

Eigenvector distribution of Wigner matrices

Antti Knowles · Jun Yin

Received: 22 March 2011 / Revised: 8 November 2011 / Published online: 14 December 2011
© Springer-Verlag 2011

Abstract We consider $N \times N$ Hermitian or symmetric random matrices with independent entries. The distribution of the (i, j) -th matrix element is given by a probability measure ν_{ij} whose first two moments coincide with those of the corresponding Gaussian ensemble. We prove that the joint probability distribution of the components of eigenvectors associated with eigenvalues close to the spectral edge agrees with that of the corresponding Gaussian ensemble. For eigenvectors associated with bulk eigenvalues, the same conclusion holds provided the first four moments of the distribution ν_{ij} coincide with those of the corresponding Gaussian ensemble. More generally, we prove that the joint eigenvector–eigenvalue distributions near the spectral edge of two generalized Wigner ensembles agree, provided that the first two moments of the entries match and that one of the ensembles satisfies a level repulsion estimate. If in addition the first four moments match then this result holds also in the bulk.

Keywords Random matrix · Universality · Eigenvector distribution

Mathematics Subject Classification (2010) 15B52 · 82B44

1 Introduction

The universality of random matrices can be roughly divided into the bulk universality in the interior of the spectrum and the edge universality near the spectral edge. Over the

A. Knowles was partially supported by NSF grant DMS-0757425 and J. Yin was partially supported by NSF grant DMS-1001655.

A. Knowles (✉) · J. Yin
Department of Mathematics, Harvard University, Cambridge, MA 02138, USA
e-mail: knowles@math.harvard.edu

J. Yin
e-mail: jyin@math.harvard.edu

past two decades, spectacular progress on bulk and edge universality has been made for invariant ensembles, see e.g. [2, 5, 6, 22] and [1, 3, 4] for a review. For non-invariant ensembles with i.i.d. matrix elements (*Standard Wigner ensembles*), edge universality can be proved via the moment method and its various generalizations; see e.g. [23–25]. In order to establish bulk universality, a new approach was developed in a series of papers [9–12, 14–17] based on three basic ingredients: (1) A local semicircle law—a precise estimate of the local eigenvalue density down to energy scales containing around N^ε eigenvalues. (2) The eigenvalue distribution of Gaussian divisible ensembles via an estimate on the rate of decay to local equilibrium of the Dyson Brownian motion [7]. (3) A density argument which shows that for any probability distribution there exists a Gaussian divisible distribution with identical eigenvalue statistics down to scales $1/N$. In [17], edge universality is established as a corollary of this approach. It asserts that, near the spectral edge, the eigenvalue distributions of two generalized Wigner ensembles are the same provided the first two moments of the two ensembles match.

Another approach to both bulk and edge universality was developed in [20, 26, 27]. Using this approach, the authors show that the eigenvalue distributions of two standard Wigner ensembles are the same in the bulk, provided that the first four moments match. They also prove a similar result at the edge, assuming that the first two moments match and the third moments vanish.

In this paper, partly based on the approach of [17], we extend edge universality to eigenvectors associated with eigenvalues near the spectral edge, assuming the matching of the first two moments of the matrix entries. We prove that, under the same two-moment condition as in [17], the edge eigenvectors of Hermitian and symmetric Wigner matrices have the same joint distribution as those of the corresponding Gaussian ensembles. The joint distribution of the eigenvectors of Gaussian ensembles is well known and can be easily computed. More generally, we prove that near the spectral edge the joint eigenvector–eigenvalue distributions of two generalized Wigner matrix ensembles coincide provided that the first two moments of the ensembles match and one of the ensembles satisfies a level repulsion condition.

We also prove similar results in the bulk, under the stronger assumption that the first four moments of the two ensembles match. In particular, we extend the result of [26] to cover the universality of bulk eigenvectors.

1.1 Setup

We now introduce the basic setup and notations. Let $H^\nu \equiv H = (h_{ij})_{i,j=1}^N$ be an $N \times N$ Hermitian or symmetric matrix whose upper-triangular matrix elements $h_{ij} = \bar{h}_{ji}$, $i \leq j$, are independent random variables with law ν_{ij} having mean zero and variance σ_{ij}^2 :

$$\mathbb{E} h_{ij} = 0, \quad \sigma_{ij}^2 := \mathbb{E} |h_{ij}|^2. \quad (1.1)$$

The law ν_{ij} and its variance σ_{ij}^2 may depend on N , but we omit this fact in the notation. We denote by $B := (\sigma_{ij}^2)_{i,j=1}^N$ the matrix of the variances. We shall always make the following three assumptions on H .

(A) For any fixed j we have

$$\sum_{i=1}^N \sigma_{ij}^2 = 1. \tag{1.2}$$

Thus B is symmetric and doubly stochastic and, in particular, satisfies $-1 \leq B \leq 1$.

(B) There exists constants $\delta_- > 0$ and $\delta_+ > 0$, independent of N , such that 1 is a simple eigenvalue of B and

$$\text{spec}(B) \subset [-1 + \delta_-, 1 - \delta_+] \cup \{1\}.$$

(C) There exists a constant C_0 , independent of N , such that $\sigma_{ij}^2 \leq C_0 N^{-1}$ for all $i, j = 1, \dots, N$.

Examples of matrices satisfying Assumptions (A)–(C) include Wigner matrices, Wigner matrices whose diagonal elements are set to zero, generalized Wigner matrices, and band matrices whose band width is of order cN for some $c > 0$. See [17], Sect. 2, for more details on these examples.

In our normalization, the matrix entries h_{ij} have a typical variance of order N^{-1} . It is well known that in this normalization the empirical eigenvalue density converges to the Wigner semicircle law $\varrho_{sc}(E) dE$ with density

$$\varrho_{sc}(E) := \frac{1}{2\pi} \sqrt{(4 - E^2)_+} \quad \text{for } E \in \mathbb{R}. \tag{1.3}$$

In particular, the spectral edge is located at ± 2 . We denote the ordered eigenvalues of H by $\lambda_1 \leq \dots \leq \lambda_N$, and their associated eigenvectors by $\mathbf{u}_1, \dots, \mathbf{u}_N$. The eigenvectors are ℓ^2 -normalized. We use the notation $\mathbf{u}_\alpha = (u_\alpha(i))_{i=1}^N$ for the components of the vector \mathbf{u}_α .

Our analysis relies on a notion of high probability which involves logarithmic factors of N . The following definitions introduce convenient shorthands.

Definition 1.1 We set $L \equiv L_N := A_0 \log \log N$ for some fixed A_0 as well as $\varphi \equiv \varphi_N := (\log N)^{\log \log N}$.

Definition 1.2 We say that an N -dependent event Ω holds *with high probability* if $\mathbb{P}(\Omega) \geq 1 - e^{-\varphi^c}$ for large enough N and some $c > 0$ independent of N .

A key assumption for our result is the following level repulsion condition, which is in particular satisfied by the Gaussian ensembles (see Remark 1.5 below). Consider a spectral window whose size is much smaller than the typical eigenvalue separation. Roughly, the level repulsion condition says that the probability of finding more than one eigenvalue in this window is much smaller than the probability of finding precisely one eigenvalue. In order to state the level repulsion condition, we introduce the following counting function. For any $E_1 \leq E_2$ we denote the number of eigenvalues in $[E_1, E_2]$ by

$$\mathcal{N}(E_1, E_2) := \#\{j : E_1 \leq \lambda_j \leq E_2\}.$$

Definition 1.3 (*Level repulsion at the edge*) The ensemble H is said to satisfy *level repulsion at the edge* if, for any $C > 0$, there is an $\alpha_0 > 0$ such that the following holds. For any α satisfying $0 < \alpha \leq \alpha_0$ there exists a $\delta > 0$ such that

$$\mathbb{P} \left(\mathcal{N}(E - N^{-2/3-\alpha}, E + N^{-2/3-\alpha}) \geq 2 \right) \leq N^{-\alpha-\delta} \tag{1.4}$$

for all E satisfying $|E + 2| \leq N^{-2/3}\varphi^C$

Definition 1.4 (*Level repulsion in the bulk*) The ensemble H is said to satisfy *level repulsion in the bulk* if, for any $\kappa > 0$, there is an $\alpha_0 > 0$ such that the following holds. For any α satisfying $0 < \alpha \leq \alpha_0$ there exists a $\delta > 0$ such that

$$\mathbb{P} \left(\mathcal{N}(E - N^{-1-\alpha}, E + N^{-1-\alpha}) \geq 2 \right) \leq N^{-\alpha-\delta} \tag{1.5}$$

for all $E \in [-2 + \kappa, 2 - \kappa]$.

Remark 1.5 Both the Gaussian unitary ensemble (GUE) and the Gaussian orthogonal ensemble (GOE) satisfy level repulsion in sense of Definitions 1.3 and 1.4. This can be established for instance as follows; see [1], Sects. 3.5 and 3.7, and in particular Lemmas 3.5.1 and 3.7.2, for full details. For GUE and GOE, the correlation functions can be explicitly expressed in terms of Hermite polynomials. Using Laplace’s method, one may then derive the large- N asymptotics of the correlation functions, from which (1.4) and (1.5) immediately follow. (Note that in [1], the exponent of N in the error estimates was not tracked in order to simplify the presentation.)

In the more general case of Wigner matrices, level repulsion in the bulk, (1.5), was proved for matrices with smooth distributions in [11] and without a smoothness assumption in [26].

We shall use the level repulsion condition of Definition 1.3 to estimate the probability of finding two eigenvalues closer to each other than the typical eigenvalue separation. For definiteness, we formulate this estimate at the lower spectral edge -2 . By partitioning the interval

$$\left[-2 - N^{-2/3}\varphi^C, -2 + N^{-2/3}\varphi^C \right]$$

into $O(\varphi^C N^\alpha)$ subintervals of size $N^{-2/3-\alpha}$, we get from (1.4) that for any sufficiently small α there exists a $\delta > 0$ such that

$$\mathbb{P} \left(\text{there exists } E \text{ with } |E + 2| \leq N^{-2/3}\varphi^C \text{ such that } \mathcal{N}(E - N^{-2/3-\alpha}, E + N^{-2/3-\alpha}) \geq 2 \right) \leq N^{-\delta}. \tag{1.6}$$

A similar result can be derived in the bulk using (1.5).

1.2 Results

Before stating our main results, we recall the definition of the classical eigenvalue locations. Let

$$n_{sc}(E) := \int_{-\infty}^E \varrho_{sc}(x) dx \tag{1.7}$$

be the integrated distribution function of the semicircle law. We use $\gamma_\alpha \equiv \gamma_{\alpha,N}$ to denote the classical location of the α -th eigenvalue under the semicircle law, defined through

$$n_{sc}(\gamma_\alpha) = \frac{\alpha}{N}. \tag{1.8}$$

To avoid unnecessary technicalities in the presentation, we shall assume that the entries h_{ij} of H have uniform subexponential decay, i.e.

$$\mathbb{P}(|h_{ij}| \geq x \sigma_{ij}) \leq \vartheta^{-1} \exp(-x^\vartheta) \tag{1.9}$$

where $\vartheta > 0$ is some fixed constant. As observed in [8], Sect. 7, one may easily check that all of our results hold provided the subexponential condition (1.9) is replaced with the weaker assumption that there is a constant C such that

$$\mathbb{E}|h_{ij} \sigma_{ij}^{-1}|^{C_0} \leq C,$$

where C_0 is a large universal constant.

Our main result on the distributions of edge eigenvectors is the following theorem.

Theorem 1.6 (Universality of edge eigenvectors) *Let H^V and H^W both satisfy Assumptions (A)–(C) as well as the uniform subexponential decay condition (1.9). Let \mathbb{E}^V and \mathbb{E}^W denote the expectations with respect to these collections of random variables. Suppose that the level repulsion estimate (1.4) holds for the ensemble H^V . Assume that the first two moments of the entries of H^V and H^W are the same, i.e.*

$$\mathbb{E}^V \bar{h}_{ij}^l h_{ij}^u = \mathbb{E}^W \bar{h}_{ij}^l h_{ij}^u \quad \text{for } 0 \leq l + u \leq 2. \tag{1.10}$$

Let ρ be a positive constant. Then for any integer k and any choice of indices $i_1, \dots, i_k, j_1, \dots, j_k, \beta_1, \dots, \beta_k$ and $\alpha_1, \dots, \alpha_k$ with $\min(|\alpha_l|, |\alpha_l - N|) + \min(|\beta_l|, |\beta_l - N|) \leq \varphi_N^\rho$ for all l we have

$$\lim_{N \rightarrow \infty} [\mathbb{E}^V - \mathbb{E}^W] \theta \left(N^{2/3}(\lambda_{\beta_1} - \gamma_{\beta_1}), \dots, N^{2/3}(\lambda_{\beta_k} - \gamma_{\beta_k}); N \bar{u}_{\alpha_1}(i_1) u_{\alpha_1}(j_1), \dots, N \bar{u}_{\alpha_k}(i_k) u_{\alpha_k}(j_k) \right) = 0, \tag{1.11}$$

where θ is a smooth function that satisfies

$$|\partial^n \theta(x)| \leq C(1 + |x|)^C \tag{1.12}$$

for some arbitrary C and all $n \in \mathbb{N}^{2k}$ satisfying $|n| \leq 3$. The convergence is uniform in all the parameters $i_l, j_l, \alpha_l, \beta_l$ satisfying the above conditions.

Remark 1.7 The scaling in front of the arguments in (1.11) is the natural scaling near the spectral edge. Indeed, for e.g. GUE or GOE it is known (see e.g. [1]) that

$(\lambda_\beta - \gamma_\beta) \sim N^{-2/3}$ near the edge, and that $u_\alpha(i) \sim N^{1/2}$ (complete delocalization of eigenvectors).

Remark 1.8 The form (1.11) characterizes the distribution of the edge eigenvectors completely. Choosing $i_l = j_l$ yields the modulus $|u_{\alpha_l}(i_l)|^2$; fixing i_l and varying j_l gives the relative phases of the entries of the vector \mathbf{u}_{α_l} , which is only defined up to a global phase.

Remark 1.9 Theorem 1.6 and Remark 1.5 imply that the joint eigenvector–eigenvalue distribution of Hermitian Wigner matrices agrees with that of GUE. In the case of GUE, it is well known that the joint distribution of the eigenvalues is given by the Airy kernel [29]. The eigenvectors are independent of the eigenvalues, and the matrix $(u_\alpha(i))_{\alpha,i}$ of the eigenvector entries is distributed according to the Haar measure on the unitary group $U(N)$. In particular, any eigenvector \mathbf{u}_α is uniformly distributed on the unit $(N - 1)$ -sphere.

Similarly, Theorem 1.6 and Remark 1.5 imply that the joint eigenvector–eigenvalue distribution of symmetric Wigner matrices agrees with that of GOE. Results similar to those outlined above on the eigenvector–eigenvalue distribution of GUE hold for GOE.

The universality of the eigenvalue distributions near the edge was already proved in [17] under the assumption that the first two moments of the matrix entries match, and in [27] under the additional assumption that the third moments vanish. Note that Theorem 1.6 holds in a stronger sense than the result in [17]: it holds for probability density functions, not just the distribution functions.

In the bulk, a result similar to Theorem 1.6 holds under the stronger assumption that four, instead of two, moments of the matrix entries match.

Theorem 1.10 (Universality of bulk eigenvectors) *Let $H^\mathbf{v}$ and $H^\mathbf{w}$ both satisfy Assumptions (A)–(C) as well as the uniform subexponential decay condition (1.9). Suppose that the level repulsion estimate (1.5) holds for the ensemble $H^\mathbf{v}$. Suppose moreover that the first four off-diagonal moments of $H^\mathbf{v}$ and $H^\mathbf{w}$ are the same, i.e.*

$$\mathbb{E}^\mathbf{v} \bar{h}_{ij}^l h_{ij}^u = \mathbb{E}^\mathbf{w} \bar{h}_{ij}^l h_{ij}^u \text{ for } i \neq j \text{ and } 0 \leq l + u \leq 4, \tag{1.13}$$

and that the first two diagonal moments of $H^\mathbf{v}$ and $H^\mathbf{w}$ are the same, i.e.

$$\mathbb{E}^\mathbf{v} \bar{h}_{ii}^l h_{ii}^u = \mathbb{E}^\mathbf{w} \bar{h}_{ii}^l h_{ii}^u \text{ for } 0 \leq l + u \leq 2. \tag{1.14}$$

Let $\rho > 0$ be fixed. Then for any integer k and any choice of indices $i_1, \dots, i_k, j_1, \dots, j_k$, as well as $\rho N \leq \alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k \leq (1 - \rho)N$, we have

$$\lim_{N \rightarrow \infty} [\mathbb{E}^\mathbf{v} - \mathbb{E}^\mathbf{w}] \theta \left(N(\lambda_{\beta_1} - \gamma_{\beta_1}), \dots, N(\lambda_{\beta_k} - \gamma_{\beta_k}); N\bar{u}_{\alpha_1}(i_1)u_{\alpha_1}(j_1), \dots, N\bar{u}_{\alpha_k}(i_k)u_{\alpha_k}(j_k) \right) = 0, \tag{1.15}$$

where θ is a smooth function that satisfies

$$|\partial^n \theta(x)| \leq C(1 + |x|)^C \tag{1.16}$$

for some arbitrary C and all $n \in \mathbb{N}^{2k}$ satisfying $|n| \leq 5$. The convergence is uniform in all the parameters $i_l, j_l, \alpha_l, \beta_l$ satisfying the above conditions.

The universality restricted to the bulk eigenvalues only has been previously established in several works. The following list provides a summary. Note that the small-scale statistics of the eigenvalues may be studied using *correlation functions*, which depend only on eigenvalue differences, or using *joint distribution functions*, as in (1.11) and (1.15), which in addition contain information about the eigenvalue locations.

- (i) In [17], bulk universality for generalized Wigner matrices was proved in the sense that correlation functions of bulk eigenvalues, averaged over a spectral window of size N^ε , converge to those of the corresponding Gaussian ensemble.
- (ii) In [26] the statement (1.15) on distribution functions, restricted to eigenvalues only, was proved for Hermitian and symmetric Wigner matrices for the case where the first four moments match as in (1.13).
- (iii) For the case of Hermitian Wigner matrices with a finite Gaussian component, it was proved in [19] that the correlation functions converge to those of GUE.
- (iv) In [18], the joint distribution function of the eigenvalues of GUE was computed. This result was extended to cover GOE in [21].

Note that (ii) and (iii) together imply the universality of the joint distribution of eigenvalues for Hermitian Wigner matrices, for which the first three moments match those of GUE and the distribution is supported on at least three points. Moreover, combining (ii) and (iv) allows one to compute the eigenvalue distribution of Hermitian and symmetric Wigner matrices, provided the four first moments match those of GUE/GOE.

Thus, Theorem 1.15 extends the results of [17] to distribution functions of individual eigenvalues as well as to eigenvectors.

Remark 1.11 A while after this paper was posted online, a result similar to Theorem 1.10 appeared in [28]. Its proof relies on a different method. The hypotheses of [28] are similar to those of Theorem 1.10, with the two following exceptions. The result of [28] is restricted to Wigner matrices instead of the generalized Wigner matrices defined by Assumptions (A)–(C). Moreover, in [28] the derivatives of the observable θ are required to be uniformly bounded in x , where this uniform bound may grow slowly with N . This latter restriction allows the authors of [28] to let k grow slowly with N .

While the results of [28] apply to eigenvectors near the spectral edge, the matching of four moments (as in Theorem 1.10) is also required for this case. As shown in Theorem 1.6, the universality of edge eigenvectors in fact only requires the first two moments to match.

1.3 Outline of the proof

The main idea behind our proof is to express the eigenvector components using matrix elements of the Green function $G(z) = (H - z)^{-1}$. To this end, we use the identity

$$\sum_{\beta} \frac{\eta/\pi}{(E - \lambda_{\beta})^2 + \eta^2} N \bar{u}_{\beta}(i) u_{\beta}(j) = \frac{N}{2\pi i} (G_{ij}(E + i\eta) - G_{ij}(E - i\eta)), \quad (1.17)$$

where $\eta > 0$. Using a good control on the matrix elements of $G(z)$, we may then apply a Green function comparison argument (similar to the Lindeberg replacement strategy) to complete the proof. For definiteness, let us consider a single eigenvalue λ_α located close to the spectral edge -2 .

In a first step, we write $N\bar{u}_\alpha(i)u_\alpha(j)$ as an integral of (1.17) over an appropriately chosen (random) domain, up to a negligible error term. We choose η in (1.17) to be much smaller than the typical eigenvalue separation, i.e. we set $\eta = N^{-2/3-\varepsilon}$ for some small $\varepsilon > 0$. Note that the fraction on the left-hand side of (1.17) is an approximate delta function on the scale η . Then the idea is to integrate (1.17) over the interval $[\lambda_\alpha - \varphi^C \eta, \lambda_\alpha + \varphi^C \eta]$ for some large enough constant C . For technical reasons related to the Green function comparison (the third step below), it turns out to be advantageous to replace the above interval with $\mathcal{I} := [\lambda_{\alpha-1} + \varphi^C \eta, \lambda_\alpha + \varphi^C \eta]$. Using eigenvalue repulsion, we infer that, with sufficiently high probability, the eigenvalues $\lambda_{\alpha-1}$ and $\lambda_{\alpha+1}$ are located at a distance greater than $\varphi^C \eta$ from λ_α . Therefore the E -integration over \mathcal{I} of the right-hand side of (1.17) yields $N\bar{u}_\alpha(i)u_\alpha(j)$ up to a negligible error term.

In a second step, we replace the sharp indicator function $\mathbf{1}(E \in \mathcal{I})$ with a smoothed indicator function expressed in terms of the Green function G . This is necessary for the Green function comparison argument, which requires all H -dependence to be expressed using Green functions. To that end, we choose a scale $\tilde{\eta} := N^{-2/3-6\varepsilon} \ll \eta$ and write

$$\mathbf{1}(E \in \mathcal{I}) \approx q \left[\text{Tr}(\mathbf{1}_{[E_L, E-\varphi^C \eta]} * \theta_{\tilde{\eta}})(H) \right] \tag{1.18}$$

where the error is negligible. Here $E_L = -2 - \varphi^C N^{-2/3}$, q is a smooth function equal to 1 in the $1/3$ -neighbourhood of $\alpha - 1$ and vanishing outside the $2/3$ -neighbourhood of $\alpha - 1$, and $\theta_{\tilde{\eta}}$ is the approximate delta function defined in (2.16) below. Thanks to the special form of the right-hand side of (2.16), we have $\theta_{\tilde{\eta}}(H) = \frac{1}{\pi} \text{Im } G(i\tilde{\eta})$. Hence the argument on right-hand side of (1.18) may be expressed as an integral over Green functions. Thus we have expressed $N\bar{u}_\alpha(i)u_\alpha(j)$ using matrix elements of G alone. Note that the above choice of \mathcal{I} was made precisely so as to make the right-hand side of (1.18) a simple function of G .

In a third step, we use a Green function comparison argument to compare the distributions of $N\bar{u}_\alpha(i)u_\alpha(j)$ under the two ensembles H^v and H^w . The basic strategy is similar to [17], but requires a more involved analysis of the resolvent expansion. The reason for this is that we need to exploit the smallness associated with off-diagonal elements of G , which requires us to keep track of their number in the power counting. This bookkeeping is complicated by the presence of the two fixed indices i and j . Another important ingredient in the error estimates of the Green function comparison argument is the restriction of the integration over E to a deterministic interval of size $\varphi^C N^{-2/3}$ around -2 . This can be done with negligible errors using the eigenvalue rigidity proved in [17]; see Theorem 2.2.

The above proof may be easily generalized to multiple eigenvector components as well as to eigenvalues; this allows us to consider observables of the form given in (1.11). The necessary changes are given in Sect. 4.

The proof for bulk eigenvectors is similar, with two major differences. At the edge, the convolution integral on the right-hand side of (1.18) was over a domain of size

$\varphi^C N^{-2/3}$. If the same expression were used in the bulk, this size would be $O(1)$ (since E is separated from the spectral edge -2 by a distance of order $O(1)$), which is not affordable in the error estimates. Instead, a more refined multiscale approach using the Helffer–Sjöstrand functional calculus is required in order to rewrite the sharp indicator function on the left-hand side of (1.18) in terms of Green functions. The second major difference for bulk eigenvectors is the power counting in the Green function comparison argument, which is in fact easier than at the edge. The main reason for this is that the smallness associated with off-diagonal elements of G is not available in the bulk. Hence we need to assume that four instead of two moments match, and the intricate bookkeeping of the number of off-diagonal resolvent elements is not required. Thus, thanks to the very strong assumption of four-moment matching, the proof of Theorem 1.10 is considerably simpler than that of Theorem 1.6. See Sect. 5 for a more detailed explanation as well the proof.

Conventions. We shall use the letters C and c to denote generic positive constants, which may depend on fixed quantities such as ϑ from (1.9), δ_{\pm} from Assumption (B), and C_0 from Assumption (C). We use C for large constants and c for small constants.

2 Local semicircle law and rigidity of eigenvalues

In this preliminary section we collect the main tools we shall need for our proof. We begin by introducing some notation and by recalling the basic results from [17] on the local semicircle law and the rigidity of eigenvalues.

We define the Green function of H by

$$G_{ij}(z) = \left(\frac{1}{H - z} \right)_{ij}, \tag{2.1}$$

where we the spectral parameter $x = E + i\eta$ satisfies $E \in \mathbb{R}$ and $\eta > 0$. The Stieltjes transform of the empirical eigenvalue distribution of H is defined as

$$m(z) := \frac{1}{N} \sum_i G_{ii}(z) = \frac{1}{N} \text{Tr} \frac{1}{H - z} = \frac{1}{N} \sum_{\alpha} \frac{1}{\lambda_{\alpha} - z}. \tag{2.2}$$

Similarly, we define $m_{sc}(z)$ as the Stieltjes transform of the local semicircle law:

$$m_{sc}(z) := \int \frac{\varrho_{sc}(\lambda) d\lambda}{\lambda - z}.$$

It is well known that $m_{sc}(z)$ can also be characterized as the unique solution of

$$m_{sc}(z) + \frac{1}{z + m_{sc}(z)} = 0 \tag{2.3}$$

with positive imaginary part for all z with $\text{Im } z > 0$. Thus,

$$m_{sc}(z) = \frac{-z + \sqrt{z^2 - 4}}{2}, \tag{2.4}$$

where the square root function is chosen with a branch cut in the segment $[-2, 2]$ so that asymptotically $\sqrt{z^2 - 4} \sim z$ at infinity. This guarantees that the imaginary part of m_{sc} is non-negative for $\eta = \text{Im } z > 0$ and in the limit $\eta \rightarrow 0$.

In order to state the local semicircle law, we introduce the control parameters

$$\Lambda_d := \max_i |G_{ii} - m_{sc}|, \quad \Lambda_o := \max_{i \neq j} |G_{ij}|, \quad \Lambda := |m - m_{sc}|, \quad (2.5)$$

where the subscripts refer to “diagonal” and “off-diagonal” matrix elements. All these quantities depend on the spectral parameter z and on N , but for simplicity we often omit the explicit mention of this dependence from the notation. The following two results were proved in [17].

Theorem 2.1 (Strong local semicircle law) *Let $H = (h_{ij})$ be a Hermitian or symmetric $N \times N$ random matrix satisfying Assumptions (A)–(C). Suppose that the distributions of the matrix elements h_{ij} have a uniformly subexponential decay in the sense of (1.9). Then there exist positive constants $A_0 > 1$, C , c , and $\tau < 1$, such that the following estimates hold for L as in Definition 1.1 and for $N \geq N_0(\vartheta, C_0, \delta_{\pm})$ large enough.*

(i) *The Stieltjes transform of the empirical eigenvalue distribution of H satisfies*

$$\mathbb{P} \left(\bigcup_{z \in \mathbf{S}_L} \left\{ \Lambda(z) \geq \frac{(\log N)^{4L}}{N\eta} \right\} \right) \leq e^{-c(\log N)^{\tau L}}, \quad (2.6)$$

where

$$\mathbf{S}_L := \left\{ z = E + i\eta : |E| \leq 5, \quad N^{-1}(\log N)^{10L} < \eta \leq 10 \right\}. \quad (2.7)$$

(ii) *The individual matrix elements of the Green function satisfy*

$$\mathbb{P} \left(\bigcup_{z \in \mathbf{S}_L} \left\{ \Lambda_d(z) + \Lambda_o(z) \geq (\log N)^{4L} \sqrt{\frac{\text{Im } m_{sc}(z)}{N\eta} + \frac{(\log N)^{4L}}{N\eta}} \right\} \right) \leq e^{-c(\log N)^{\tau L}}. \quad (2.8)$$

(iii) *The norm of H is bounded by $2 + N^{-2/3}(\log N)^{9L}$ in the sense that*

$$\mathbb{P} \left(\|H\| \geq 2 + N^{-2/3}(\log N)^{9L} \right) \leq e^{-c(\log N)^{\tau L}}. \quad (2.9)$$

The local semicircle law implies that the eigenvalues are close to their classical locations with high probability. Recall that $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ are the ordered eigenvalues of H . The classical location γ_α of the α -th eigenvalue was defined in (1.7).

Theorem 2.2 (Rigidity of eigenvalues) *Under the assumptions of Theorem 2.1 there exist positive constants $A_0 > 1$, C , c , and $\tau < 1$, depending only on ϑ in (1.9), δ_{\pm} in*

Assumption **(B)**, and C_0 in Assumption **(C)**, such that such that

$$\mathbb{P}\left\{\exists\alpha : |\lambda_\alpha - \gamma_\alpha| \geq (\log N)^L [\min(\alpha, N - \alpha + 1)]^{-1/3} N^{-2/3}\right\} \leq e^{-c(\log N)^{\tau L}}, \tag{2.10}$$

where L is given in Definition 1.1.

A simple consequence of Theorem 2.1 is that the eigenvectors of H are completely delocalized.

Theorem 2.3 (Complete delocalization of eigenvectors) *Under the assumptions of Theorem 2.1 we have*

$$\mathbb{P}\left\{\exists\alpha, i : |u_\alpha(i)|^2 \geq \frac{\varphi^C}{N}\right\} \leq e^{-c(\log N)^{\tau L}} \tag{2.11}$$

for some positive constants C and c .

Proof Using (2.8) and (2.9) we have, with probability greater than $1 - e^{-c(\log N)^{\tau L}}$,

$$C \geq \text{Im } G_{ii}(\lambda_\alpha + i\eta) = \sum_\beta \frac{\eta |u_\beta(i)|^2}{(\lambda_\alpha - \lambda_\beta)^2 + \eta^2} \geq \frac{|u_\alpha(i)|^2}{\eta}.$$

Choosing $\eta = N^{-1}(\log N)^{20L}$ yields the claim. □

2.1 Stability of the level repulsion condition

In this section we prove that level repulsion, in the sense of (1.4) (respectively (1.5)), holds for the ensemble H^W provided it holds for the ensemble H^V and the first two (respectively four) moments of the entries of H^V and H^W match.

Proposition 2.4 (Stability of level repulsion at the edge) *Let H^V and H^W both satisfy Assumptions **(A)**–**(C)** as well as the uniform subexponential decay condition (1.9). Assume moreover that the first two moments of the entries of H^V and H^W are the same, in the sense of (1.10). If the level repulsion estimate (1.4) holds for H^V then it holds for H^W .*

Proposition 2.5 (Stability of level repulsion in the bulk) *Let H^V and H^W both satisfy Assumptions **(A)**–**(C)** as well as the uniform subexponential decay condition (1.9). Assume moreover that the first four moments of the entries of H^V and H^W are the same, in the sense of (1.13) and (1.14). If the level repulsion estimate (1.5) holds for H^V then it holds for H^W .*

The proofs of Propositions 2.4 and 2.5 are very similar. For definiteness, we give the details for the edge case (Proposition 2.4). The rest of this section is devoted to the proof of Proposition 2.4. The main tool is the following Green function comparison theorem, which was proved in [17], Theorem 6.3.

Lemma 2.6 (Green function comparison theorem at the edge) *Suppose that the assumptions of Theorem 2.1 hold for both ensembles H^V and H^W . Let $F : \mathbb{R} \rightarrow \mathbb{R}$ be a function whose derivatives satisfy*

$$\max_x |F^{(n)}(x)| (1 + |x|)^{-C_1} \leq C_1 \quad \text{for } n = 1, 2, 3, 4, \tag{2.12}$$

with some constant $C_1 > 0$. Then there exists a constant $\varepsilon_0 > 0$, depending only on C_1 , such that for any $\varepsilon < \varepsilon_0$ and for any real numbers E_1 and E_2 satisfying

$$|E_1 + 2| \leq N^{-2/3+\varepsilon}, \quad |E_2 + 2| \leq N^{-2/3+\varepsilon},$$

we have

$$\left| [\mathbb{E}^V - \mathbb{E}^W] F \left(N \int_{E_1}^{E_2} dy \operatorname{Im} m(y + iN^{-2/3-\varepsilon}) \right) \right| \leq CN^{-1/6+C\varepsilon} \tag{2.13}$$

and

$$\left| [\mathbb{E}^V - \mathbb{E}^W] F \left(N\eta \operatorname{Im} m(E_1 + iN^{-2/3+\varepsilon}) \right) \right| \leq CN^{-1/6+C\varepsilon}, \tag{2.14}$$

for some constant C and large enough N , depending only on C_1 , ϑ in (1.9), δ_{\pm} in Assumption (B), and C_0 in Assumption (C).

The basic idea behind the proof of Proposition 2.4 is to first cast the level repulsion estimate into an estimate in terms of Green functions and then use the Green function comparison theorem. Recalling L from Definition 1.1, we set

$$E_L := -2 - 2(\log N)^L N^{-2/3}. \tag{2.15}$$

For any $E \geq E_L$ let

$$\chi_E := \mathbf{1}_{[E_L, E]}$$

be the characteristic function of the interval $[E_L, E]$. For any $\eta > 0$ we define the approximate delta function θ_η on the scale η through

$$\theta_\eta(x) := \frac{\eta}{\pi(x^2 + \eta^2)} = \frac{1}{\pi} \operatorname{Im} \frac{1}{x - i\eta}. \tag{2.16}$$

The following result provides a tool for estimating the number operator using Green functions. It is proved in [17], Lemma 6.1 and Corollary 6.2.

Lemma 2.7 *Suppose that the assumptions of Theorem 2.1 hold, and let A_0 and τ be as in Theorem 2.1. For any $\varepsilon > 0$, set $\ell_1 := N^{-2/3-3\varepsilon}$ and $\eta := N^{-2/3-9\varepsilon}$. Then there exist constants C, c such that, for any E satisfying*

$$|E + 2|N^{2/3} \leq \frac{3}{2}(\log N)^L, \tag{2.17}$$

we have

$$|\text{Tr } \chi_E(H) - \text{Tr } \chi_E * \theta_\eta(H)| \leq C \left(N^{-2\varepsilon} + \mathcal{N}(E - \ell_1, E + \ell_1) \right) \tag{2.18}$$

with high probability.

Moreover, let $\ell := \frac{1}{2}\ell_1 N^{2\varepsilon} = \frac{1}{2}N^{-2/3-\varepsilon}$. Then under the above assumptions the inequalities

$$\text{Tr}(\chi_{E-\ell} * \theta_\eta)(H) - N^{-\varepsilon} \leq \mathcal{N}(-\infty, E) \leq \text{Tr}(\chi_{E+\ell} * \theta_\eta)(H) + N^{-\varepsilon} \tag{2.19}$$

hold with high probability.

After these preparations we may complete the proof of Proposition 2.4.

Proof of Proposition 2.4 Assume that H^V satisfies the level repulsion assumption (1.4) with constant α_0 . We shall show that H^W satisfies also satisfies Definition (1.4) with the same constant α_0 . Fix α satisfying $0 < \alpha \leq \alpha_0$, and let $\delta > 0$ be as chosen so that (1.4) holds for the ensemble H^V .

Abbreviate $E_\pm = E \pm N^{-2/3-\alpha}$ and set $\varepsilon := 2\alpha$. By using (2.19) for $E = E_+$ and $E = E_-$, and subtracting the resulting two inequalities, we get, with high probability,

$$\begin{aligned} \text{Tr}(\mathbf{1}_{[E_+ - \ell, E_+ - \ell]} * \theta_\eta)(H) - 2N^{-\varepsilon} &\leq \mathcal{N}(E_-, E_+) \\ &\leq \text{Tr}(\mathbf{1}_{[E_- - \ell, E_+ + \ell]} * \theta_\eta)(H) + 2N^{-\varepsilon}. \end{aligned} \tag{2.20}$$

Let F be a nonnegative increasing smooth function satisfying $F(x) = 1$ for $x \geq 2$ and $F(x) = 0$ for $x \leq 3/2$. Then, using (2.20) and Lemma 2.6, we have

$$\begin{aligned} \mathbb{E}^W F(\mathcal{N}(E_-, E_+)) &\leq \mathbb{E}^W F(\text{Tr}(\mathbf{1}_{[E_- - \ell, E_+ + \ell]} * \theta_\eta)(H) + 2N^{-\varepsilon}) \\ &\leq \mathbb{E}^V F(\text{Tr}(\mathbf{1}_{[E_- - \ell, E_+ + \ell]} * \theta_\eta)(H)) + CN^{-\varepsilon} + CN^{-1/6+C\varepsilon} \\ &\leq \mathbb{E}^V F(\mathcal{N}(E_- - 2\ell, E_+ + 2\ell) + N^{-\varepsilon}) + CN^{-\varepsilon} + CN^{-1/6+C\varepsilon} \\ &\leq N^{-\alpha-\delta} + CN^{-\varepsilon} + CN^{-1/6+C\varepsilon}. \end{aligned}$$

Since $\varepsilon = 2\alpha$, we get that (1.4) holds for the ensemble H^W with exponent $\delta' = \min\{\delta, \alpha\} > 0$. □

3 Proof of Theorem 1.6

To simplify presentation, in this section we prove Theorem 1.6 in the special case $\theta = \theta(N\bar{u}_\alpha(i)u_\alpha(j))$, where $\alpha \leq \varphi^\rho$. The proof of the general case is analogous; see Sect. 4 for more details.

In a first step we convert the eigenvector problem into a problem involving the Green function G_{ij} . To that end, we define

$$\tilde{G}_{ij}(z) := \frac{1}{2i}(G_{ij}(z) - G_{ij}(\bar{z})) = \eta \sum_k G_{ik}(z)\overline{G_{jk}(z)}, \tag{3.1}$$

where the second equality follows easily by spectral decomposition, $G_{ij}(z) = \sum_\beta \frac{\bar{u}_\beta(i)u_\beta(j)}{\lambda_\beta - z}$. Note that

$$\tilde{G}_{ij}(E + i\eta) = \sum_\beta \frac{\eta}{(E - \lambda_\beta)^2 + \eta^2} \bar{u}_\beta(i)u_\beta(j)$$

as well as $\text{Im } G_{ii}(z) = \tilde{G}_{ii}(z)$. It is a triviality that all of the results from Sect. 2 hold with z replaced with \bar{z} .

The following lemma expresses the eigenvector components as an integral of the Green function over an appropriate random interval.

Lemma 3.1 *Under the assumptions of Theorem 1.6, for any $\varepsilon > 0$ there exist constants C_1, C_2 such that for $\eta = N^{-2/3-\varepsilon}$ we have*

$$\lim_{N \rightarrow \infty} \max_{\alpha \leq \varphi^\rho} \max_{i,j} \left\{ \mathbb{E}^{\mathbf{u}} \theta(N\bar{u}_\alpha(i)u_\alpha(j)) - \mathbb{E}^{\mathbf{u}} \theta \left[\frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) \mathbf{1}(\lambda_{\alpha-1} \leq E^- \leq \lambda_\alpha) dE \right] \right\} = 0, \tag{3.2}$$

where

$$E^\pm := E \pm \varphi^{C_1} \eta, \quad I := \left[-2 - N^{-2/3} \varphi^{C_2}, -2 + N^{-2/3} \varphi^{C_2} \right] \tag{3.3}$$

and we introduce the convention $\lambda_0 := -\infty$. Here \mathbf{u} stands for either \mathbf{v} or \mathbf{w} .

Proof We shall fix i, j and $\alpha \leq \varphi^\rho$; it is easy to check that all constants in the following are uniform in i, j , and $\alpha \leq \varphi^\rho$. We write

$$\bar{u}_\alpha(i)u_\alpha(j) = \frac{\eta}{\pi} \int \frac{\bar{u}_\alpha(i)u_\alpha(j)}{(E - \lambda_\alpha)^2 + \eta^2} dE. \tag{3.4}$$

Using Theorem 2.3 it is easy to prove that for C_1 large enough we have

$$\bar{u}_\alpha(i)u_\alpha(j) = \frac{\eta}{\pi} \int_a^b \frac{\bar{u}_\alpha(i)u_\alpha(j)}{(E - \lambda_\alpha)^2 + \eta^2} dE + O\left(\frac{1}{N\varphi^{C_1/2}}\right) \tag{3.5}$$

holds with high probability for some $c > 0$, as long as

$$a \leq \lambda_\alpha^-, \quad b \geq \lambda_\alpha^+, \tag{3.6}$$

where we use the notation (3.3), i.e. $\lambda_\alpha^\pm := \lambda_\alpha \pm \varphi^{C_1}\eta$. We now choose

$$a := \min\{\lambda_\alpha^-, \lambda_{\alpha-1}^+\}, \quad b := \lambda_\alpha^+.$$

By the assumption on θ and using Theorem 2.3, we therefore find

$$\mathbb{E}^{\mathbf{u}}\theta(N\bar{u}_\alpha(i)u_\alpha(j)) = \mathbb{E}^{\mathbf{u}}\theta\left(\frac{N\eta}{\pi} \int_a^b \frac{\bar{u}_\alpha(i)u_\alpha(j)}{(E - \lambda_\alpha)^2 + \eta^2} dE\right) + o(1). \tag{3.7}$$

Now we split

$$\int_a^b dE = \int_{\lambda_{\alpha-1}^+}^{\lambda_\alpha^+} dE + \mathbf{1}(\lambda_{\alpha-1}^+ > \lambda_\alpha^-) \int_{\lambda_\alpha^-}^{\lambda_{\alpha-1}^+} dE$$

to get

$$\begin{aligned} \mathbb{E}^{\mathbf{u}}\theta(N\bar{u}_\alpha(i)u_\alpha(j)) &= \mathbb{E}^{\mathbf{u}}\theta\left(\frac{N\eta}{\pi} \int_{\lambda_{\alpha-1}^+}^{\lambda_\alpha^+} \frac{\bar{u}_\alpha(i)u_\alpha(j)}{(E - \lambda_\alpha)^2 + \eta^2} dE\right) \\ &\quad + O(\varphi^{C_0}\mathbb{E}^{\mathbf{u}}\mathbf{1}(\lambda_{\alpha-1}^+ > \lambda_\alpha^-)) + o(1) \end{aligned} \tag{3.8}$$

for some constant C_0 , where we used Theorem 2.3 and the assumption on θ . Now the level repulsion estimate (1.6) implies that the second term of (3.8) is $o(1)$. We now observe that, by (2.10), we have $\lambda_\alpha^+ \leq -2 + N^{-2/3}\varphi^{C_2}$ and $\lambda_{\alpha-1}^+ \geq -2 - N^{-2/3}\varphi^{C_2}$ with high probability. It therefore easy to see that

$$\begin{aligned} \mathbb{E}^{\mathbf{u}}\theta(N\bar{u}_\alpha(i)u_\alpha(j)) &= \mathbb{E}^{\mathbf{u}}\theta\left(\frac{N\eta}{\pi} \int_I \frac{\bar{u}_\alpha(i)u_\alpha(j)}{(E - \lambda_\alpha)^2 + \eta^2} \mathbf{1}(\lambda_{\alpha-1} \leq E^- \leq \lambda_\alpha) dE\right) \\ &\quad + o(1). \end{aligned} \tag{3.9}$$

Next, we replace the integrand in (3.9) by $\tilde{G}_{ij}(E + i\eta)$. By definition, we have

$$\frac{1}{\eta} \tilde{G}_{ij}(E + i\eta) = \sum_{\beta \neq \alpha} \frac{\bar{u}_\beta(i)u_\beta(j)}{(E - \lambda_\beta)^2 + \eta^2} + \frac{\bar{u}_\alpha(i)u_\alpha(j)}{(E - \lambda_\alpha)^2 + \eta^2}. \tag{3.10}$$

In order to be able to apply the mean value theorem to θ with the decomposition (3.10), we need an upper bound on

$$\sum_{\beta} \frac{N\eta}{\pi} \int_I \frac{|\bar{u}_\beta(i)u_\beta(j)|}{(E - \lambda_\beta)^2 + \eta^2} dE \leq \varphi^{C_0+C_3} + \varphi^{C_0} \sum_{\beta \geq \varphi^{C_3}} \int_I \frac{\eta}{(E - \lambda_\beta)^2 + \eta^2} dE, \tag{3.11}$$

where the inequality holds with high probability for any C_3 ; here we used Theorem 2.3. Using $\gamma_\beta \geq -2 + c(\beta/N)^{2/3}$ as well as (2.10), we find with high probability for large enough C_3

$$\varphi^{C_0} \sum_{\beta \geq \varphi^{C_3}} \int_I \frac{\eta}{(E - \lambda_\beta)^2 + \eta^2} dE \leq \varphi^{C_0+C_2} N^{-2/3} \sum_{\beta \geq \varphi^{C_3}} \frac{\eta}{(\beta/N)^{4/3}} \leq N^{-\varepsilon/2}. \tag{3.12}$$

Thus the left-hand side of (3.11) is bounded by $\varphi^{C_0+C_3+1}$.

Let us abbreviate $\chi(E) := \mathbf{1}(\lambda_{\alpha-1} \leq E^- \leq \lambda_\alpha)$. Now, recalling the assumption on θ , we may apply the mean value theorem as well as Theorem 2.3 to get

$$\begin{aligned} & \left| \mathbb{E}^{\mathbf{u}} \theta \left(\frac{N\eta}{\pi} \int_I \frac{\bar{u}_\alpha(i)u_\alpha(j)}{(E - \lambda_\alpha)^2 + \eta^2} \chi(E) dE \right) - \mathbb{E}^{\mathbf{u}} \theta \left(\frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) \chi(E) dE \right) \right| \\ & \leq \varphi^{\tilde{C}} \mathbb{E}^{\mathbf{u}} \sum_{\beta \neq \alpha} \frac{N\eta}{\pi} \int_I \frac{|\bar{u}_\beta(i)u_\beta(j)|}{(E - \lambda_\beta)^2 + \eta^2} \chi(E) dE \end{aligned} \tag{3.13}$$

for some constant $\tilde{C} \leq C(C_0 + C_3 + 1)$ independent of C_1 . We now estimate the right-hand side of (3.13). Exactly as in (3.12), one finds that there exists C_4 such that the contribution of $\beta \geq \varphi^{C_4}$ to the right-hand side of (3.13) vanishes in the limit $N \rightarrow \infty$. Next, we deal with the eigenvalues $\beta < \alpha$ (in the case $\alpha > 1$). Using Theorem 2.3 we get

$$\sum_{\beta < \alpha} \frac{N\eta}{\pi} \mathbb{E}^{\mathbf{u}} \int_I \frac{|\bar{u}_\beta(i)u_\beta(j)|}{(E - \lambda_\beta)^2 + \eta^2} \chi(E) dE \leq \varphi^C \mathbb{E}^{\mathbf{u}} \int_{\lambda_{\alpha-1}^+}^\infty \frac{\eta}{(E - \lambda_{\alpha-1})^2 + \eta^2} dE \leq \varphi^{-\tilde{C}-c},$$

where $c > 0$ for C_1 large enough.

What remains is the estimate of the terms $\alpha < \beta \leq \varphi^{C_4}$ in (3.13). For a given constant $C_5 > 0$ we partition $I = I_1 \cup I_2$ with $I_1 \cap I_2 = \emptyset$ and

$$I_1 := \left\{ E \in I : \exists \beta, \alpha < \beta \leq \varphi^{C_4}, |E - \lambda_\beta| \leq \eta \varphi^{C_5} \right\}. \tag{3.14}$$

It is easy to see that, for large enough C_5 , we have

$$\sum_{\beta: \alpha < \beta \leq \varphi^{C_4}} \frac{N\eta}{\pi} \mathbb{E}^{\mathbf{u}} \int_{I_2} \frac{|\bar{u}_\beta(i)u_\beta(j)|}{(E - \lambda_\beta)^2 + \eta^2} \chi(E) dE \leq \varphi^{-\tilde{c}-c}$$

where $c > 0$. Let us therefore consider the integral over I_1 . One readily finds, for $\lambda_\alpha \leq \lambda_{\alpha+1} \leq \lambda_\beta$, that

$$\frac{1}{(E - \lambda_\beta)^2 + \eta^2} \mathbf{1}(E^- \leq \lambda_\alpha) \leq \frac{\varphi^{2C_1}}{(\lambda_\beta - \lambda_\alpha)^2 + \eta^2} \leq \frac{\varphi^{2C_1}}{(\lambda_{\alpha+1} - \lambda_\alpha)^2 + \eta^2}. \tag{3.15}$$

From Theorem 2.3 we therefore find that there exists a constant C_6 , depending on C_1 , such that

$$\sum_{\beta: \alpha < \beta \leq \varphi^{C_4}} \frac{N\eta}{\pi} \mathbb{E}^{\mathbf{u}} \int_{I_1} \frac{|\bar{u}_\beta(i)u_\beta(j)|}{(E - \lambda_\beta)^2 + \eta^2} \chi(E) dE \leq \varphi^{C_6} \mathbb{E}^{\mathbf{u}} \frac{\eta^2}{(\lambda_{\alpha+1} - \lambda_\alpha)^2 + \eta^2}. \tag{3.16}$$

The right-hand side of (3.16) is bounded by $\mathbb{E}^{\mathbf{u}} \mathbf{1}(|\lambda_{\alpha+1} - \lambda_\alpha| \leq N^{-1/3} \eta^{1/2}) + O(N^{-\varepsilon})$. Using (1.6) we now obtain

$$\sum_{\beta: \alpha < \beta \leq \varphi^{C_4}} \frac{N\eta}{\pi} \mathbb{E}^{\mathbf{u}} \int_{I_1} \frac{|\bar{u}_\beta(i)u_\beta(j)|}{(E - \lambda_\beta)^2 + \eta^2} \chi(E) dE \leq \varphi^{-\tilde{c}-c} \tag{3.17}$$

where $c > 0$. This concludes the proof. □

In a second step we convert the cutoff function in lemma 3.1 into a function of \tilde{G}_{ij} .

Lemma 3.2 *Recall the definition (2.16) of the approximate delta function θ_η on the scale η . Let $\alpha \leq \varphi^\rho$ and $q \equiv q_\alpha : \mathbb{R} \rightarrow \mathbb{R}_+$ be a smooth cutoff function concentrated around $\alpha - 1$, satisfying*

$$q(x) = 1 \text{ if } |x - \alpha + 1| \leq 1/3, \quad q(x) = 0 \text{ if } |x - \alpha + 1| \geq 2/3.$$

Let

$$\chi := \mathbf{1}_{[E_L, E^-]}, \quad E_L := -2 - 2N^{-2/3}(\log N)^L \tag{3.18}$$

where

$$\eta := N^{-2/3-\varepsilon}, \quad \tilde{\eta} := N^{-2/3-6\varepsilon} \tag{3.19}$$

for $\varepsilon > 0$. Then for ε small enough we have (recall the Definition (3.3))

$$\lim_{N \rightarrow \infty} \max_{\alpha \leq \varphi^\rho} \max_{i,j} \left\{ \mathbb{E}^{\mathbf{u}} \theta(N\bar{u}_\alpha(i)u_\alpha(j)) - \mathbb{E}^{\mathbf{u}} \theta \left[\frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) q[\text{Tr}(\chi * \theta_{\tilde{\eta}})(H)] dE \right] \right\} = 0. \tag{3.20}$$

Here \mathbf{u} stands for either \mathbf{v} or \mathbf{w} .

Proof Note first that

$$\begin{aligned} & \frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) \mathbf{1}(\lambda_{\alpha-1} \leq E^- \leq \lambda_\alpha) dE \\ &= \frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) \mathbf{1}(\mathcal{N}(-\infty, E^-) = \alpha - 1) dE \\ &= \frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) q[\text{Tr} \chi(H)] dE \end{aligned}$$

with high probability.

Next, recall that (2.18) asserts that for $\ell = N^{-2/3-2\varepsilon}$ we have

$$|\text{Tr} \chi(H) - \text{Tr}(\chi * \theta_{\tilde{\eta}})(H)| \leq C(N^{-\varepsilon} + \mathcal{N}(E^- - \ell, E^- + \ell)) \tag{3.21}$$

with high probability for sufficiently large N . We therefore find that

$$\begin{aligned} & \left| \frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) \mathbf{1}(\lambda_{\alpha-1} \leq E^- \leq \lambda_\alpha) dE - \frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) q[\text{Tr}(\chi * \theta_{\tilde{\eta}})(H)] dE \right| \\ & \leq CN \sum_{\beta=1}^N \int_I |\tilde{G}_{ij}(E + i\eta)| \mathbf{1}(|E^- - \lambda_\beta| \leq \ell) dE + \varphi^C N^{-\varepsilon} \\ & \leq CN \sum_{\beta=1}^{\varphi^C} \int_I |\tilde{G}_{ij}(E + i\eta)| \mathbf{1}(|E^- - \lambda_\beta| \leq \ell) dE + \varphi^C N^{-\varepsilon} \\ & \leq C\varphi^C N \ell \sup_{E \in I} |\tilde{G}_{ij}(E + i\eta)| + \varphi^C N^{-\varepsilon} \end{aligned}$$

holds with high probability, where in the first inequality we estimated the integral $\int_I |\tilde{G}_{ij}(E + i\eta)| dE$ exactly as (3.11), and in the second inequality we used (2.10). Using the definition of I and (2.8) we get

$$\sup_{E \in I} |\tilde{G}_{ij}(E + i\eta)| \leq \varphi^C \left(N^{-1/3} + N^{-2/3} \eta^{-1/2} + N^{-1} \eta^{-1} \right) \leq N^{-1/3+\varepsilon}.$$

Together with (3.2), the claim follows. □

In a third and final step, we use the Green function comparison method to show the following statement.

Lemma 3.3 *Under the assumptions of Lemma 3.2, we have*

$$\lim_{N \rightarrow \infty} \max_{i,j} (\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}) \theta \left[\frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) q[\text{Tr}(\chi * \theta_{\tilde{\eta}})(H)] dE \right] = 0.$$

The rest of this section is devoted to the proof of Lemma 3.3.

3.1 Green function comparison: proof of Lemma 3.3

The claimed uniformity in i and j is easy to check in our proof, and we shall not mention it anymore. Throughout the following we rename $i = \alpha$ and $j = \beta$ in order to use i and j as summation indices. We now fix α and β for the whole proof. (Note that α and β need not be different.)

We use the identity (see (3.1))

$$\tilde{G}_{ij}(z) = (\text{Im } z) \sum_k X_{ij,k}(z), \quad X_{ij,k}(z) := G_{ik}(z) \overline{G_{jk}(z)}. \tag{3.22}$$

We begin by dropping the diagonal terms in (3.22).

Lemma 3.4 *For small enough $\varepsilon > 0$ we have*

$$\mathbb{E}^{\mathbf{u}} \theta \left[\frac{N}{\pi} \int_I \tilde{G}_{\alpha\beta}(E + i\eta) q[\text{Tr}(\chi * \theta_{\tilde{\eta}})(H)] dE \right] - \mathbb{E}^{\mathbf{u}} \theta \left[\int_I x(E) q(y(E)) dE \right] = o(1), \tag{3.23}$$

where \mathbf{u} stands for either \mathbf{v} or \mathbf{w} , and

$$x(E) := \frac{N\eta}{\pi} \sum_{k \neq \alpha, \beta} X_{\alpha\beta,k}(E + i\eta), \quad y(E) := \tilde{\eta} \int_{E_L}^{E^-} \sum_{i \neq k} X_{ii,k}(\tilde{E} + i\tilde{\eta}) d\tilde{E}. \tag{3.24}$$

Proof We estimate

$$\left| \frac{N}{\pi} \tilde{G}_{\alpha\beta}(E + i\eta) - x(E) \right| \leq \varphi^C N \eta \leq \varphi^C N^{1/3-\varepsilon}$$

with high probability and, recalling that q' is bounded,

$$|q[\text{Tr}(\chi * \theta_{\tilde{\eta}})(H)] - y(E)| \leq \varphi^C \tilde{\eta} N N^{-2/3} \leq \varphi^C N^{-1/3-6\varepsilon}$$

with high probability. Therefore the difference of the arguments of θ in (3.23) is bounded by $\varphi^C N^{-1/3-\varepsilon}$ with high probability. (Recall that $|I| \leq \varphi^C N^{-2/3}$.) Moreover, since q is bounded, it is easy to see that both arguments of θ in (3.23) are bounded with high probability by

$$\varphi^C N \eta N^{-2/3} \left(1 + N \sup_{E \in I} \Lambda_0^2(E + i\eta) \right) \leq \varphi^C N^\varepsilon,$$

where we used Theorem 2.1. The claim now follows from the mean value theorem and the assumption on θ . □

For the following we work on the product probability space of the ensembles H^V and H^W . To distinguish them we denote the elements of H^V by $N^{1/2}v_{ij}$ and the elements of H^W by $N^{-1/2}w_{ij}$. We fix a bijective ordering map Φ on the index set of the independent matrix elements,

$$\Phi : \{(i, j) : 1 \leq i \leq j \leq N\} \rightarrow \{1, \dots, \gamma_{\max}\}, \quad \gamma_{\max} := \frac{N(N+1)}{2},$$

and denote by H_γ the generalized Wigner matrix whose matrix elements h_{ij} follow the v -distribution if $\Phi(i, j) \leq \gamma$ and the w -distribution otherwise. In particular, $H_0 = H^V$ and $H_{\gamma_{\max}} = H^W$. Hence

$$\begin{aligned} & [\mathbb{E}^V - \mathbb{E}^W] \theta \left[\int_I x(E) q(y(E)) dE \right] \\ &= \sum_{\gamma=1}^{\gamma_{\max}} \left[\mathbb{E}^{(H_{\gamma-1})} - \mathbb{E}^{(H_\gamma)} \right] \theta \left[\int_I x(E) q(y(E)) dE \right] \end{aligned} \tag{3.25}$$

(in self-explanatory notation).

Let us now fix a γ and let (a, b) be determined by $\Phi(a, b) = \gamma$. Throughout the following we consider α, β, a, b to be arbitrary but fixed and often omit dependence on them from the notation. Our strategy is to compare $H_{\gamma-1}$ with H_γ for each γ . In the end we shall sum up the differences in the telescopic sum (3.25).

Note that $H_{\gamma-1}$ and H_γ differ only in the matrix elements indexed by (a, b) and (b, a) . Let $E^{(ij)}$ denote the matrix whose matrix elements are zero everywhere except at position (i, j) where it is 1; in other words, $E_{k\ell}^{(ij)} = \delta_{ik}\delta_{j\ell}$. Thus we have

$$\begin{aligned} H_{\gamma-1} &= Q + \frac{1}{\sqrt{N}}V, & V &:= v_{ab}E^{(ab)} + v_{ba}E^{(ba)}, \\ H_\gamma &= Q + \frac{1}{\sqrt{N}}W, & W &:= w_{ab}E^{(ab)} + w_{ba}E^{(ba)}. \end{aligned} \tag{3.26}$$

Here Q is the matrix obtained from H_γ (or, equivalently, from $H_{\gamma-1}$) by setting the matrix elements indexed by (a, b) and (b, a) to zero. Next, we define the Green functions

$$R := \frac{1}{Q - z}, \quad S := \frac{1}{H_{\gamma-1} - z}. \tag{3.27}$$

We shall show that the difference between the expectation $\mathbb{E}^{(H_{\gamma-1})}$ and $\mathbb{E}^{(Q)}$ depends only on the second moments of v_{ab} , up to an error term that is affordable even after summation over γ . Together with same argument applied to $\mathbb{E}^{(H_\gamma)}$, and the fact that the second moments of v_{ab} and w_{ab} are identical, this will prove Lemma 3.3.

For the estimates we need the following basic result, proved in [17] (Eq. 6.32).

Lemma 3.5 *For any $\eta' := N^{-2/3-\delta}$ we have with high probability*

$$\sup_{E \leq N^{-2/3+\varepsilon}} \max_{i,j} |R_{ij}(E + i\eta') - \delta_{ij}m_{sc}(E + i\eta')| \leq \Lambda_\delta := N^{-1/3+2\delta}.$$

The same estimates hold for S instead of R .

Our comparison is based on the resolvent expansion

$$S = R - N^{-1/2}RV R + N^{-1}(RV)^2R - N^{-3/2}(RV)^3R + N^{-2}(RV)^4S. \tag{3.28}$$

Using Lemma 3.5 we easily get with high probability, for $i \neq j$,

$$|S_{ij} - R_{ij}| \leq \varphi^C N^{-1/2} \Lambda_\varepsilon^{2-r} \quad \text{where } r := \mathbf{1}(i \in \{a, b\}) + \mathbf{1}(j \in \{a, b\}). \tag{3.29}$$

Defining

$$\Delta X_{ij,k} := S_{ik}\overline{S_{jk}} - R_{ik}\overline{R_{jk}}, \tag{3.30}$$

we therefore have the trivial bound with high probability

$$|\Delta X_{ij,k}| \leq \varphi^C N^{-1/2} \Lambda_\varepsilon^{3-s} \quad (k \neq i, j), \tag{3.31}$$

where we abbreviated

$$s := \max \left\{ \mathbf{1}(i \in \{a, b\}) + \mathbf{1}(k \in \{a, b\}), \mathbf{1}(j \in \{a, b\}) + \mathbf{1}(k \in \{a, b\}) \right\} \\ = \mathbf{1}(\{i, j\} \cap \{a, b\} \neq \emptyset) + \mathbf{1}(k \in \{a, b\}). \tag{3.32}$$

The variable s counts the maximum number of diagonal resolvent matrix elements in $\Delta X_{ij,k}$. The bookkeeping of s will play a crucial role in our proof, since the smallness associated with off-diagonal elements (see Lemma 3.5) is needed to control the resolvent expansion (3.28) under the two-moment matching assumption.

From now on it is convenient to modify slightly our notation, and to write $\mathbb{E}F(H_\gamma)$ instead of $\mathbb{E}^{(H_\gamma)}F(H)$. We also use \mathbb{E}_{ab} to denote partial expectation obtained by integrating out the variables v_{ab} and w_{ab} .

By applying (3.28) to (3.30) and taking the partial expectation \mathbb{E}_{ab} , one finds, as above, that there exists a random variable A_1 , which depends on the randomness only through Q and the first two moments of v_{ab} , such that for $k \neq i, j$ and s as in (3.32) we have with high probability

$$|\mathbb{E}_{ab} \Delta X_{ij,k} - A_1| \leq \varphi^C N^{-3/2} \Lambda_\varepsilon^{3-s}. \tag{3.33}$$

Using this bound we may estimate

$$\Delta x(E) := x^S(E) - x^R(E), \quad \Delta y(E) := y^S(E) - y^R(E), \tag{3.34}$$

with the convention that a superscript S denotes a quantity defined in terms of the matrix $H_{\gamma-1}$, and a superscript R a quantity defined in terms of the matrix Q .

Lemma 3.6 *For fixed α, β, a, b there exists a random variable A , which depends on the randomness only through Q and the first two moments of v_{ab} , such that*

$$\mathbb{E}\theta \left[\int_I x^S(E)q(y^S(E))dE \right] - \mathbb{E}\theta \left[\int_I x^R(E)q(y^R(E))dE \right] \\ = A + o(N^{-2+t} + N^{-2+1(a=b)}), \tag{3.35}$$

where $t := |\{a, b\} \cap \{\alpha, \beta\}|$.

Before proving Lemma 3.6, we show how it implies Lemma 3.3.

Proof of Lemma 3.3 It suffices to prove that each summand in (3.25) is bounded by $o(N^{-2+t} + N^{-2+1(a=b)})$. This follows immediately by applying Lemma 3.6 to $S = (H_{\gamma-1} - z)^{-1}$ and $S' := (H_\gamma - z)^{-1}$ and subtracting the statements; note that the random variables A in the statement of Lemma 3.6 are by definition the same for S and S' . □

Proof of Lemma 3.6 Throughout the proof of Lemma 3.6 we shall abbreviate $H \equiv H_{\gamma-1} = (h_{ij})$, as well as $S \equiv S(z) = (H - z)^{-1}$.

Since $E \in I$ (recall (3.3)) we get from Theorem 2.1 that with high probability

$$|x(E)| \leq \varphi^C N^2 \eta \Lambda_\varepsilon^2 \leq \frac{N^{C\varepsilon}}{\eta}, \tag{3.36}$$

which implies

$$\int_I |x(E)| dE \leq N^{C\varepsilon}. \tag{3.37}$$

Here we adopt the convention that if x or y appears without a superscript, the claim holds for both superscripts R and S . Similarly, we find with high probability

$$|y(E)| \leq \tilde{\eta} \varphi^C N^{-2/3} N^2 \Lambda_{6\varepsilon}^2 \leq N^{C\varepsilon}. \tag{3.38}$$

Next, in the definition of $x(E)$ and $y(E)$ we condition on the variable s defined in (3.32) by introducing, for $s = 0, 1, 2$,

$$x_s(E) := \frac{N\eta}{\pi} \sum_{k \neq \alpha, \beta} X_{\alpha\beta,k}(E + i\eta) \mathbf{1}(s = \mathbf{1}(\{\alpha, \beta\} \cap \{a, b\} \neq \emptyset) + \mathbf{1}(k \in \{a, b\})),$$

$$y_s(E) := \tilde{\eta} \int_{E_L}^{E^-} \sum_{i \neq k} X_{ii,k}(\tilde{E} + i\tilde{\eta}) d\tilde{E} \mathbf{1}(s = \mathbf{1}(i \in \{a, b\}) + \mathbf{1}(k \in \{a, b\})).$$

As above, s is a bookkeeping index that bounds the number of diagonal resolvent matrix elements appearing in the resolvent expansion.

We abbreviate $\Delta x_s(E) := x_s^S(E) - x_s^R(E)$ and $\Delta y_s(E) = y_s^S(E) - y_s^R(E)$. Recalling the definition $t = |\{a, b\} \cap \{\alpha, \beta\}|$, we find with high probability

$$|\Delta x_s(E)| \leq \varphi^C N \eta N^{-1/2} \Lambda_\varepsilon^{3-s} N^{\mathbf{1}(s=\mathbf{1}(t>0))} \leq \frac{\eta^{s-2}}{N^{3/2-t-C\varepsilon}}, \tag{3.39}$$

where we used Theorem 2.1 and the elementary inequality $s + \mathbf{1}(s = \mathbf{1}(t > 0)) \leq t + 1$ which holds if $x_s(E) \neq 0$. Thus we get with high probability

$$\int_I |\Delta x_s(E)| dE \leq \frac{\eta^{s-1}}{N^{3/2-t-C\varepsilon}} = N^{-5/6} N^{-2s/3+t+C\varepsilon}. \tag{3.40}$$

Now we may argue similarly to (3.33). We find that, for any E -dependent random variable $f \equiv f(E)$ independent of h_{ab} , there exists a random variable A_2 , which depends on the randomness only through Q, f , and the first two moments of h_{ab} , such

that with high probability

$$\left| \int_I (\mathbb{E}_{ab} \Delta x_s(E)) f(E) dE - A_2 \right| \mathbf{1}(\Omega) \leq \|f\mathbf{1}(\Omega)\|_\infty N^{-11/6} N^{-2s/3+t+C\varepsilon}, \tag{3.41}$$

where Ω is any event. Note that, as in (3.33), we find that (3.41) is suppressed by a factor N^{-1} compared to (3.31). This may be easily understood, as the leading order error term in the resolvent expansion of (3.31) is of order 1 in H , whereas the leading order error term in (3.41) is of order 3 in H . These error terms have the same number of off-diagonal elements (estimated using Lemma 3.5), and the same entropy factor of the summation indices.

We may derive similar bounds for $y_s(E)$. As in (3.31), we have with high probability

$$|\Delta y_s(E)| \leq \varphi^C \tilde{\eta} N^{-2/3} N^{2-s} N^{-1/2} \Lambda_{6\varepsilon}^{3-s} \leq N^{-5/6} N^{-2s/3+C\varepsilon}. \tag{3.42}$$

Furthermore, we find that there exists an E -dependent random variable $A_3(E)$, which depends on the randomness only through Q and the first two moments of h_{ab} , such that with high probability

$$\left| \mathbb{E}_{ab}^{(H_{y^{-1}})} \Delta y_s(E) - A_3(E) \right| \leq N^{-11/6} N^{-2s/3+C\varepsilon}. \tag{3.43}$$

After these preparations, we may now estimate the error resulting from setting h_{ab} to zero in the expression $\mathbb{E}\theta \left[\int_I x(E)q(y(E))dE \right]$. Recalling the conditioning over $s = 0, 1, 2$, we find

$$\begin{aligned} & \theta \left[\int_I x^S q(y^S) dE \right] \\ &= \theta \left[\int_I \left(x^R + \Delta x_0 + \Delta x_1 + \Delta x_2 \right) q \left(y^R + \Delta y_0 + \Delta y_1 + \Delta y_2 \right) dE \right]; \end{aligned}$$

here and in the following we omit the argument E unless it is needed. Using (3.42) we have with high probability

$$\begin{aligned} & \theta \left[\int_I \left(x^R + \Delta x_0 + \Delta x_1 + \Delta x_2 \right) q \left(y^R + \Delta y_0 + \Delta y_1 + \Delta y_2 \right) dE \right] \\ &= \theta \left[\int_I \left(x^R + \Delta x_0 + \Delta x_1 + \Delta x_2 \right) \left(q(y^R) + q'(y^R) (\Delta y_0 + \Delta y_1) \right. \right. \\ & \quad \left. \left. + q''(y^R) (\Delta y_0)^2 \right) dE \right] + o(N^{-2}). \end{aligned}$$

The use of the mean value theorem for ε small enough is easy to justify using the assumption on θ and the bounds (3.37) and (3.38). In the following we shall no longer mention such estimates of the argument of derivatives of θ , which can always be easily checked in a similar fashion.

Recall that an error of order $o(N^{-2+t})$ is affordable in the error estimate. Thus, using the basic power counting given by (3.37), (3.38), (3.40), and (3.42), we find with high probability

$$\begin{aligned} & \theta \left[\int_I x^S q(y^S) dE \right] - \theta \left[\int_I x^R q(y^R) dE \right] = \theta' \left[\int_I x^R q(y^R) dE \right] \\ & \times \left[\int_I \left((\Delta x_0 + \Delta x_1) q(y^R) + x^R q'(y^R) (\Delta y_0 + \Delta y_1) \right. \right. \\ & \left. \left. + \Delta x_0 q'(y^R) \Delta y_0 + x^R q''(y^R) (\Delta y_0)^2 \right) dE \right] \\ & + \frac{1}{2} \theta'' \left[\int_I x^R q(y^R) dE \right] \left[\int_I \left(\Delta x_0 q(y^R) + x^R q'(y^R) \Delta y_0 \right) dE \right]^2 + o(N^{-2+t}). \end{aligned} \tag{3.44}$$

We now start dealing with the individual terms on the right-hand side of (3.44).

First, we consider the terms containing Δx_1 and Δy_1 . Applying (3.41) and (3.43) we find that there exists a random variable A_4 , which depends on the randomness only through Q and the first two moments of h_{ab} , such that

$$\left| \mathbb{E}_{ab} \int_I \left(\Delta x_1 q(y^R) + x^R q'(y^R) \Delta y_1 \right) dE - A_4 \right| = o(N^{-2+t}) \tag{3.45}$$

with high probability. Inserting this into (3.44), we find with high probability

$$\begin{aligned} & \mathbb{E}_{ab} \theta \left[\int_I x^S q(y^S) dE \right] - \mathbb{E}_{ab} \theta \left[\int_I x^R q(y^R) dE \right] \\ & = \mathbb{E}_{ab} \theta' \left[\int_I x^R q(y^R) dE \right] \times \left[\int_I \left(\Delta x_0 q(y^R) \right. \right. \\ & \left. \left. + x^R q'(y^R) \Delta y_0 + \Delta x_0 q'(y^R) \Delta y_0 + x^R q''(y^R) (\Delta y_0)^2 \right) dE \right] \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{2} \mathbb{E}_{ab} \theta'' \left[\int_I x^R q(y^R) dE \right] \left[\int_I (\Delta x_0 q(y^R) + x^R q'(y^R) \Delta y_0) dE \right]^2 \\
 & + A_4 + o(N^{-2+t}).
 \end{aligned} \tag{3.46}$$

Thus we only need to focus on the error terms Δx_0 and Δy_0 . Note that we have

$$\Delta x_0(E) = \mathbf{1}(t = 0) \frac{N\eta}{\pi} \sum_{k \neq \alpha, \beta, a, b} \Delta X_{\alpha\beta, k}(E + i\eta) \tag{3.47}$$

$$\Delta y_0(E) = \tilde{\eta} \int_{E_L}^{E^-} d\tilde{E} \sum_{i \neq k} \mathbf{1}(i, k \notin \{a, b\}) \Delta X_{ii, k}(\tilde{E} + i\tilde{\eta}). \tag{3.48}$$

Recall that the (i, j) -component of the resolvent expansion (3.28) reads

$$S_{ij} = \left(R - N^{-1/2} R V R + N^{-1} (R V)^2 R - N^{-3/2} (R V)^3 R + N^{-2} (R V)^4 S \right)_{ij}. \tag{3.49}$$

Now we assume that $i \neq j$ and $\{i, j\} \cap \{a, b\} = \emptyset$. It is easy to see that this assumption holds for any matrix element in the formulas (3.47) and (3.48). Then we can use Lemma 3.5 to estimate the m -th term as follows:

$$\left| N^{-m/2} [(R V)^m R]_{ij} \right| \leq N^{-m/2+C\varepsilon} N^{-2/3}, \quad \left| N^{-2} [(R V)^4 S]_{ij} \right| \leq N^{-8/3+C\varepsilon}, \tag{3.50}$$

with high probability.

Next, we apply the resolvent expansion to $X_{ij, k}$. Note that in our applications errors of size $O(N^{-8/3-c})$ are affordable in $\Delta X_{ij, k}$ for some $c > 0$ independent of ε (see (3.23) and (3.24)). Now let us assume that the indices i, j, a, b, k satisfy the condition (*) $\{i, j\} \cap \{a, b\} = \emptyset$ and $k \neq i, j, a, b$.

In the applications we shall set $i = \alpha$ and $j = \beta$ in (3.47), and $i = j$ in (3.48). In both cases, it is easy to check that the condition (*) is satisfied for nonvanishing summands.

We can therefore separate $\Delta X_{ij, k}$ into three parts, indexed according to how many V -matrix elements they contain,

$$\Delta X_{ij, k} = \Delta X_{ij, k}^{(1)} + \Delta X_{ij, k}^{(2)} + \Delta X_{ij, k}^{(3)} + O(N^{-3+C\varepsilon}) \tag{3.51}$$

with high probability; here we defined

$$\Delta X_{ij, k}^{(1)} := -N^{-1/2} R_{ik} \overline{(R V R)}_{jk} + [C]_1, \tag{3.52}$$

$$\Delta X_{ij, k}^{(2)} := N^{-1} R_{ik} \overline{(R V R V R)}_{jk} + N^{-1} (R V R)_{ik} \overline{(R V R)}_{jk} + [C]_1, \tag{3.53}$$

$$\Delta X_{ij, k}^{(3)} := -N^{-3/2} R_{ik} \overline{(R V R V R V R)}_{jk} - N^{-3/2} (R V R)_{ik} \overline{(R V R V R)}_{jk} + [C]_2, \tag{3.54}$$

where $[C]_l, l = 1, 2$, means the complex conjugate of the first l terms on the right-hand side with i and j exchanged. Furthermore, it easy to see that the second term on the right-hand side of (3.54) is of order $O(N^{-17/6+C\varepsilon})$. Thus we find with high probability

$$\begin{aligned}
 -\Delta X_{ij,k}^{(3)} &= N^{-3/2} R_{ik} \overline{(RVRVRV R)}_{jk} + N^{-3/2} (RVRVRV R)_{ik} \overline{R}_{jk} \\
 &\quad + O(N^{-17/6+C\varepsilon}) \\
 &= Y + O(N^{-17/6+C\varepsilon}),
 \end{aligned}
 \tag{3.55}$$

where Y is a finite sum of terms of the form

$$N^{-3/2} R_{ik} \overline{(R_{ja} V_{ab} R_{bb} V_{ba} R_{aa} V_{ab} R_{bk})}
 \tag{3.56}$$

and terms obtained from (3.56) by (i) taking the complex conjugate and exchanging i and j , and (ii) exchanging a and b . Using Lemma 3.5 we find that (3.56) is equal to

$$\begin{aligned}
 N^{-3/2} R_{ik} \overline{(R_{ja} V_{ab} R_{bb} V_{ba} R_{aa} V_{ab} R_{bk})} &= N^{-3/2} \overline{m_{sc}^2} R_{ik} \overline{R_{ja} V_{ab} V_{ba} V_{ab} R_{bk}} \\
 &\quad + O(N^{-17/6+C\varepsilon})
 \end{aligned}$$

with high probability. The splitting (3.51) induces a splitting

$$\Delta x_0 = \Delta x_0^{(1)} + \Delta x_0^{(2)} + \Delta x_0^{(3)} + O(N^{-5/3+C\varepsilon})
 \tag{3.57}$$

with high probability in self-explanatory notation. It is easy to see that

$$|\Delta x_0^{(1)}| \leq N^{-1/6+C\varepsilon}, \quad |\Delta x_0^{(2)}| \leq N^{-4/6+C\varepsilon}, \quad |\Delta x_0^{(3)}| \leq N^{-7/6+C\varepsilon}.
 \tag{3.58}$$

From (3.47) and (3.56), we find that $\Delta x_0^{(3)}$ is a finite sum of terms of the form

$$\mathbf{1}(t = 0) \sum_{k \neq \alpha, \beta, a, b} \frac{\overline{\eta m_{sc}^2}}{\pi N^{1/2}} R_{\alpha k} \overline{R_{\beta a} V_{ab} V_{ba} V_{ab} R_{bk}} + O(N^{-3/2+C\varepsilon})
 \tag{3.59}$$

with high probability, where the other terms are obtained from (3.59) as described after (3.56).

Similarly, we find

$$\Delta y_0 = \Delta y_0^{(1)} + \Delta y_0^{(2)} + \Delta y_0^{(3)} + O(N^{-7/3+C\varepsilon})
 \tag{3.60}$$

and

$$|\Delta y_0^{(1)}| \leq N^{-5/6+C\varepsilon}, \quad |\Delta y_0^{(2)}| \leq N^{-8/6+C\varepsilon}, \quad |\Delta y_0^{(3)}| \leq N^{-11/6+C\varepsilon}.
 \tag{3.61}$$

Now we insert these bounds into (3.46). Recall that the upper index l in $\Delta x_0^{(l)}$ and $\Delta y_0^{(l)}$ counts the number of V -matrix elements. Thus we find, recalling (3.46) and the power counting estimates (3.58) and (3.61), that there is a random variable A_5 , depending on the randomness only through Q and the two first moments of h_{ab} , such that

$$\begin{aligned} & \mathbb{E}_{ab}\theta \left[\int_I x^S q(y^S) dE \right] - \mathbb{E}_{ab}\theta \left[\int_I x^R q(y^R) dE \right] \\ &= \mathbb{E}_{ab}\theta' \left[\int_I x^R q(y^R) dE \right] \left[\int_I \left(\Delta x_0^{(3)} q(y^R) + x^R q'(y^R) \Delta y_0^{(3)} \right) dE \right] \\ & \quad + A_4 + A_5 + o(N^{-2+t}). \end{aligned} \tag{3.62}$$

with high probability. Moreover, by the same power counting estimates we find that the second line of (3.62) is bounded by $o(N^{-1})$. We use this rough bound in the case $a = b$, and get

$$\begin{aligned} & \mathbb{E}_{ab}\theta \left[\int_I x^S q(y^S) dE \right] - \mathbb{E}_{ab}\theta \left[\int_I x^R q(y^R) dE \right] \\ &= \mathbf{1}(a \neq b) \mathbb{E}_{ab}\theta' \left[\int_I x^R q(y^R) dE \right] \left[\int_I \left(\Delta x_0^{(3)} q(y^R) + x^R q'(y^R) \Delta y_0^{(3)} \right) dE \right] \\ & \quad + A_4 + A_5 + o(N^{-2+t}) + o(N^{-2+\mathbf{1}(a=b)}) \end{aligned} \tag{3.63}$$

with high probability.

Hence Lemma 3.6 is proved if we can show that, for $a \neq b$, we have

$$\mathbb{E}\theta' \left[\int_I x^R q(y^R) dE \right] \left[\int_I \left(\Delta x_0^{(3)} q(y^R) + x^R q'(y^R) \Delta y_0^{(3)} \right) dE \right] = o(N^{-2}) \tag{3.64}$$

with high probability. This is proved below. □

Proof of (3.64) We shall prove, for $a \neq b$, that

$$\mathbb{E}\theta' \left[\int_I x^R q(y^R) dE \right] \left[\int_I \Delta x_0^{(3)} q(y^R) dE \right] = o(N^{-2}). \tag{3.65}$$

The other term on the left-hand side of (3.64) is estimated similarly. Let us abbreviate

$$B^R := \theta' \left[\int_I x^R q(y^R) dE \right]. \tag{3.66}$$

From (3.36) and the assumption on θ , we find that $|B^R| \leq N^{C\varepsilon}$ with high probability.

We shall estimate the contribution to (3.65) of one term of the form (3.59). Recalling that $\mathbb{E}_{ab}|V_{ab}|^3 = O(1)$ and $m_{sc} = O(1)$, we find the bound

$$\begin{aligned} & N^{-1/6+C\varepsilon} \max_{k \neq \alpha, \beta, a, b} \left| \mathbb{E} B^R \int_I R_{\alpha k} \overline{R_{\beta a} R_{bk}} q(y^R) dE \right| + o(N^{-2}) \\ & \leq N^{-5/6+C\varepsilon} \max_{k \neq \alpha, \beta, a, b} \sup_{E \in I} \left| \mathbb{E} B^R R_{\alpha k} \overline{R_{\beta a} R_{bk}} q(y^R) \right| + o(N^{-2}). \end{aligned} \tag{3.67}$$

The proof of (3.64) is therefore complete if we can show that, assuming the sets $\{\alpha, \beta\}, \{a\}, \{b\}, \{k\}$ are disjoint, we have

$$\left| \mathbb{E} R_{\alpha k} \overline{R_{\beta a} R_{bk}} B^R q(y^R) \right| \leq N^{-4/3+C\varepsilon}. \tag{3.68}$$

In order to prove (3.68), we first use a simple resolvent expansion to show that with high probability

$$\left| R_{\alpha k} \overline{R_{\beta a} R_{bk}} B^R q(y^R) - S_{\alpha k} \overline{S_{\beta a} S_{bk}} B^S q(y^S) \right| \leq N^{-4/3+C\varepsilon}, \tag{3.69}$$

where B^S is defined analogously to (3.66) with R replaced by S . Therefore it suffices to prove

$$\left| \mathbb{E} S_{\alpha k} \overline{S_{\beta a} S_{bk}} B^S q(y^S) \right| \leq N^{-4/3+C\varepsilon}. \tag{3.70}$$

In order to complete the proof, we introduce some notation. Recall that $H \equiv H_{\gamma-1}$ and $S = (H - z)^{-1}$. We define $H^{(a)}$ as the matrix obtained from H by setting its a -th column and a -th row to be zero. For any function $F \equiv F(H)$ we define $F^{(a)} := F(H^{(a)})$. We now remove the a -th row and column from H in (3.70), which we can do with a negligible error. The key identity is the following resolvent identity, proved in Lemma 4.2 of [15]: For $k \neq i, j$ we have

$$S_{ij} = S_{ij}^{(k)} + \frac{S_{ik} S_{kj}}{S_{kk}}. \tag{3.71}$$

Using (3.71), one readily sees that

$$\left| S_{\alpha k} \overline{S_{\beta a} S_{\beta k}} B^S q(y^S) - S_{\alpha k}^{(a)} \overline{S_{\beta a} S_{\beta k}^{(a)}} \theta' \left[\int_I (x^S)^{(a)} q((y^S)^{(a)}) dE' \right] q((y^S)^{(a)}) \right| \leq N^{-4/3+C\epsilon}. \tag{3.72}$$

Moreover, we have

$$\begin{aligned} & S_{\alpha k}^{(a)} \overline{S_{\beta a} S_{\beta k}^{(a)}} \theta' \left[\int_I (x^S)^{(a)} q((y^S)^{(a)}) dE' \right] q((y^S)^{(a)}) \\ &= \left(S_{\alpha k} \overline{S_{\beta k} \theta'} \left[\int_I x^S q(y^S) dE' \right] q(y^S) \right)^{(a)} \overline{S_{\beta a}}. \end{aligned} \tag{3.73}$$

Next, we claim that the conditional expectation – with respect to the variables in the a -th column of H – of $S_{\beta a}$ is much smaller than its typical size. To that end, we use the identities, valid for $i \neq j$,

$$S_{ij} = -S_{ii} \sum_{k \neq i} h_{ik} S_{kj}^{(i)}, \quad S_{ij} = -S_{jj} \sum_{k \neq j} S_{ik}^{(j)} h_{kj}, \tag{3.74}$$

proved in [8], Lemma 6.10. Now using (3.74) we find

$$-S_{\beta a} = \sum_{j \neq a} S_{aa} S_{\beta j}^{(a)} h_{ja} = \sum_{j \neq a} m_{sc} S_{\beta j}^{(a)} h_{ja} + (S_{aa} - m_{sc}) \sum_{j \neq a} S_{\beta j}^{(a)} h_{ja}. \tag{3.75}$$

The conditional expectation with respect to the variables in the a -th column of H applied to the first term on the right-hand side of (3.75) vanishes; hence its contribution to the expectation of (3.73) also vanishes. In order to estimate the second term on the right-hand side of (3.73), we note that with high probability

$$|S_{aa} - m_{sc}| \leq N^{-1/3+C\epsilon},$$

by Lemma 3.5. Moreover, using the large deviation bound (3.9) in [17], we get with high probability

$$\begin{aligned} \left| \sum_{j \neq a} S_{\beta j}^{(a)} h_{ja} \right| &\leq N^{-1/2+\epsilon} \left(\sum_{j \neq a} |S_{\beta j}^{(a)}| \right)^2 \Big|^{1/2} \leq N^{-1/2+\epsilon} |S_{\beta\beta}^{(a)}| + N^\epsilon \max_{j \neq a, \beta} |S_{\beta j}^{(a)}| \\ &\leq N^{-1/3+C\epsilon}, \end{aligned}$$

where in the last step we used (3.71) and Lemma 3.5. Putting everything together, we find that the expectation of (3.73) is bounded in absolute value by $N^{-4/3+C\varepsilon}$. By (3.72), this completes the proof of (3.70), and hence of (3.64).

4 Extension to eigenvalues and several arguments

In this section we describe how the arguments of Sect. 3 extend to general functions θ as in (1.11).

Consider first the case of a single eigenvalue, λ_β , in which case the claim reads

$$\lim_{N \rightarrow \infty} [\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}] \theta(N^{2/3}(\lambda_\beta - \gamma_\beta)) = 0, \tag{4.1}$$

uniformly in $\beta \leq \varphi^\rho$. Denote by $\varrho^{\mathbf{v}}$ and $\varrho^{\mathbf{w}}$ the laws of λ_β in the ensembles $H^{\mathbf{v}}$ and $H^{\mathbf{w}}$ respectively. Using Theorem 2.2 we find

$$\mathbb{E}^{\mathbf{u}} \theta(N^{2/3}(\lambda_\beta - \gamma_\beta)) = \int_I \theta(N^{2/3}(E - \gamma_\beta)) \varrho^{\mathbf{u}}(dE) + O(e^{-\varphi^c}), \tag{4.2}$$

where \mathbf{u} stands for either \mathbf{v} or \mathbf{w} , and I was defined in (3.3). Now integration by parts yields

$$\begin{aligned} & [\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}] \theta(N^{2/3}(\lambda_\beta - \gamma_\beta)) \\ &= -[\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}] \int_I N^{2/3} \theta'(N^{2/3}(E - \gamma_\beta)) \mathbf{1}(\lambda_\beta \leq E) dE + O(e^{-\varphi^c}), \end{aligned} \tag{4.3}$$

where the boundary terms are of order $O(e^{-\varphi^c})$ by Theorem 2.2. Next, we choose a smooth nondecreasing function r_β that vanishes on the interval $(-\infty, \beta - 2/3]$ and is equal to 1 on the interval $[\beta - 1/3, \infty)$. Recalling the definition (2.15), we get from (4.3)

$$\begin{aligned} & \mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}} \theta(N^{2/3}(\lambda_\beta - \gamma_\beta)) \\ &= -\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}} \int_I N^{2/3} \theta'(N^{2/3}(E - \gamma_\beta)) r_\beta(\mathcal{N}(E_L, E)) dE + O(e^{-\varphi^c}) \\ &= -[\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}] \int_I N^{2/3} \theta'(N^{2/3}(E - \gamma_\beta)) \\ & \quad \times r_\beta(\text{Tr}(\mathbf{1}_{[E_L, E]} * \theta_{\tilde{\eta}})(H)) dE + O(\varphi^C N^{-\varepsilon}), \end{aligned}$$

where in the second step we used the assumption on θ , that r'_β is bounded, and Lemma 2.7 with $\tilde{\eta} := N^{-2/3-6\varepsilon}$. More precisely, we apply Lemma 2.7 to estimate, with high

probability,

$$\begin{aligned}
 & \varphi^C N^{2/3} \int_I dE |\text{Tr}(\mathbf{1}_{[E_L, E]} * \theta_{\tilde{\eta}})(H) - \mathcal{N}(E_L, E)| \\
 & \leq \varphi^C N^{2/3} \int_I dE \left(N^{-\varepsilon} + \mathcal{N}(E - N^{-2/3-\varepsilon}, E + N^{-2/3-\varepsilon}) \right) \\
 & \leq \varphi^C N^{-\varepsilon} + \varphi^C N^{2/3} \sum_{\alpha=1}^{\varphi^C} \int_I dE \mathbf{1}(|E - \lambda_{\alpha}| \leq N^{-2/3-\varepsilon}) \\
 & \leq \varphi^C N^{-\varepsilon},
 \end{aligned} \tag{4.4}$$

where the first step follows from (2.18) and the second from Theorem 2.2.

Integrating by parts again, we find with high probability

$$\begin{aligned}
 \mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}} \theta(N^{2/3}(\lambda_{\beta} - \gamma_{\beta})) &= \mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}} \int_I \theta(N^{2/3}(E - \gamma_{\beta})) r'_{\beta} (\text{Tr}(\mathbf{1}_{[E_L, E]} * \theta_{\tilde{\eta}})(H)) \\
 & \times N \text{Im } m(E + i\tilde{\eta}) dE + O(\varphi^C N^{-\varepsilon}).
 \end{aligned}$$

Now we may apply the Green function comparison method from Sect. 3.1. In fact, in this case the analysis is easier as we have no fixed indices i and j to keep track of.

The general case, θ as in (1.11), is treated similarly. Repeating successively the above procedure for each argument $\lambda_{\beta_1}, \dots, \lambda_{\beta_k}$, we find that there is a constant C_k , depending on k , such that

$$\begin{aligned}
 & [\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}] \theta \left(N^{2/3}(\lambda_{\beta_1} - \gamma_{\beta_1}), \dots, N^{2/3}(\lambda_{\beta_k} - \gamma_{\beta_k}); N\tilde{u}_{\alpha_1}(i_1)u_{\alpha_1}(j_1), \dots, N\tilde{u}_{\alpha_k}(i_k)u_{\alpha_k}(j_k) \right) \\
 & = [\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}] \int_{I^k} dE_1 \cdots dE_k \theta \left(N^{2/3}(E_1 - \gamma_{\beta_1}), \dots, N^{2/3}(E_k - \gamma_{\beta_k}); \zeta_1, \dots, \zeta_k \right) \\
 & \times \prod_{l=1}^k \left[r'_{\beta_l} (\text{Tr}(\mathbf{1}_{[E_L, E_l]} * \theta_{\tilde{\eta}})(H)) N \text{Im } m(E_l + i\tilde{\eta}) \right] + O(\varphi^{C_k} N^{-\varepsilon}),
 \end{aligned} \tag{4.5}$$

where we introduced the shorthand

$$\zeta_l := \frac{N}{\pi} \int_I d\tilde{E} \tilde{G}_{i_l j_l}(\tilde{E} + i\eta) q_{\alpha_l} [\text{Tr} \mathbf{1}_{[E_L, \tilde{E}-1]} * \theta_{\tilde{\eta}}(H)],$$

and set $\eta := N^{-2/3-\varepsilon}$; q_{α} is the function from Lemma 3.2. Here at each step we used the assumption on θ , that r'_{β} is bounded, and the estimate

$$\begin{aligned}
 \int_I dE N \text{Im } m(E + i\tilde{\eta}) & \leq \mathcal{N}(-\infty, E + N^{-2/3-\varepsilon/10}) + N^{-\varepsilon/10} \\
 & \leq N n_{sc}(-2 + \varphi^C N^{-2/3}) + 1 \leq \varphi^C,
 \end{aligned} \tag{4.6}$$

where in the first step we used (2.19), in the second Theorem 2.2, and in the third the definition (1.7) of n_{sc} .

The randomness on the right-hand side of (4.5) is expressed entirely in terms of Green functions; hence (4.5) is amenable to the Green function comparison method from Sect. 3.1. The complications are merely notational, as we now have $2k$ fixed indices $i_1, j_1, \dots, i_k, j_k$ instead of just the two i, j .

5 Eigenvectors in the bulk: proof of Theorem 1.10

In this section we prove Theorem 1.10. In the bulk the eigenvalue spacing is of order N^{-1} as opposed to $N^{-2/3}$ at the edge. Thus, we shall have to take spectral windows of size $\eta = N^{-1-\varepsilon}$. To that end, we begin by extending the strong local semicircle law from Theorem 2.1 to arbitrarily small values of $\eta > 0$. Recall the notation $z = E + i\eta$.

Lemma 5.1 *For any $|E| \leq 5$ and $0 < \eta \leq 10$, we have with high probability*

$$\max_{ij} |G_{ij}(z) - \delta_{ij}m_{sc}(z)| \leq \varphi^C \left(\sqrt{\frac{\text{Im } m_{sc}(z)}{N\eta}} + \frac{1}{N\eta} \right) \tag{5.1}$$

for large enough N .

Proof By Theorem 2.1, we only need to consider $\eta \leq y := \varphi^{C_1} N^{-1}$ for some $C_1 > 0$. We use the trivial bound

$$\text{Im } G_{ii}(E + i\eta) \leq \frac{y}{\eta} \text{Im } G_{ii}(E + iy) \quad \text{for } \eta \leq y,$$

as well as

$$|G_{ij}(E + i\eta)| \leq C \log N \max_k \text{Im } G_{kk}(E + i\eta),$$

which follows by a simple dyadic decomposition; see [15], Eq. (4.9). Thus we get

$$|G_{ij}(E + i\eta)| \leq C \log N \frac{y}{\eta} \max_k \text{Im } G_{kk}(E + iy) \leq \varphi^C \frac{y}{\eta} \leq \frac{\varphi^C}{N\eta}.$$

This completes the proof. □

The strategy behind the proof of Theorem 1.10 is very similar to that of Theorem 1.6, given in Sect. 3. In a first step, we express the eigenvector components using integrals involving resolvent matrix elements G_{ij} ; in a second step, we replace the sharp indicator functions in the integrand by smoothed out functions which depend only on the resolvent; in a third step, we use the Green function comparison method to complete the proof.

For ease of presentation, we shall give the proof for the case $\theta = \theta(N\bar{u}_\alpha(i)u_\alpha(j))$; we show that

$$\lim_{N \rightarrow \infty} [\mathbb{E}^{\mathbf{V}} - \mathbb{E}^{\mathbf{W}}] \theta(N\bar{u}_\alpha(i)u_\alpha(j)) = 0, \tag{5.2}$$

where $\rho N \leq \alpha \leq (1-\rho)N$. As outlined in Sect. 4, the extension to general functions θ , as given in (1.15), is an easy extension which we sketch briefly at the end of this section.

We now spell out the three steps mentioned above.

Step 1. The analogue of Lemma 3.1 in the bulk is the following result whose proof uses (1.5) and Lemma 2.5, and is very similar to the proof of Lemma 3.1 (in fact somewhat easier). We omit further details.

Lemma 5.2 *Under the assumption of Theorem 1.10, for any $\varepsilon > 0$ there exist constants C_1, C_2 such that for $\eta = N^{-1-\varepsilon}$ we have*

$$\lim_{N \rightarrow \infty} \max_{\rho N \leq \alpha \leq (1-\rho)N} \max_{i,j} \left\{ \mathbb{E}^{\mathbf{u}} \theta (N \bar{u}_\alpha(i) u_\alpha(j)) - \mathbb{E}^{\mathbf{u}} \theta \left[\frac{N}{\pi} \int_{I_\alpha} \tilde{G}_{ij}(E + i\eta) \mathbf{1}(\lambda_{\alpha-1} \leq E^- \leq \lambda_\alpha) dE \right] \right\} = 0, \tag{5.3}$$

where

$$E^\pm := E \pm (\varphi_N)^{C_1} \eta, \quad I_\alpha := \left[\gamma_\alpha - N^{-1}(\varphi_N)^{C_2}, \gamma_\alpha + N^{-1}(\varphi_N)^{C_2} \right] \tag{5.4}$$

and we introduce the convention $\lambda_0 = -\infty$. Here \mathbf{u} stands for either \mathbf{v} or \mathbf{w} .

Step 2. We choose $\eta = N^{-1-\varepsilon}$ for some small enough $\varepsilon > 0$ and express the indicator function in

$$\mathbb{E}^{\mathbf{u}} \theta \left[\frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) \mathbf{1}(\lambda_{\alpha-1} \leq E^- \leq \lambda_\alpha) dE \right] \tag{5.5}$$

using Green functions (as before, we write $I_\alpha \equiv I$). Using Theorem 2.2, we know that

$$(5.5) = \mathbb{E}^{\mathbf{u}} \theta \left[\frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta) \mathbf{1}(\mathcal{N}(E_L, E^-) = \alpha - 1) dE \right] + o(1), \tag{5.6}$$

where $E_L := -2 - \varphi^C N^{-2/3}$.

As explained in Sect. 1.3, the approach in Step 2 has to be modified slightly from the one employed in Sect. 3. The reason is that the size of the interval $[E_L, E^-]$ is no longer small, but of order one.

For any $E_1, E_2 \in [-3, 3]$ and $\eta_d > 0$ we define $f(\lambda) \equiv f_{E_1, E_2, \eta_d}(\lambda)$ to be the characteristic function of $[E_1, E_2]$ smoothed on scale η_d ; i.e. $f = 1$ on $[E_1, E_2]$, $f = 0$ on $\mathbb{R} \setminus [E_1 - \eta_d, E_2 + \eta_d]$ and $|f'| \leq C\eta_d^{-1}$, $|f''| \leq C\eta_d^{-2}$. Let $q \equiv q_\alpha : \mathbb{R} \rightarrow \mathbb{R}_+$ be a smooth cutoff function concentrated around $\alpha - 1$, satisfying

$$q(x) = 1 \quad \text{if } |x - \alpha + 1| \leq 1/3, \quad q(x) = 0 \quad \text{if } |x - \alpha + 1| \geq 2/3.$$

Now we choose $\eta_d := N^{-1-d\epsilon}$, for some fixed $d > 2$. Then, using Lemma 5.1 and an argument similar to the proof of Lemma 3.2, we find that

$$(5.6) = \mathbb{E}^{\mathbf{u}\theta} \left[\frac{N}{\pi} \int_I \tilde{G}_{ij}(E + i\eta)q (\text{Tr } f_{E_L, E^-, \eta_d}(H)) \, dE \right] + o(1). \tag{5.7}$$

To simplify notation, we follow the conventions of Sect. 3 in writing $I \equiv I_\alpha, q \equiv q_\alpha$ and $f_E \equiv f_{E_L, E^-, \eta_d}$, and set $\alpha = i$ and $\beta = j$. In this notation, we need to estimate

$$[\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}]\theta \left[\frac{N}{\pi} \int_I \tilde{G}_{\alpha\beta}(E + i\eta)q (\text{Tr } f_E(H)) \, dE \right]. \tag{5.8}$$

Now we express $f_E(H)$ in terms of Green functions using Helffer–Sjöstrand functional calculus. Let $\chi(y)$ be a smooth cutoff function with support in $[-1, 1]$, with $\chi(y) = 1$ for $|y| \leq 1/2$ and with bounded derivatives. Then we have (see e.g. Equation (B.12) of [13])

$$f_E(\lambda) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \frac{i\sigma f_E''(e)\chi(\sigma) + i f_E(e)\chi'(\sigma) - \sigma f_E'(e)\chi'(\sigma)}{\lambda - e - i\sigma} \, ded\sigma. \tag{5.9}$$

Thus we get

$$\begin{aligned} \text{Tr } f_E(H) &= \frac{N}{2\pi} \int_{\mathbb{R}^2} \left(i\sigma f_E''(e)\chi(\sigma) + i f_E(e)\chi'(\sigma) - \sigma f_E'(e)\chi'(\sigma) \right) m(e + i\sigma) \, ded\sigma \\ &= \frac{N}{2\pi} \int_{\mathbb{R}^2} \left(i f_E(e)\chi'(\sigma) - \sigma f_E'(e)\chi'(\sigma) \right) m(e + i\sigma) \, ded\sigma \\ &\quad + \frac{iN}{2\pi} \int_{|\sigma| > \tilde{\eta}_d} \, d\sigma \chi(\sigma) \int \, de f_E''(e)\sigma m(e + i\sigma) \\ &\quad + \frac{iN}{2\pi} \int_{-\tilde{\eta}_d}^{\tilde{\eta}_d} \, d\sigma \int \, de f_E''(e)\sigma m(e + i\sigma), \end{aligned} \tag{5.10}$$

where we introduced the parameter $\tilde{\eta}_d := N^{-1-(d+1)\epsilon}$. We shall treat the last term of (5.10) as an error term. From (5.1) we find with high probability

$$\sigma m(e + i\sigma) \leq \frac{\varphi^C}{N}.$$

Therefore the third term of (5.10) is bounded, with high probability, by

$$\left| \frac{iN}{\pi} \int_{-\tilde{\eta}_d}^{\tilde{\eta}_d} d\sigma \int de f_E''(e)\sigma m(e + i\sigma) \right| \leq \varphi^C \tilde{\eta}_d \eta_d^{-1} = \varphi^C N^{-\varepsilon}, \tag{5.11}$$

where in the first step we used that $\int |f_E''(e)|de = O(\eta_d^{-1})$.

Step 3 We estimate (5.8) using a Green function comparison argument, similarly to Sect. 3.1. As in Sect. 3.1, we use the notation

$$x(E) = \frac{N\eta}{\pi} \sum_{k \neq \alpha, \beta} G_{\alpha k}(E + i\eta) \overline{G_{\beta k}(E + i\eta)}. \tag{5.12}$$

Similarly to Lemma 3.4, we begin by dropping the diagonal terms. Using Lemma 5.1 we find

$$\int_I \left| \frac{N}{\pi} \tilde{G}_{\alpha\beta}(E + i\eta) - x(E) \right| dE \leq \varphi^C N\eta^2 \leq N^{-1+C\varepsilon} \tag{5.13}$$

with high probability, so that it suffices to prove

$$[\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}]\theta \left[\int_I x(E)q(\text{Tr } f_E(H)) dE \right] = o(1). \tag{5.14}$$

Using (5.11) we find that it suffices to prove

$$[\mathbb{E}^{\mathbf{v}} - \mathbb{E}^{\mathbf{w}}]\theta \left[\int_I x(E)q(y(E) + \tilde{y}(E))dE \right] = o(1), \tag{5.15}$$

where

$$y(E) := \frac{N}{2\pi} \int_{\mathbb{R}^2} i\sigma f_E''(e)\chi(\sigma)m(e + i\sigma)\mathbf{1}(|\sigma| \geq \tilde{\eta}_d)ded\sigma, \tag{5.16}$$

$$\tilde{y}(E) := \frac{N}{2\pi} \int_{\mathbb{R}^2} \left(i f_E(e)\chi'(\sigma) - \sigma f_E'(e)\chi'(\sigma) \right) m(e + i\sigma)ded\sigma. \tag{5.17}$$

By a telescopic expansion similar to (3.25), we find that (5.15) follows if we can prove, with high probability,

$$\begin{aligned} & \mathbb{E}\theta \left[\int_I x^S(E)q((y + \tilde{y})^S(E))dE \right] - \mathbb{E}\theta \left[\int_I x^R(E)q((y + \tilde{y})^R(E))dE \right] \\ &= A + o(N^{1(a=b)-2}), \end{aligned} \tag{5.18}$$

where we use the notation of Sect. 3.1; here A is a random variable that depends on the randomness only through Q and the first four moments of h_{ab} if $a \neq b$, and the first two moments of h_{ab} if $a = b$. (As in Sect. 3.1, \mathbb{E} denotes expectation with respect to the product measure of the \mathbf{v} and \mathbf{w} ensembles.)

Now we prove (5.18). We use the resolvent expansion

$$\begin{aligned} S &= R - N^{-1/2}RV R + N^{-1}(RV)^2R - N^{-3/2}(RV)^3R \\ &\quad + N^{-2}(RV)^4R - N^{-5/2}(RV)^5S. \end{aligned} \tag{5.19}$$

Similarly to Sect. 3.1, we decompose

$$\Delta m := m^S - m^R = \Delta m_0 + \Delta m_1,$$

where

$$\Delta m_r := \frac{1}{N} \sum_i (S_{ii} - R_{ii}) \mathbf{1}(r = \mathbf{1}(i \in \{a, b\})).$$

Using (5.19) we can expand Δm_r , for $|\sigma| \geq \tilde{\eta}_d$ and with high probability,

$$\Delta m_r(e + i\sigma) = \sum_{p=1}^4 \Delta m_r^{(p)}(e + i\sigma) + O(N^{-5/2+C\varepsilon} \Lambda_\sigma^{2-2r} N^{-r}), \tag{5.20}$$

where

$$|\Delta m_r^{(p)}| \leq N^{-p/2+C\varepsilon} \Lambda_\sigma^{2-2r} N^{-r} \tag{5.21}$$

with high probability, and $\Delta m_r^{(p)}$ is a polynomial in the matrix elements of R and V , each term containing precisely p matrix elements of V ; here we set $\Lambda_\sigma := \sup_{|e| \leq 5} \max_{i \neq j} |G_{ij}(e + i\sigma)|$. Putting both cases $r = 1, 2$ together, we get, for $|\sigma| \geq \tilde{\eta}_d$ and with high probability,

$$\begin{aligned} \Delta m &= \sum_{p=1}^4 \Delta m^{(p)} + O(N^{-5/2+C\varepsilon} (\Lambda_\sigma^2 + N^{-1})), \quad |\Delta m^{(p)}| \leq N^{-p/2+C\varepsilon} (\Lambda_\sigma^2 + N^{-1}). \end{aligned} \tag{5.22}$$

We may now estimate the variables x , y , and \tilde{y} . Let us first consider the variables \tilde{y} . From the definition of χ , we find that in the integrand of (5.17) we have $\sigma \geq c$ and

therefore by Theorem 2.1 we have $\Lambda_\sigma \leq \varphi^C N^{-1/2}$ with high probability. Thus we get from (5.17)

$$\Delta \tilde{y}(E) = \sum_{p=1}^4 \Delta \tilde{y}^{(p)}(E) + O(N^{-5/2+C\varepsilon}), \quad |\Delta \tilde{y}^{(p)}(E)| \leq N^{-p/2+C\varepsilon} \tag{5.23}$$

with high probability.

In order to estimate the contributions of the variables y , we integrate by parts, first in e and then in σ , to obtain

$$\begin{aligned} & \left| N \int_{\mathbb{R}^2} \sigma f_E''(e) \chi(\sigma) \Delta m^{(p)}(e + i\sigma) \mathbf{1}(|\sigma| \geq \tilde{\eta}_d) de d\sigma \right| \\ & \leq CN \left| \int de f_E'(e) \tilde{\eta}_d \Delta m^{(p)}(e + i\tilde{\eta}_d) \right| \\ & \quad + CN \left| \int de f_E'(e) \int_{\tilde{\eta}_d}^\infty d\sigma \chi'(\sigma) \sigma \Delta m^{(p)}(e + i\sigma) \right| \\ & \quad + CN \left| \int de f_E'(e) \int_{\tilde{\eta}_d}^\infty d\sigma \chi(\sigma) \Delta m^{(p)}(e + i\sigma) \right|. \end{aligned} \tag{5.24}$$

Using (5.22), it is easy to see that the sum of the two first terms of (5.24) is bounded by $N^{-p/2+C\varepsilon}$. In order to estimate the last term of (5.24), we use (5.22) and (5.1) to get the bound

$$CN \int_{\tilde{\eta}_d}^1 d\sigma \left(\frac{1}{N\sigma} + \frac{1}{(N\sigma)^2} + \frac{1}{N} \right) N^{-p/2+C\varepsilon} \leq N^{-p/2+C\varepsilon}.$$

Thus we find that

$$\Delta y(E) = \sum_{p=1}^4 \Delta y^{(p)}(E) + O(N^{-5/2+C\varepsilon}), \quad |\Delta y^{(p)}(E)| \leq N^{-p/2+C\varepsilon} \tag{5.25}$$

with high probability.

Finally, as in (3.40), we find that

$$\Delta x(E) = \sum_{p=1}^4 \Delta x^{(p)}(E) + O(N^{-3/2+C\varepsilon}) \tag{5.26}$$

with high probability. Moreover, we have the bound

$$\int_I |\Delta x^{(p)}(E)| dE \leq N^{-p/2+C\varepsilon} \quad (5.27)$$

with high probability. This concludes our estimate of the terms in the resolvent expansion of x , y , and \tilde{y} .

Now using the power counting bounds from (5.23), (5.25), (5.26), and (5.27), we may easily complete the Green function comparison argument to prove (5.18), as in Sect. 3.1.

Finally, we comment on how to deal with more general observables θ , as in (1.15). The basic strategy is the same as in Sect. 4. In fact, the argument is simpler because the errors made in replacing sharp indicator functions with smooth indicator functions are easier to control in the bulk. In Sect. 4, the relatively large errors arising from the soft edge of the function $\mathbf{1}_{[E_L, E]} * \theta_{\tilde{\eta}}$ were controlled by Lemma 2.7. In the bulk, we replace $\mathbf{1}_{[E_L, E]}$ with the function f_E whose edges are sharper. Thus, in the bulk the error resulting from this replacement is bounded by $\mathcal{N}(E^- - \eta_d, E^- + \eta_d)$, whose integral may be estimated exactly as in (4.4). The estimate (4.6) is replaced by the trivial estimate $\int_I dE |\partial_E \text{Tr} f_E(H)| \leq \varphi^C$.

Acknowledgments The authors would like to thank L. Erdős and H.T. Yau for many insights and helpful discussions.

References

1. Anderson, G., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices Studies in advanced mathematics, **118**. Cambridge University Press, Cambridge (2009)
2. Bleher, P., Its, A.: Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model. *Ann. Math.* **150**, 185–266 (1999)
3. Deift, P.: Orthogonal polynomials and random matrices: a Riemann–Hilbert approach. *Courant Lecture Notes in Mathematics*, vol. 3. American Mathematical Society, Providence (1999)
4. Deift, P., Gioev, D.: Random Matrix Theory: Invariant Ensembles and Universality. *Courant Lecture Notes in Mathematics*, vol. 18. American Mathematical Society, Providence (2009)
5. Deift, P., Kriecherbauer, T., McLaughlin, K.T.-R., Venakides, S., Zhou, X.: Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Commun. Pure Appl. Math.* **52**, 1335–1425 (1999)
6. Deift, P., Kriecherbauer, T., McLaughlin, K.T.-R., Venakides, S., Zhou, X.: Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.* **52**, 1491–1552 (1999)
7. Dyson, F.J.: A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3**, 1191–1198 (1962)
8. Erdős, L., Knowles, A., Yau, H.-T., Yin, J.: Spectral statistics of Erdős–Rényi graphs II: eigenvalue spacing and the extreme eigenvalues. arXiv:1103.3869 (2011, preprint)
9. Erdős, L., Schlein, B., Yau, H.-T.: Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.* **37**(3), 815–852 (2009)
10. Erdős, L., Schlein, B., Yau, H.-T.: Local semicircle law and complete delocalization for Wigner random matrices. *Commun. Math. Phys.* **287**, 641–655 (2009)
11. Erdős, L., Schlein, B., Yau, H.-T.: Wegner estimate and level repulsion for Wigner random matrices. *Int. Math. Res. Notices* **2010**(3), 436–479 (2010)
12. Erdős, L., Schlein, B., Yau, H.-T.: Universality of random matrices and local relaxation flow. *Invent. Math.* arXiv:0907.5605 (2011, to appear, preprint)

13. Erdős, L., Ramirez, J., Schlein, B., Yau, H.-T.: Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation. *Electr. J. Probab.* **15**, 526–604 (2010)
14. Erdős, L., Schlein, B., Yau, H.-T., Yin, J.: The local relaxation flow approach to universality of the local statistics for random matrices. *Ann. Inst. H. Poincaré Probab. Stat.* arXiv:0911.3687 (2011, to appear, preprint)
15. Erdős, L., Yau, H.-T., Yin, J.: Bulk universality for generalized Wigner matrices. arXiv:1001.3453 (2011, preprint)
16. Erdős, L., Yau, H.-T., Yin, J.: Universality for generalized Wigner matrices with Bernoulli distribution. *J. Combin.* arXiv:1003.3813 (2011, to appear, preprint)
17. Erdős, L., Yau, H.-T., Yin, J.: Rigidity of eigenvalues of generalized Wigner matrices. arXiv:1007.4652 (2011, preprint)
18. Gustavsson, J.: Gaussian Fluctuations of Eigenvalues in the GUE. *Ann. Inst. H. Poincaré Probab. Stat.* **41**(2), 151–178 (2005)
19. Johansson, K.: Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Commun. Math. Phys.* **215**(3), 683–705 (2001)
20. Johansson, K.: Universality for certain Hermitian Wigner matrices under weak moment conditions. arXiv:0910.4467 (2011, preprint)
21. O’Rourke, S.: Gaussian fluctuations of eigenvalues in Wigner random matrices. *J. Stat. Phys.* **138**(6), 1045–1066 (2009)
22. Pastur, L., Shcherbina, M.: Bulk universality and related properties of Hermitian matrix models. *J. Stat. Phys.* **130**(2), 205–250 (2008)
23. Sinai, Y., Soshnikov, A.: A refinement of Wigner’s semicircle law in a neighborhood of the spectrum edge. *Funct. Anal. Appl.* **32**(2), 114–131 (1998)
24. Sodin, S.: The spectral edge of some random band matrices. arXiv:0906.4047 (2011, preprint)
25. Soshnikov, A.: Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.* **207**(3), 697–733 (1999)
26. Tao, T., Vu, V.: Random matrices: universality of the local eigenvalue statistics. *Acta Math.* arXiv:0906.0510 (2011, to appear, preprint)
27. Tao, T., Vu, V.: Random matrices: universality of local eigenvalue statistics up to the edge. arXiv:0908.1982 (2011, preprint)
28. Tao, T., Vu, V.: Random matrices: universal properties of eigenvectors. arXiv:1103.2801 (2011, preprint)
29. Tracy, C., Widom, H.: Level-spacing distributions and the Airy kernel. *Commun. Math. Phys.* **159**, 151–174 (1994)