqrt{\kappa}}{y^2} \mathrm{d}y \le C\sqrt{\kappa+\eta}\,.\end{aligned}$$

Case 2. When $\kappa < \eta < 1/2$, a calculation similar to *Case 1* proves the same bound.

Case 3. When $\kappa \ge 1/2$, we have

$$\operatorname{Im} m_{fc}(z) \le C \int \frac{\eta\, \mathrm{d}x}{(x - L_2 + \kappa)^2 + \eta^2} \le C \le C\sqrt{\kappa+\eta}\,.$$

Case 4. When $\eta \ge 1/2$, we have the trivial bound

$$\operatorname{Im} m_{fc}(z) \le |m_{fc}(z)| \le \frac{1}{\eta} \le C\sqrt{\kappa+\eta}\,.$$

This completes the proof of statement *i*. To prove *ii*, we proceed similarly:

Case 1. When $\kappa > \eta$, computing the integral from $x = L_2 - \kappa$ to $x = L_2 - 2\kappa$, we get

$$\operatorname{Im} m_{fc}(z) = \int \frac{\eta\, \mathrm{d}\mu_{fc}(x)}{(x - L_2 - \kappa)^2 + \eta^2} \ge C \int_{L_2-\kappa}^{L_2-2\kappa} \frac{\eta\sqrt{\kappa}}{\kappa^2} \mathrm{d}x \ge \frac{C\eta}{\sqrt{\kappa}} \ge \frac{C\eta}{\sqrt{\eta+\kappa}}\,.$$

For the upper bound, we find that

$$\begin{aligned}\operatorname{Im} m_{fc}(z) &= \int \frac{\eta\, \mathrm{d}\mu_{fc}(x)}{(x - L_2 - \kappa)^2 + \eta^2} \le C\eta \int_{L_2-\kappa}^{L_2} \frac{\sqrt{\kappa}}{\kappa^2} \mathrm{d}x + C\eta \int_{-L_1}^{L_2-\kappa} \frac{\sqrt{L_2 - x}}{(x - L_2)^2} \mathrm{d}x \\ &\le \frac{C\eta}{\sqrt{\kappa}} + C\eta \int_{\kappa}^{L_1+L_2} \frac{\sqrt{y}}{y^2} \le \frac{C\eta}{\sqrt{\kappa}} \le \frac{C\eta}{\sqrt{\eta+\kappa}}.\end{aligned}$$

Case 2. When $\kappa \le \eta$, computing the integral from $x = L_2 - (\eta/2)$ to $x = L_2 - \eta$, we obtain

$$\operatorname{Im} m_{fc}(z) = \int \frac{\eta\, \mathrm{d}\mu_{fc}(x)}{(x - L_2 - \kappa)^2 + \eta^2} \ge C \int_{L_2-(\eta/2)}^{L_2-\eta} \frac{\eta\sqrt{\eta}}{\eta^2} \mathrm{d}x \ge C\sqrt{\eta} \ge \frac{C\eta}{\sqrt{\eta+\kappa}}\,.$$

For the upper bound, we find that

$$\begin{aligned}\operatorname{Im} m_{fc}(z) &= \int \frac{\eta\, \mathrm{d}d\mu_{fc}(x)}{(x - L_2 - \kappa)^2 + \eta^2} \le C\eta \int_{L_2-\eta}^{L_2} \frac{\sqrt{\eta}}{\eta^2} \mathrm{d}x + C\eta \int_{-L_1}^{L_2-\eta} \frac{\sqrt{L_2 - x}}{(x - L_2)^2} \mathrm{d}x \\ &\le C\sqrt{\eta} + C\eta \int_{\eta}^{L_1+L_2} \frac{\sqrt{y}}{y^2} \le C\sqrt{\eta} \le \frac{C\eta}{\sqrt{\eta+\kappa}}\,.\end{aligned}$$

This completes the proof of the lemma. □

Finally, we show that $|1 - R_2(z)| \sim \sqrt{\kappa_E + \eta}$ and $R_3(z) = \mathcal{O}(1)$; see (3.32) for the definitions.

Lemma A.6. Assume that μ_{fc} has support $[L_1, L_2]$ and satisfies

$$C^{-1}\sqrt{\kappa_E} \le \mu_{fc}(E) \le C\sqrt{\kappa_E}\,,$$

with $C \ge 1$, where $\kappa_E := \min\{|E - L_1|, |E - L_2|\}$, denotes the distance to the endpoints of the support of μ_{fc}. Moreover, assume the stability bound

$$c < |\lambda v - z - m_{fc}(z)| \le C_0,$$

with $C_0, c > 0$, for any $|E| \le E_0$, $0 < \eta \le 3$, and $|v| \le 1$. Then, we have the followings:

i. *There exists a constant $C \ge 1$ such that for any $|E| \le E_0$, $0 < \eta \le 3$,*

$$C^{-1}\sqrt{\kappa_E + \eta} \le \left| 1 - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} \right| \le C\sqrt{\kappa_E + \eta}.$$

ii. *There exists a constant C such that $|R_3(z)| \le C$ uniformly in $z \in \mathcal{D}_L$ and $\lambda \in \mathcal{D}_{\lambda_0}$. Moreover, there exist constants c and ϵ_0 such that $|R_3(z)| \ge c$ whenever $z \in \mathcal{D}_L$ satisfies $|z - L_i| < \epsilon_0$, $i = 1, 2$.*

Proof. Since $c \le |\lambda v - z - m_{fc}(z)|$, it is easy to see that $|R_3| < C$. Furthermore, it is proved in Ref. 36 that $R_3(L_2) > 0$. Since $R_3(z)$ is an analytic function of z in a neighborhood of L_i, $i = 1, 2$, this proves the second part of the lemma.

In order to prove the first part of the lemma, we first consider the following decomposition:

$$\left| 1 - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} \right| \le \left| 1 - \int \frac{\mathrm{d}\mu(v)}{|\lambda v - z - m_{fc}(z)|^2} \right| + \left| \int \frac{\mathrm{d}\mu(v)}{|\lambda v - z - m_{fc}(z)|^2} - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} \right|.$$

For the first term in the right side of the decomposition (4), we have

$$1 - \frac{\operatorname{Im} m_{fc}(z)}{\operatorname{Im}(z + m_{fc}(z))} = \frac{\eta}{\operatorname{Im}(z + m_{fc}(z))} \le \frac{\eta}{C\sqrt{\kappa_E + \eta}} \le C\sqrt{\kappa_E + \eta}\,,$$

if $E \in [L_1, L_2]$ and

$$1 - \frac{\operatorname{Im} m_{fc}(z)}{\operatorname{Im}(z + m_{fc}(z))} = \frac{\eta}{\operatorname{Im}(z + m_{fc}(z))} \le \frac{\eta}{C\eta/\sqrt{\kappa_E + \eta}} \le C\sqrt{\kappa_E + \eta},$$

if $E \in [L_1, L_2]^c$. Since

$$|\operatorname{Re}(\lambda v - z - m_{fc}(z))|\,, \operatorname{Im}(z + m_{fc}(z)) \le |z + m_{fc}(z)| + \lambda < C,$$

we also find that

$$\begin{aligned} &\left| \int \frac{\mathrm{d}\mu(v)}{|\lambda v - z - m_{fc}(z)|^2} - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} \right| \\ &\qquad = 2\left| \int \frac{(\operatorname{Im}(z + m_{fc}(z)))^2 + \mathrm{i}\operatorname{Re}(\lambda v - z - m_{fc}(z)) \cdot \operatorname{Im}(z + m_{fc}(z))}{|\lambda v - z - m_{fc}(z)|^4} \mathrm{d}\mu(v) \right| \\ &\qquad \le C\ \operatorname{Im}(z + m_{fc}(z)) \le C\sqrt{\kappa_E + \eta}\,. \end{aligned} \tag{A16}$$

Thus,

$$\left| 1 - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} \right| \le C\sqrt{\kappa_E + \eta},$$

which proves the upper bound.

For the lower bound, we first consider the case $E \in [L_1, L_2]$: If $|\operatorname{Re}(z + m_{fc}(z))| > \lambda$, we get

$$\begin{aligned} \left| \operatorname{Im} \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} \right| &= 2\left| \int \frac{\operatorname{Re}(\lambda v - z - m_{fc}(z)) \cdot \operatorname{Im}(z + m_{fc}(z))}{|\lambda v - z - m_{fc}(z)|^4} \mathrm{d}\mu(v) \right| \\ &= 2\int \frac{|\operatorname{Re}(\lambda v - z - m_{fc}(z))| \cdot \operatorname{Im}(z + m_{fc}(z))}{|\lambda v - z - m_{fc}(z)|^4} \mathrm{d}\mu(v) \\ &\ge C\operatorname{Im}(z + m_{fc}(z)) \ge C\sqrt{\kappa_E + \eta}\,. \end{aligned}$$

Hence,

$$\left| 1 - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} \right| \ge \left| \operatorname{Im} \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} \right| \ge C\sqrt{\kappa_E + \eta}\,.$$

If $|\mathrm{Re}(z+m_{fc}(z))| < \lambda$, then the stability bound $|\lambda v - z - m_{fc}(z)| > c$ implies that $\mathrm{Im}(z + m_{fc}(z)) > c$. Then, we get

$$\begin{aligned}\mathrm{Re}\int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2} &= \int \frac{[\mathrm{Re}(\lambda v - z - m_{fc}(z))]^2 - [\mathrm{Im}(z+m_{fc}(z))]^2}{|\lambda v - z - m_{fc}(z)|^4}\mathrm{d}\mu(v)\\ &\le \int \frac{\mathrm{d}\mu(v)}{|\lambda v - z - m_{fc}(z)|^2} - C.\end{aligned}$$

Thus,

$$\begin{aligned}\left|1 - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2}\right| &\ge 1 - \mathrm{Re}\int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2}\\ &\ge \int \frac{\mathrm{d}\mu(v)}{|\lambda v - z - m_{fc}(z)|^2} - \mathrm{Re}\int \frac{\mathrm{d}\mu(v)}{|\lambda v - z - m_{fc}(z)|^2}\\ &\ge C \ge C\sqrt{\kappa_E + \eta}.\end{aligned}$$

In case $E \in [L_1, L_2]^c$, we obtain a lower bound from

$$\left|1 - \int \frac{\mathrm{d}\mu(v)}{(\lambda v - z - m_{fc}(z))^2}\right| \ge 1 - \int \frac{\mathrm{d}\mu(v)}{|\lambda v - z - m_{fc}(z)|^2}$$

and

$$1 - \int \frac{\mathrm{d}\mu(v)}{|\lambda v - z - m_{fc}(z)|^2} = 1 - \frac{\mathrm{Im}\, m_{fc}(z)}{\mathrm{Im}(z + m_{fc}(z))} = \frac{\eta}{\mathrm{Im}(z+m_{fc}(z))} \ge \frac{C\eta}{\eta/\sqrt{\kappa_E+\eta}} = C\sqrt{\kappa_E + \eta}.$$

This completes the proof. □

[1] Anderson, G. W., Guionnet, A., and Zeitouni, O., *An Introduction to Random Matrices* (Cambridge University Press, 2010).

[2] Belinschi, S. T., Benaych-Georges, F., and Guionnet, A., "Regularization by free additive convolution, square and rectangular cases," Complex Anal. Oper. Theory **3**, 611–660 (2009).

[3] Belinschi, S. T. and Bercovici, H., "A new approach to subordination results in free probability," J. Anal. Math. **101**, 357–365 (2007).

[4] Belinschi, S. T., Bercovici, H., Capitaine, M., and Février, M., "Outliers in the spectrum of large deformed unitarily invariant models," preprint arXiv:1207.5443 (2012).

[5] Biane, P., "On the free convolution with a semi-circular distribution," Indiana Univ. Math. J. **46**, 705–718 (1997).

[6] Biane, P., "Processes with free increments," Math. Z. **227**, 143–174 (1998).

[7] Bryc, W., Dembo, A., and Jiang, T., "Spectral measure of large random Hankel, Markov and Toeplitz matrices," Ann. Probab. **34**, 1–38 (2006).

[8] Capitaine, M., Donati-Martin, C., Féral, D., and Février, M., "Free convolution with a semi-circular distribution and eigenvalues of spiked deformations of Wigner matrices," Electron. J. Probab. **16**, 1750–1792 (2011).

[9] Chistyakov, G. P. and Götze, F., "The arithmetic of distributions in free probability theory," Cent. Eur. J. Math. **9**, 997–1050 (2011).

[10] Dyson, F., "A Brownian motion model for the eigenvalues of a random matrix," J. Math. Phys. **3**, 1191 (1962).

[11] Erdős, L., "Universality of Wigner random matrices: A survey of recent results," preprint arXiv:1004.0861v2 (2010).

[12] Erdős, L., Knowles, A., Yau, H.-T., and Yin, J., "Spectral statistics of Erdős-Rényi graphs I: Local semicircle law," Ann. Probab. **41**(3B), 2279–2375 (2013).

[13] Erdős, L., Knowles, A., Yau, H.-T., and Yin, J., "The local semicircle law for a general class of random matrices," Electron. J. Probab. **18**(59), 1–58 (2013).

[14] Erdős, L., Knowles, A., and Yau, H.-T., "Averaging Fluctuations in resolvents of random band matrices," Ann. Henri Poincare (published online); preprint arXiv:1205.5664 (2012).

[15] Erdős, L., Ramirez, J., Schlein, B., and Yau, H.-T., "Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation," Electron. J. Probab. **15**(18), 526–603 (2010).

[16] Erdős, L., Schlein, B., and Yau, H.-T., "Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices," Ann. Probab. **37**, 815–852 (2009).

[17] Erdős, L., Schlein, B., and Yau, H.-T., "Local semicircle law and complete delocalization for Wigner random matrices," Commun. Math. Phys. **287**, 641–655 (2009).

[18] Erdős, L., Schlein, B., and Yau, H.-T., "Wegner estimate and level repulsion for Wigner random matrices," Int. Math. Res. Notices **2010**, 436–479.

[19] Erdős, L., Schlein, B., and Yau, H.-T., "Universality of random matrices and local relaxation flow," Invent. Math. **185**, 75–119 (2011).

[20] Erdős, L., Schlein, B., Yau, H.-T., and Yin, J., "The local relaxation flow approach to universality of the local statistics for random matrices," Ann. I.H.P. Probab. Stat. **48**, 1–46 (2012).

[21] Erdős, L. and Yau, H.-T., "Universality of local spectral statistics of random matrices," Bull. Am. Math. Soc. **49**, 377–414 (2012).
[22] Erdős, L., Yau, H.-T., and Yin, J., "Bulk universality for generalized Wigner matrices," Probab. Theory Relat. Fields **154**, 341–407 (2012).
[23] Erdős, L., Yau, H.-T., and Yin, J., "Universality for generalized Wigner matrices with Bernoulli distribution," J. Comb. **2**, 15–81 (2012).
[24] Erdős, L., Yau, H.-T., and Yin, J., "Rigidity of eigenvalues of generalized Wigner matrices," Adv. Math. **229**, 1435–1515 (2012).
[25] Forrester, P. J. and Nagao, T., "Correlations for the circular Dyson brownian motion model with Poisson initial conditions," Nucl. Phys. B **532**, 733–752 (1998).
[26] Hiai, F. and Petz, D., *The Semicircle Law, Free Random Variables and Entropy* (American Mathematical Society, 2006).
[27] Johansson, K., "Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices," Commun. Math. Phys. **215**, 683–705 (2001).
[28] Johansson, K., "From gumbel to Tracy-Widom," Probab. Theory Relat. Fields **138**, 75–112 (2007).
[29] Kargin, V., "Subordination of the resolvent for a sum of random matrices," preprint arXiv:1109.5818 (2011).
[30] Nica, A. and Speicher, R., *Lectures on the Combinatorics of Free Probability* (Cambridge University Press, 2006).
[31] Olver, S. and Nadakuditi, R. R., "Numerical computation of convolutions in free probability theory," preprint arXiv:1203.1958 (2012).
[32] Pandey, A., "Statistical properties of many-particle spectra. IV. New ensembles by Stieltjes transform methods," Ann. Phys. **134**, 110–127 (1981).
[33] Pastur, L., "On the spectrum of random matrices," Theor. Math. Phys. **10**, 67–74 (1972).
[34] Pastur, L. and Vasilchuk, V., "On the law of addition of random matrices," Commun. Math. Phys. **214**, 249–286 (2000).
[35] Shcherbina, T., "On universality of bulk local regime of the deformed Gaussian unitary ensemble," Math. Phys., Anal. Geom. **5**, 396–433 (2009).
[36] Shcherbina, T., "On universality of local edge regime for the deformed Gaussian unitary ensemble," J. Stat. Phys. **143**, 455–481 (2011).
[37] Tao, T. and Vu, V., "Random matrices: Universality of the local eigenvalue statistics," Acta Math **206**, 127–204 (2011).
[38] Voiculescu, D. V., "The analogues of entropy and of Fisher's information measure in free probability theory. I," Commun. Math. Phys. **155**, 71–92 (1993).
[39] Voiculescu, D. V., Dykema, K. J., and Nica, A., *Free Random Variables: A Noncommutative Probability Approach to Free Products with Applications to Random Matrices, Operator Algebras and Harmonic Analysis on Free Groups* (American Mathematical Society, 1992).
[40] Wigner, E. P., "Characteristic vectors of bordered matrices with infinite dimensions," Ann. Math. **62**, 548–564 (1955).