The Annals of Probability
2013, Vol. 41, No. 3B, 2279–2375
DOI: 10.1214/11-AOP734

SPECTRAL STATISTICS OF ERDŐS–RÉNYI GRAPHS I: LOCAL SEMICIRCLE LAW

BY LÁSZLÓ ERDŐS[1], ANTTI KNOWLES[2], HORNG-TZER YAU[2,3] AND JUN YIN[4]

University of Munich, Harvard University, Harvard University and University of Wisconsin

We consider the ensemble of adjacency matrices of Erdős–Rényi random graphs, that is, graphs on N vertices where every edge is chosen independently and with probability $p \equiv p(N)$. We rescale the matrix so that its bulk eigenvalues are of order one. We prove that, as long as $pN \to \infty$ (with a speed at least logarithmic in N), the density of eigenvalues of the Erdős–Rényi ensemble is given by the Wigner semicircle law for spectral windows of length larger than N^{-1} (up to logarithmic corrections). As a consequence, all eigenvectors are proved to be completely delocalized in the sense that the ℓ^∞-norms of the ℓ^2-normalized eigenvectors are at most of order $N^{-1/2}$ with a very high probability. The estimates in this paper will be used in the companion paper [Spectral statistics of Erdős–Rényi graphs II: Eigenvalue spacing and the extreme eigenvalues (2011) Preprint] to prove the universality of eigenvalue distributions both in the bulk and at the spectral edges under the further restriction that $pN \gg N^{2/3}$.

1. Introduction. The universality of random matrices has been a central subject since the pioneering work of Wigner [40], Gaudin [27], Mehta [30] and Dyson [12]. The problem can roughly be divided into the bulk universality in the interior of the spectrum and the edge universality near the spectral edge. The bulk and edge universalities for invariant ensembles have been extensively studied; see, for example, [4, 7, 8, 31] and [1, 6, 9] for a review. A key contributing factor to the progress in the study of invariant ensembles is the existence of explicit formulas for the joint density function of the eigenvalues. There is no such explicit formula for noninvariant ensembles and, hence, our understanding of them is much more limited. The most prominent examples for noninvariant ensembles are the Wigner matrices with i.i.d. non-Gaussian matrix elements. The edge universality of Wigner matrices can be proved via the moment method and its various generalizations; see,

Received April 2011; revised November 2011.
[1]Supported in part by SFB-TR 12 Grant of the German Research Council.
[2]Supported in part by NSF Grant DMS-07-57425.
[3]Supported in part by NSF Grant DMS-08-04279.
[4]Supported in part by NSF Grant DMS-10-01655.
MSC2010 subject classifications. 15B52, 82B44.
Key words and phrases. Erdős–Rényi graphs, local semicircle law, density of states.

for example, [33–35]. The bulk universality for general classes of Wigner matrices was listed in Mehta's book [30] as Conjectures 1.2.1 and 1.2.2 on page 7. We shall refer to these two conjectures collectively as the Wigner–Dyson–Gaudin–Mehta conjecture, in recognition of the pioneering works of Wigner, Dyson, Gaudin and Mehta listed above. It remained unsolved until very recently. This is mainly due to the fact that all existing methods on local eigenvalue statistics depended on explicit formulas, which are not available for Wigner matrices. In a series of papers [14, 16–23], a new approach to understanding the local eigenvalue statistics was developed. In particular, it led to the first proof [14] of the Wigner–Dyson–Gaudin–Mehta conjecture for Hermitian Wigner matrices whose entries have smooth distributions. This approach is based on three basic ingredients: (1) a local semicircle law—a precise estimate of the local eigenvalue density down to energy scales containing around $(\log N)^C$ 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 [12]; (3) a density argument which shows that for any probability distribution of the matrix elements there exists a Gaussian divisible distribution such that the two associated Wigner ensembles have identical local eigenvalue statistics down to the scale $1/N$. Furthermore, the edge universality can also be obtained by some modifications of (1) and (3) [23]. The class of ensembles to which this method applies is extremely general; in particular, it includes any (generalized) Wigner matrices under the sole assumption that the distributions of the matrix elements have a uniform subexponential decay. We remark that the universality of Wigner matrices, under certain restrictions on the distribution of the matrix entries, was also established in [36, 37]. We shall discuss these results in the companion paper [13].

In this paper and its companion [13], we extend the approach (1)–(3) to cover a class of sparse matrices. This class includes the Erdős–Rényi matrices, which we now introduce. Symmetric $N \times N$ matrices with 0–1 entries arise naturally as adjacency matrices of graphs on N vertices. Since every nonoriented graph can be uniquely characterized by its adjacency matrix, we shall from now talk about matrix ensembles (with 0–1 entries) and graph ensembles interchangeably. We shall always normalize the matrices so that their spectra typically lie in an interval of length of order one. One common random graph ensemble is the Erdős–Rényi graph [24, 25]. In it each edge is chosen independently and with probability $p \equiv p(N)$. Since each row and column of the adjacency matrix has typically pN nonzero entries, it is sparse as long as $p \ll 1$. We shall refer to pN as the sparseness parameter of the matrix.

Our goal in this paper, and in its companion [13], is to establish both the bulk and edge universalities for the Erdős–Rényi ensemble under the restriction $pN \gg N^{2/3}$. In other words, we prove that the eigenvalue gap distributions in the bulk and near the edges are given by those of the Gaussian Orthogonal Ensemble (GOE) provided that $pN \gg N^{2/3}$. We remark that the law of the Erdős–Rényi ensemble is even more singular than that of the Bernoulli Wigner matrices, since the matrix elements are highly concentrated around 0. Another way of expressing the singular nature of the Erdős–Rényi ensemble is to say that the moments of the matrix entries decay much more slowly than in the case of Wigner matrices.

The matrix elements of the Erdős–Rényi ensemble take on the values 0 and 1. Hence, they do not satisfy the mean zero condition which typically appears in the random matrix literature. Due to the nonzero mean of the entries, the largest eigenvalue of the Erdős–Rényi ensemble is very large and far away from the rest of the spectrum, which by our normalization lies in the interval $[-2, 2]$. By the edge universality of the Erdős–Rényi ensemble, we therefore mean that the probability distribution of the second largest eigenvalue is given by the distribution of the largest eigenvalue of the GOE, which is the well-known Tracy–Widom distribution.

As the first step of the general strategy to establish universality, we shall prove a local semicircle law, Theorem 2.9, stating that the eigenvalue distribution of the Erdős–Rényi ensemble in any spectral window of size η containing on average $N\eta \sim (\log N)^C$ eigenvalues is given by the Wigner semicircle law. Theorem 2.9 is valid in the bulk and at the edges as long as the parameter $p \equiv p(N)$ satisfies $pN \to \infty$ with a rate at least logarithmic in N. Similar results but for much larger spectral windows [of lengths at least $\eta \sim (pN)^{-1/10}$] were recently proved in [38].

We note that the semicircle law for Wigner matrices in spectral windows of size $\eta \sim N^{-1/2}$ has been known for some time [2, 29]. The semicircle law in the smallest possible spectral window (of size $\eta \gtrsim N^{-1}$ in the bulk) was established in [17, 18]. This estimate, referred to as the local semicircle law, has become a fundamental tool in the proofs of the universality of random matrices in [14] as well as in the subsequent works [19, 37]. The local semicircle law in [17, 18] is optimal in terms of the range of η, but the error estimates, of order $(N\eta)^{-1/2}$ in the bulk and with a coefficient deteriorating near the spectral edges, were not optimal. Optimal error estimates, uniform throughout the entire spectrum and valid for more general classes of Wigner matrices, were obtained in [23]. The local semicircle law proved in this paper can also be viewed as a generalization of the results in [23] in two unrelated directions: (a) the laws of the matrix entries are much more singular, and (b) the matrix entries have nonzero mean.

Besides eigenvalues, eigenvectors also play a fundamental role in the theory of random matrices. One important motivation for their study is that random matrices can be viewed as mean-field approximations of random Schrödinger operators where delocalization of eigenfunctions is a key signature for the metallic or conducting phase. Another question about eigenvectors of random graphs is the size of their nodal domains, which can studied using delocalization bounds [10]. It was first proved in [16] that eigenvectors for Wigner matrices are completely delocalized, partly motivated by a conjecture of T. Spencer. The method was refined in [17, 18], and was also adapted in [37, 38]. The key observation behind the proof is that the delocalization estimate for eigenvectors follows from the local semicircle law provided that the spectral windows can be made sufficiently small. Thus, Theorem 2.9 also implies, with $\mathbf{v}_\alpha$ denoting the ℓ^2-normalized eigenvectors, that

$$\max\{\|\mathbf{v}_\alpha\|_\infty : 1 \le \alpha \le N\} \le \frac{(\log N)^C}{\sqrt{N}} \tag{1.1}$$

holds with exponentially high probability with some constant C (Theorem 2.16). This establishes the complete delocalization of all eigenvectors as long as the sparseness parameter pN increases at least logarithmically in N. In particular, this result gives the optimal answer to a question posed in Section 3.3 of [10], asking whether $\|\mathbf{v}_\alpha\|_\infty \le N^{-1/2+o(1)}$ holds for all eigenvectors $\mathbf{v}_\alpha$. In fact, this question was originally posed for fixed p, but our result shows that the bound conjectured in [10] holds even for $p \ge (\log N)^C N^{-1}$. It was recently proved in [38] that $\|\mathbf{v}_\alpha\|_\infty \le (pN)^{-1/2}$ away from the spectral edge; some earlier results were obtained in [11]. These results established only the lower bound pN on the localization length; the complete delocalization (1.1) corresponds to the optimal localization length of order N.

Our main result on the bulk and edge universalities will require a further condition

$$pN \gg N^{2/3}. \tag{1.2}$$

This and related issues will be discussed in the second paper [13].

2. Definitions and results. We begin this section by introducing a class of $N \times N$ sparse random matrices $A \equiv A_N$. Here N is a large parameter. (Throughout the following we shall often refrain from explicitly indicating N-dependence.)

The motivating example is the *Erdős–Rényi matrix*, or the adjacency matrix of the *Erdős–Rényi random graph*. Its entries are independent (up to the constraint that the matrix be symmetric), and equal to 1 with probability p and 0 with probability $1-p$. For our purposes it is convenient to replace p with the new parameter $q \equiv q(N)$, defined through $p = q^2/N$. Moreover, we rescale the matrix in such a way that its bulk eigenvalues typically lie in an interval of size of order one.

Thus, we are led to the following definition. Let $A = (a_{ij})$ be the symmetric $N \times N$ matrix whose entries a_{ij} are independent (up to the symmetry constraint $a_{ij} = a_{ji}$) and each element is distributed according to

$$a_{ij} = \frac{\gamma}{q} \begin{cases} 1, & \text{with probability } \dfrac{q^2}{N}, \\ 0, & \text{with probability } 1 - \dfrac{q^2}{N}. \end{cases} \tag{2.1}$$

Here $\gamma := (1 - q^2/N)^{-1/2}$ is a scaling introduced for convenience. The parameter $q \le N^{1/2}$ expresses the sparseness of the matrix; it may depend on N. Since A typically has $q^2 N$ nonvanishing entries, we find that if $q \ll N^{1/2}$, then the matrix is sparse.

We extract the mean of each matrix entry and write

$$A = H + \gamma q |\mathbf{e}\rangle\langle\mathbf{e}|,$$

where the entries of H (given by $h_{ij} = a_{ij} - \gamma q/N$) have mean zero, and we defined the vector

$$\mathbf{e} \equiv \mathbf{e}_N := \frac{1}{\sqrt{N}}(1, \dots, 1)^T. \tag{2.2}$$

(As above, we often neglect the subscript N of $\mathbf{e}$; the precise value of this subscript will always be clear from the context.) Here we use the notation $|\mathbf{e}\rangle\langle\mathbf{e}|$ to denote the orthogonal projection onto $\mathbf{e}$, that is, $(|\mathbf{e}\rangle\langle\mathbf{e}|)_{ij} := N^{-1}$.

One readily finds that the matrix elements of H satisfy the moment bounds

$$\mathbb{E}h_{ij}^2 = \frac{1}{N}, \qquad \mathbb{E}|h_{ij}|^p \le \frac{1}{Nq^{p-2}}, \tag{2.3}$$

where $p \ge 2$.

More generally, we consider the following class of random matrices with non-centered entries characterized by two parameters q and f, which may be N-dependent. The parameter q expresses how singular the distribution of h_{ij} is; in particular, it expresses the sparseness of A for the special case (2.1). The parameter f determines the nonzero expectation value of the matrix elements.

Throughout the following we shall make use of a (possibly N-dependent) quantity $\xi \equiv \xi_N$ satisfying

$$1 + a_0 \le \xi \le A_0 \log\log N \tag{2.4}$$

for some fixed positive constants $a_0 > 0$ and $A_0 \ge 10$. The parameter ξ will be used as an exponent in logarithmic corrections as well as probability estimates.

DEFINITION 2.1 (H). Fix a parameter $\xi \equiv \xi_N$ satisfying (2.4). We consider $N \times N$ random matrices $H = (h_{ij})$ whose entries are real and independent up to the symmetry constraint $h_{ij} = h_{ji}$. We assume that the elements of H satisfy the moment conditions

$$\mathbb{E}h_{ij} = 0, \qquad \mathbb{E}|h_{ij}|^2 = \frac{1}{N}, \qquad \mathbb{E}|h_{ij}|^p \le \frac{C^p}{Nq^{p-2}} \tag{2.5}$$

for $1 \le i, j \le N$ and $3 \le p \le (\log N)^{A_0 \log\log N}$, where C is a positive constant. Here $q \equiv q(N)$ satisfies

$$(\log N)^{3\xi} \le q \le CN^{1/2} \tag{2.6}$$

for some positive constant C.

Note that the entries of H exhibit a slow decay of moments. The variance is of order N^{-1}, but higher moments decay at a rate proportional to inverse powers of q and not $N^{1/2}$. Thus, unlike for Wigner matrices, the entries of sparse matrices satisfying Definition 2.1 do not have a natural scale. (The entries of a Wigner matrix live on the scale $N^{-1/2}$, which means that the high moments decay at a rate proportional to inverse powers of $N^{1/2}$. See Remark 2.5 below for a more precise statement.)

DEFINITION 2.2 (A). Let H satisfy Definition 2.1. Define

$$A := H + f|\mathbf{e}\rangle\langle\mathbf{e}|, \tag{2.7}$$

where $f \equiv f(N)$ is a deterministic number that satisfies

$$0 \le f \le N^C \tag{2.8}$$

for some constant $C > 0$.

REMARK 2.3. For definiteness, and bearing the Erdős–Rényi matrix in mind, we restrict ourselves to real symmetric matrices satisfying Definition 2.2. However, our proof applies equally to complex Hermitian sparse matrices.

REMARK 2.4. To simplify the presentation, we assume that all matrix elements of H have identical variance $1/N$. As in [22], Section 5, one may, however, easily generalize this condition and require that the variances be bounded by C/N and their column sums (hence, also the row sums) be equal to 1. Thus, one may, for instance, consider Erdős–Rényi graphs in which a vertex cannot link to itself (i.e., the diagonal elements of A vanish).

REMARK 2.5. In particular, we may take H to be a Wigner matrix whose entries have subexponential decay,

$$\mathbb{P}(N^{1/2}|h_{ij}| \ge x) \le C \exp(-x^{1/\theta})$$

for some positive constants θ and C. Indeed, in this case we get

$$\mathbb{E}h_{ij} = 0, \qquad \mathbb{E}|h_{ij}|^2 = \frac{1}{N}, \qquad \mathbb{E}|h_{ij}|^p \le C\frac{(\theta p)^{\theta p}}{N^{p/2}} \qquad (p \ge 3).$$

Now we choose

$$q := N^{1/2}(\theta(\log N)^{A_0 \log\log N})^{-\theta}.$$

Since $q^{-1} \le (\log N)^{C \log\log N} N^{-1/2}$, we find that all factors q^{-1} in error estimates such as (2.16) and (2.17) below may be replaced with $N^{-1/2}$ at the expense of a larger exponent in the preceding logarithmic factors. In fact, using Lemma 3.2 below, it is easy to see that in this case all terms depending on q in estimates such as (2.16) and (2.17) may dropped, as they are bounded by the other error terms. In particular, Theorem 2.8 generalizes Theorem 2.1 of [23].

We shall frequently have to deal with events of very high probability, for which the following definition is useful. It is characterized by two positive parameters, ξ and ν, where ξ is subject to (2.4).

DEFINITION 2.6 (High probability events). We say that an N-dependent event Ω holds with (ξ,ν)-*high probability* if

$$\mathbb{P}(\Omega^c) \le \mathrm{e}^{-\nu(\log N)^\xi} \tag{2.9}$$

for $N \ge N_0(\nu, a_0, A_0)$.

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

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

for $N \ge N_0(\nu, a_0, A_0)$.

REMARK 2.7. In the following we shall not keep track of the explicit value of ν; in fact, we allow ν to decrease from one line to another without introducing a new notation. It will be clear from the proof that such reductions of ν occur only at a few, finitely many steps. Hence, all of our results will hold for $\nu \le \nu_0$, where ν_0 depends only on the constants C in Definition 2.1 and the parameter Σ in (2.10) below. (In particular, ν is independent of ξ.)

We shall use C and c to denote generic positive constants which may only depend on the constants in assumptions such as (2.4) and (2.5). Typically, C denotes a large constant and c a small constant. Note that the fundamental large parameter of our model is N, and the notation $\gg, \ll, O(\cdot), o(\cdot)$ always refers to the limit $N \to \infty$. Here $a \ll b$ means $a = o(b)$. We write $a \sim b$ for $C^{-1}a \le b \le Ca$.

We now list our results. We introduce the spectral parameter

$$z = E + \mathrm{i}\eta,$$

where $E \in \mathbb{R}$ and $\eta > 0$. Let $\Sigma \ge 3$ be a fixed but arbitrary constant and define the domain

$$D := \{z \in \mathbb{C} : |E| \le \Sigma, 0 < \eta \le 3\}. \tag{2.10}$$

We define the density of the semicircle law

$$\varrho_{\mathrm{sc}}(x) := \frac{1}{2\pi}\sqrt{[4 - x^2]_+}, \tag{2.11}$$

and, for $\operatorname{Im} z > 0$, its Stieltjes transform

$$m_{\mathrm{sc}}(z) := \int_{\mathbb{R}} \frac{\varrho_{\mathrm{sc}}(x)}{x - z}\,\mathrm{d}x. \tag{2.12}$$

The Stieltjes transform $m_{\mathrm{sc}}(z) \equiv m_{\mathrm{sc}}$ may also be characterized as the unique solution of

$$m_{\mathrm{sc}} + \frac{1}{z + m_{\mathrm{sc}}} = 0 \tag{2.13}$$

satisfying $\operatorname{Im} m_{\mathrm{sc}}(z) > 0$ for $\operatorname{Im} z > 0$. We define the resolvent

$$G(z) := (H - z)^{-1}$$

as well as the Stieltjes transform of the empirical eigenvalue density

$$m(z) := \frac{1}{N} \operatorname{Tr} G(z).$$

For $x \in \mathbb{R}$ we define the distance κ_x to the spectral edge through

$$\kappa_x := \big||x| - 2\big|. \tag{2.14}$$

THEOREM 2.8 (Local semicircle law for H). *There are universal constants $C_1, C_2 > 0$ such that the following holds. Suppose that H satisfies Definition 2.1. Moreover, assume that*

$$\xi = \frac{A_0(1 + o(1))}{2} \log\log N, \qquad q \geq (\log N)^{C_1 \xi}. \tag{2.15}$$

Then there is a constant $\nu > 0$, depending on A_0, Σ and the constants C in (2.5) and (2.6), such that the following holds.

We have the local semicircle law: the event

$$\bigcap_{z \in D} \left\{ |m(z) - m_{\mathrm{sc}}(z)| \leq (\log N)^{C_2 \xi} \left(\min\left\{ \frac{1}{q^2 \sqrt{\kappa_E + \eta}}, \frac{1}{q} \right\} + \frac{1}{N\eta} \right) \right\} \tag{2.16}$$

holds with (ξ, ν)-high probability. Moreover, we have the following estimate on the individual matrix elements of G. The event

$$\begin{aligned} \bigcap_{z \in D} \Bigg\{ & \max_{1 \leq i,j \leq N} |G_{ij}(z) - \delta_{ij} m_{\mathrm{sc}}(z)| \\ & \leq (\log N)^{C_2 \xi} \left(\frac{1}{q} + \sqrt{\frac{\operatorname{Im} m_{\mathrm{sc}}(z)}{N\eta}} + \frac{1}{N\eta} \right) \Bigg\} \end{aligned} \tag{2.17}$$

holds with (ξ, ν)-high probability.

The results of Theorem 2.8 may be interpreted as follows. Consider first the bulk, that is, $\kappa_E \geq c > 0$. Then Theorem 2.8 states roughly that

$$|m(z) - m_{\mathrm{sc}}(z)| \lesssim \frac{1}{q^2} + \frac{1}{N\eta}, \qquad |G_{ij}(z) - \delta_{ij} m_{\mathrm{sc}}(z)| \lesssim \frac{1}{q} + \frac{1}{\sqrt{N\eta}}, \tag{2.18}$$

up to logarithmic factors. Since $|m_{sc}(z)| \sim 1$, both estimates are stable in a sense that they identify the leading order terms of m and G_{ii} down to the optimal scale $\eta \gtrsim N^{-1}$. Note that choosing $\eta \lesssim N^{-1}$ in

$$\operatorname{Im} m(z) = \frac{1}{N\eta} \sum_{\alpha} \frac{\eta^2}{(E-\lambda_\alpha)^2 + \eta^2}$$

allows one to resolve individual eigenvalues λ_α of H. Therefore, below the scale $\eta \lesssim N^{-1}$ the quantities m and G_{ii} become strongly fluctuating and these fluctuations are larger than the main term. In the regime $\eta \geq (\log N)^{C\xi} N^{-1}$ in which the fluctuations are smaller than the main term, a spectral window of size η contains at least $(\log N)^{C\xi}$ eigenvalues, hence, an averaging takes place.

The factor $1/q$ on the right-hand side of the second inequality of (2.18) arises from the strong fluctuations of the matrix entries h_{ij}, which take on values of size q^{-1} with probability of order $q^2 N^{-1}$. Indeed, it is apparent from the representations (3.13) and (3.23) that $G_{ij} = m_{sc}^2 h_{ij} + \cdots$, that is, G_{ij} has a component that fluctuates on the scale q^{-1}. The improvement from q^{-1} to q^{-2} in the first inequality of (2.18) arises from an averaging in the summation $m = N^{-1} \sum_i G_{ii}$. If the random variables in the average were independent, one would expect the averaging to yield an improvement of order $N^{-1/2}$; however, in our case there are strong dependencies, which result in the more modest gain of order q^{-1}.

At the edge ($\kappa_E = 0$), the estimates (2.16) and (2.17) may be roughly stated as

$$|m(z) - m_{sc}(z)| \lesssim \frac{1}{q} + \frac{1}{N\eta}, \qquad |G_{ij}(z) - \delta_{ij} m_{sc}(z)| \lesssim \frac{1}{q} + \frac{\eta^{1/4}}{\sqrt{N\eta}} + \frac{1}{N\eta}.$$

Now we formulate the local semicircle law for the matrix A given in Definition 2.2. Define the quantities

$$\widetilde{G}(z) := (A - z)^{-1}, \qquad \widetilde{m}(z) := \frac{1}{N} \operatorname{Tr} \widetilde{G}(z). \tag{2.19}$$

THEOREM 2.9 (Local semicircle law for A). *There are universal constants $C_1, C_2 > 0$ such that the following holds. Suppose that A satisfies Definition 2.2, and that ξ and q satisfy (2.15). Then there is a constant $\nu > 0$—depending on A_0, Σ and the constants C in (2.5), (2.6) and (2.8)—such that the following holds.*

We have the local semicircle law: the event

$$\bigcap_{z \in D} \left\{ |\widetilde{m}(z) - m_{sc}(z)| \leq (\log N)^{C_2 \xi} \left(\min\left\{ \frac{1}{q^2 \sqrt{\kappa_E + \eta}}, \frac{1}{q} \right\} + \frac{1}{N\eta} \right) \right\} \tag{2.20}$$

holds with (ξ, ν)-high probability. Moreover, we have the following estimate on the individual matrix elements of $\widetilde{G}$. If the assumption (2.8) is strengthened to

$$0 \leq f \leq C_0 N^{1/2} \tag{2.21}$$

for some constant C_0, then the event

$$\bigcap_{z\in D}\left\{\max_{1\le i,j\le N}|\widetilde{G}_{ij}(z)-\delta_{ij}m_{\mathrm{sc}}(z)| \le (\log N)^{C_2\xi}\left(\frac{1}{q}+\sqrt{\frac{\operatorname{Im} m_{\mathrm{sc}}(z)}{N\eta}}+\frac{1}{N\eta}\right)\right\} \tag{2.22}$$

holds with (ξ,ν)-high probability, where ν also depends on C_0.

Next, let $\lambda_1\le\cdots\le\lambda_N$ be the ordered family of eigenvalues of H, and let $\mathbf{u}_1,\dots,\mathbf{u}_N$ denote the associated eigenvectors. Similarly, we denote the ordered eigenvalues of A by $\mu_1\le\cdots\le\mu_N$ and the associated eigenvectors by $\mathbf{v}_1,\dots,\mathbf{v}_N$. We use the notation $\mathbf{u}_\alpha=(u_\alpha(i))_{i=1}^N$ and $\mathbf{v}_\alpha=(v_\alpha(i))_{i=1}^N$ for the vector components. All eigenvectors are ℓ^2-normalized and have real components.

We state our main result about the local density of states of A. For $E_1<E_2$ define the counting functions

$$\mathcal{N}_{\mathrm{sc}}(E_1,E_2):=N\int_{E_1}^{E_2}\varrho_{\mathrm{sc}}(x)\,\mathrm{d}x,\qquad \widetilde{\mathcal{N}}(E_1,E_2):=|\{\alpha:E_1<\mu_\alpha\le E_2\}|. \tag{2.23}$$

THEOREM 2.10 (Local density of states). *Suppose that A satisfies Definition 2.2 and that ξ and q satisfy (2.15). Then there is a constant $\nu>0$—depending on A_0, Σ and the constants C in (2.5), (2.6) and (2.8)—as well as a constant $C>0$ such that the following holds.*

For any E_1 and E_2 satisfying $E_2\ge E_1+(\log N)^{C\xi}N^{-1}$ we have

$$\widetilde{\mathcal{N}}(E_1,E_2) = \mathcal{N}_{\mathrm{sc}}(E_1,E_2) \times\left[1+O\left((\log N)^{C\xi}\left(\frac{1}{N(E_2-E_1)^{3/2}}+\frac{1}{q^2(E_2-E_1)}\right)\right)\right] \tag{2.24}$$

with (ξ,ν)-high probability.

Away from the spectral edge we have a stronger statement. Fix $\kappa_*>0$. Then, for any E_1 and E_2 satisfying $E_2\ge E_1+(\log N)^{C\xi}N^{-1}$ as well as $\kappa_{E_1}\ge\kappa_*$ and $\kappa_{E_2}\ge\kappa_*$, we have

$$\widetilde{\mathcal{N}}(E_1,E_2)=\mathcal{N}_{\mathrm{sc}}(E_1,E_2)\left[1+O\left((\log N)^{C\xi}\left(\frac{1}{N(E_2-E_1)}+\frac{1}{q^2}\right)\right)\right] \tag{2.25}$$

with (ξ,ν)-high probability, where the constant in $O(\cdot)$ depends on κ_*.

REMARK 2.11. Both results (2.24) and (2.25) are special cases of a more general, uniform, estimate; see Proposition 8.2.

In the recent work [38], the asymptotics of the local density of states was also established, but only in much larger spectral windows, of size at least $(Np)^{-1/10} = q^{-1/5}$, and with a weaker error estimate.

Our next result concerns the integrated densities of states,

$$n_{\mathrm{sc}}(E) := \frac{1}{N}\mathcal{N}_{\mathrm{sc}}(-\infty, E), \qquad \widetilde{\mathfrak{n}}(E) := \frac{1}{N}\widetilde{\mathcal{N}}(-\infty, E). \tag{2.26}$$

THEOREM 2.12 (Integrated density of states). *Suppose that A satisfies Definition 2.2 and that ξ and q satisfy (2.15). Then there is a constant $\nu > 0$—depending on A_0, Σ and the constants C in (2.5), (2.6) and (2.8)—as well as a constant $C > 0$ such that the event*

$$\bigcap_{E\in[-\Sigma,\Sigma]} \left\{ |\widetilde{\mathfrak{n}}(E) - n_{\mathrm{sc}}(E)| \le (\log N)^{C\xi}\left(\frac{1}{N} + \frac{1}{q^3} + \frac{\sqrt{\kappa_E}}{q^2}\right)\right\} \tag{2.27}$$

holds with (ξ, ν)-high probability.

Next, we prove that the $N - 1$ first eigenvalues of A are close to their classical locations predicted by the semicircle law. Denote by γ_α the classical location of the αth eigenvalue, defined through

$$n_{\mathrm{sc}}(\gamma_\alpha) = \frac{\alpha}{N} \qquad \text{for } \alpha = 1, \dots, N. \tag{2.28}$$

The following theorem compares the locations of the eigenvalues $\mu_1, \dots, \mu_{N-1}$ to their classical locations $\gamma_1, \dots, \gamma_{N-1}$. It is well known that the largest eigenvalue μ_N of the Erdős–Rényi matrix is much larger than γ_N. This holds for more general sparse matrices as well; more precisely, if $f \ge 1 + c$, then $\mu_N \approx f + f^{-1}$ is separated from $\mu_{N-1} \approx 2$ by a gap of order one. The precise behavior of μ_N in this regime is established in Theorem 6.2 below.

THEOREM 2.13 (Eigenvalue locations). *Suppose that A satisfies Definition 2.2 and that ξ satisfies (2.15). Let ϕ be an exponent satisfying $0 < \phi \le 1/2$, and set $q = N^\phi$. Then there is a constant $\nu > 0$—depending on A_0, Σ and the constants C in (2.5), (2.6) and (2.8)—as well as a constant $C > 0$ such that the following holds.*

We have with (ξ, ν)-high probability that

$$\sum_{\alpha=1}^{N-1} |\mu_\alpha - \gamma_\alpha|^2 \le (\log N)^{C\xi}(N^{1-4\phi} + N^{4/3-8\phi}). \tag{2.29}$$

Moreover, for all $\alpha = 1, \dots, N - 1$ we have with (ξ, ν)-high probability that

$$\begin{aligned} |\mu_\alpha - \gamma_\alpha| \le (\log N)^{C\xi}\big(&N^{-2/3}[\widehat{\alpha}^{-1/3} + \mathbf{1}(\widehat{\alpha} \le (\log N)^{C\xi}(1 + N^{1-3\phi}))] \\ &+ N^{2/3-4\phi}\widehat{\alpha}^{-2/3} + N^{-2\phi}\big), \end{aligned} \tag{2.30}$$

where we abbreviated $\widehat{\alpha} := \min\{\alpha, N - \alpha\}$.

REMARK 2.14. Under the assumption $\phi \geq 1/3$, the estimate (2.30) simplifies to

$$|\mu_\alpha - \gamma_\alpha| \leq (\log N)^{C\xi}(N^{-2/3}\widehat{\alpha}^{-1/3} + N^{-2\phi}), \tag{2.31}$$

which holds with (ξ, ν)-high probability.

REMARK 2.15. Theorems 2.10, 2.12 and 2.13 also hold—with the same proof—for the matrix H. More precisely, Theorem 2.10 holds with $\widetilde{\mathcal{N}}(E_1, E_2)$ replaced with

$$\mathcal{N}(E_1, E_2) := |\{\alpha : E_1 < \lambda_\alpha \leq E_2\}|,$$

Theorem 2.12 holds with $\widetilde{\mathfrak{n}}(E)$ replaced with

$$\mathfrak{n}(E) := \frac{1}{N}\mathcal{N}(-\infty, E),$$

and Theorem 2.13 holds with μ_α replaced with λ_α.

Our final result shows that the eigenvectors of A are *completely delocalized*.

THEOREM 2.16 (Complete delocalization of eigenvectors). *Suppose that A satisfies Definition 2.2 and (2.21). Then there is a constant $\nu > 0$—depending on A_0, Σ and the constants C in (2.5), (2.6) and (2.8)—such that the following statements hold for any ξ satisfying (2.4).*

We have with (ξ, ν)-high probability

$$\max_{\alpha < N} \|\mathbf{v}_\alpha\|_\infty \leq \frac{(\log N)^{4\xi}}{\sqrt{N}}. \tag{2.32}$$

Moreover, we have with (ξ, ν)-high probability

$$\|\mathbf{v}_N - \mathbf{e}\|_2 = \frac{1}{f} + O\left(\sqrt{\frac{1}{f^3} + \frac{(\log N)^{\xi}}{f\sqrt{N}}}\right). \tag{2.33}$$

If additionally $f \leq C$ for some constant C, then we have with (ξ, ν)-high probability

$$\|\mathbf{v}_N\|_\infty \leq \frac{(\log N)^{4\xi}}{\sqrt{N}}. \tag{2.34}$$

Finally, there exists positive constants C, C_0 such that if $f \geq C_0(\log N)^{\xi}$, then we have with (ξ, ν)-high probability

$$\|\mathbf{v}_N - \mathbf{e}\|_\infty \leq C\frac{(\log N)^{\xi}}{\sqrt{N} f}. \tag{2.35}$$

REMARK 2.17. If f does not grow with N, then the components $v_N(i)$ of the largest eigenvector fluctuate, and we do not expect (2.35) to hold. However, a delocalization bound similar to (2.34) holds for all f. In (2.34) this bound was proved for $f \le C$. In fact, a slight modification of our proof yields complete delocalization for the values of f not covered by Theorem 2.16, that is, $1 \ll f \le C_0(\log N)^{\xi}$. We claim that in this case we have with (ξ, ν)-high probability

$$\|\mathbf{v}_N\|_\infty \le \frac{(\log N)^{C\xi}}{\sqrt{N}}. \tag{2.36}$$

The required modifications are sketched at the end of Section 7.3 below.

REMARK 2.18. Similarly, if H satisfies Definition 2.1, all of its eigenvectors are delocalized in the sense that

$$\max_\alpha \|\mathbf{u}_\alpha\|_\infty \le \frac{(\log N)^{4\xi}}{\sqrt{N}}$$

with (ξ, ν)-high probability. The proof is a straightforward application of (3.4) below and the estimate (7.25) applied to G_{jj}.

In the recent work [38], a weaker upper bound of size $(Np)^{-1/2} = q^{-1}$ was established for the ℓ^∞-norm of the eigenvectors of A associated with eigenvalues away from the spectral edge.

3. The weak local semicircle law for *H*. In this section we introduce and prove a weak version of the local semicircle law for the matrix H. This result is weaker than our final result for H, Theorem 2.8, but it will be used as an a priori bound for the proof of Theorem 2.8. Moreover, Theorem 3.1 holds under slightly weaker assumptions on ξ than Theorem 2.8, and is for this reason a more suitable tool for proving eigenvector delocalization, Theorem 2.16; see Section 7.3 for details.

We shall prove Theorem 3.1 (the weak local semicircle law) for spectral parameters z in the set

$$D_L := \{z \in \mathbb{C} : |E| \le \Sigma, (\log N)^L N^{-1} \le \eta \le 3\} \subset D, \tag{3.1}$$

where the parameter $L \equiv L(N)$ will always satisfy

$$L \ge 8\xi. \tag{3.2}$$

THEOREM 3.1 (Weak local semicircle law for H). *Let H satisfy Definition 2.1. Then there are constants $\nu > 0$ and $C > 0$ such that the following statements hold for any ξ satisfying (2.4) and L satisfying (3.2).*

The events

$$\bigcap_{z \in D_L} \left\{ \max_{i \ne j} |G_{ij}(z)| \le \frac{C}{q} + \frac{C(\log N)^{2\xi}}{\sqrt{N\eta}} \right\} \tag{3.3}$$

and

$$\bigcap_{z \in D_L} \left\{ \max_i |G_{ii}(z) - m(z)| \le \frac{C(\log N)^{\xi}}{q} + \frac{C(\log N)^{2\xi}}{\sqrt{N\eta}} \right\} \tag{3.4}$$

hold with (ξ, ν)*-high probability. Furthermore, we have the weak local semicircle law: the event*

$$\bigcap_{z \in D_L} \left\{ |m(z) - m_{\mathrm{sc}}(z)| \le \frac{C(\log N)^{\xi}}{\sqrt{q}} + \frac{C(\log N)^{2\xi}}{(N\eta)^{1/3}} \right\} \tag{3.5}$$

holds with (ξ, ν)*-high probability.*

Roughly, Theorem 3.1 states that

$$|G_{ij} - \delta_{ij} m(z)| \lesssim \frac{1}{q} + \frac{1}{\sqrt{N\eta}} \tag{3.6}$$

and

$$|m(z) - m_{\mathrm{sc}}(z)| \lesssim \frac{1}{\sqrt{q}} + \frac{1}{(N\eta)^{1/3}}. \tag{3.7}$$

Comparing with the strong local semicircle law, Theorem 2.8, we note that the error bound in (3.6) for G_{ij} is already optimal in the bulk. However, unlike Theorem 2.8, the quantity G_{ii} is compared to m and not m_{sc}.

On the other hand, the estimate (3.7) is considerably weaker than the corresponding bound in (2.18). The smaller power $1/3$ in the factor $(N\eta)^{-1/3}$ reflects the instability near the edge; it appears because we insist on having uniform bounds up to the edge. If we were interested only in the bulk, it would be easy to repeat the proof of Theorem 3.1 to obtain $(N\eta\kappa)^{-1/2}$, thus replacing the power $1/3$ with $1/2$. The price would be a coefficient which blows up at the edge.

As in Theorem 2.8, the estimates of Theorem 3.1 are stable down to the optimal scale $\eta \gtrsim N^{-1}$, uniformly up to the edge. Thus, the difference between Theorems 2.8 and 3.1 lies only in the precision of the estimates.

In order to prove Theorem 3.1, we first collect some basic tools and notation.

3.1. *Preliminaries.* The following lemma collects some useful properties of m_{sc} defined in (2.13).

LEMMA 3.2. *For* $z = E + \mathrm{i}\eta \in D_L$ *abbreviate* $\kappa \equiv \kappa_E$. *Then we have*

$$|m_{\mathrm{sc}}(z)| \sim 1, \qquad |1 - m_{\mathrm{sc}}(z)^2| \sim \sqrt{\kappa + \eta}. \tag{3.8}$$

Moreover,

$$\operatorname{Im} m_{\mathrm{sc}}(z) \sim \begin{cases} \sqrt{\kappa + \eta}, & \text{if } |E| \le 2, \\ \dfrac{\eta}{\sqrt{\kappa + \eta}}, & \text{if } |E| \ge 2. \end{cases}$$

Here the implicit constants in $\sim$ *depend on* Σ *in* (2.10).

PROOF. The proof is an elementary calculation; see Lemma 4.2 in [21]. □

In order to streamline notation, we shall often omit the explicit dependence of quantities on the spectral parameter $z \in D_L$; thus, we write, for instance, $G_{ij}(z) \equiv G_{ij}$. Define the z-dependent quantities

$$\Lambda_o := \max_{i \neq j} |G_{ij}|, \qquad \Lambda_d := \max_i |G_{ii} - m_{\mathrm{sc}}|,$$
$$\Lambda := |m - m_{\mathrm{sc}}|, \qquad v_i := G_{ii} - m_{\mathrm{sc}}. \tag{3.9}$$

DEFINITION 3.3. Let $\mathbb{T} \subset \{1, \ldots, N\}$. Then we define $H^{(\mathbb{T})}$ as the $(N - |\mathbb{T}|) \times (N - |\mathbb{T}|)$ minor of H obtained by removing all rows and columns of H indexed by $i \in \mathbb{T}$. Note that we keep the names of indices of H when defining $H^{(\mathbb{T})}$.

More formally, for $i \in \{1, \ldots, N\}$ we define the operation π_i on the probability space by

$$(\pi_i(H))_{kl} := \mathbf{1}(k \neq i)\mathbf{1}(l \neq i)h_{kl}. \tag{3.10}$$

For $\mathbb{T} \subset \{1, \ldots, N\}$ we also write $\pi_{\mathbb{T}} := \prod_{i \in \mathbb{T}} \pi_i$. Then we define

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

The quantities $G^{(\mathbb{T})}(z)$, $\lambda_\alpha^{(\mathbb{T})}$, $\mathbf{u}_\alpha^{(\mathbb{T})}$, etc. are defined in the obvious way using $H^{(\mathbb{T})}$. Here $\alpha = 1, \ldots, \alpha_{\max}$, where $\alpha_{\max} := N - |\mathbb{T}|$.

Moreover, we use the notation

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

and abbreviate $(i) = (\{i\})$ as well as $(\mathbb{T}i) = (\mathbb{T} \cup \{i\})$.

We also set

$$m^{(\mathbb{T})} := \frac{1}{N} \sum_i^{(\mathbb{T})} G_{ii}^{(\mathbb{T})}. \tag{3.11}$$

Note that we choose the normalization N^{-1} instead of the more natural $(N - |\mathbb{T}|)^{-1}$ in (3.11); this is simply a convenient choice for later applications.

The next lemma collects the main identities of the resolvent matrix elements $G_{ij}^{(\mathbb{T})}$. Its proof is elementary linear algebra; see, for example, [22].

LEMMA 3.4. *For $i, j \neq k$ we have*

$$G_{ij} = G_{ij}^{(k)} + \frac{G_{ik} G_{kj}}{G_{kk}}. \tag{3.12}$$

For $i \neq j$ we have

$$G_{ij} = -G_{ii} G_{jj}^{(i)} (h_{ij} - Z_{ij}), \qquad G_{ii} = (h_{ii} - z - Z_{ii})^{-1}, \tag{3.13}$$

where we defined, for arbitrary $i, j \in \{1, \ldots, N\}$,

$$Z_{ij} := \mathbf{h}_i \cdot G^{(ij)} \mathbf{h}_j = \sum_{k,l}^{(ij)} h_{ik} G_{kl}^{(ij)} h_{lj}. \tag{3.14}$$

Here $\mathbf{h}_i$ denotes the vector given by the ith column of H. Note that in expressions of the form (3.14) it is implied that the ith and jth entries of $\mathbf{h}_i$ and $\mathbf{h}_j$ have been removed; we do not indicate this explicitly, as it is always clear from the context.

REMARK 3.5. Lemma 3.4 remains trivially valid for the minors $H^{(\mathbb{T})}$ of H. For instance, (3.12) reads

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

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

DEFINITION 3.6. We denote by $\mathbb{E}_i$ the partial expectation with respect to the variables $\mathbf{h}_i = (h_{ij})_{j=1}^N$, and set $\mathbb{IE}_i X := X - \mathbb{E}_i X$.

We abbreviate

$$Z_i := \mathbb{IE}_i Z_{ii} = \mathbb{IE}_i \sum_{k,l}^{(i)} h_{ik} G_{kl}^{(i)} h_{li} = \sum_{k,l}^{(i)} \left(h_{ik} h_{li} - \frac{1}{N} \delta_{kl} \right) G_{kl}^{(i)}. \tag{3.15}$$

The following trivial large deviation estimate provides a bound on the matrix elements of H.

LEMMA 3.7. *For C large enough we have with (ξ, ν)-high probability*

$$|h_{ij}| \leq \frac{C}{q}.$$

PROOF. The claim follows by choosing $p = \nu (\log N)^{\xi}$ in (2.5) and applying Markov's inequality. □

We collect here the large deviation estimates for random variables whose moments decay slowly. Their proof is given in the Appendix.

LEMMA 3.8. (i) *Let* (a_i) *be a family of centered and independent random variables satisfying*

$$\mathbb{E}|a_i|^p \le \frac{C^p}{N^\gamma q^{\alpha p+\beta}} \tag{3.16}$$

for all $2 \le p \le (\log N)^{A_0 \log\log N}$, *where* $\alpha \ge 0$ *and* $\beta, \gamma \in \mathbb{R}$. *Then there is a* $\nu > 0$, *depending only on* C *in* (3.16), *such that for all* ξ *satisfying* (2.4) *we have with* (ξ, ν)*-high probability*

$$\left|\sum_i A_i a_i\right| \le (\log N)^\xi \left[\frac{\sup_i |A_i|}{q^\alpha} + \left(\frac{1}{N^\gamma q^{\beta+2\alpha}} \sum_i |A_i|^2\right)^{1/2}\right]. \tag{3.17}$$

(ii) *Let* $a_1, \dots, a_N$ *be centered and independent random variables satisfying*

$$\mathbb{E}|a_i|^p \le \frac{C^p}{N q^{p-2}} \tag{3.18}$$

for $2 \le p \le (\log N)^{A_0 \log\log N}$. *Then there is a* $\nu > 0$, *depending only on* C *in* (3.18), *such that for all* ξ *satisfying* (2.4), *and for any* $A_i \in \mathbb{C}$ *and* $B_{ij} \in \mathbb{C}$, *we have with* (ξ, ν)*-high probability*

$$\left|\sum_{i=1}^N A_i a_i\right| \le (\log N)^\xi \left[\frac{\max_i |A_i|}{q} + \left(\frac{1}{N} \sum_{i=1}^N |A_i|^2\right)^{1/2}\right], \tag{3.19}$$

$$\left|\sum_{i=1}^N \bar{a}_i B_{ii} a_i - \sum_{i=1}^N \sigma_i^2 B_{ii}\right| \le (\log N)^\xi \frac{B_d}{q}, \tag{3.20}$$

$$\left|\sum_{1 \le i \ne j \le N} \bar{a}_i B_{ij} a_j\right| \le (\log N)^{2\xi} \left[\frac{B_o}{q} + \left(\frac{1}{N^2} \sum_{i \ne j} |B_{ij}|^2\right)^{1/2}\right], \tag{3.21}$$

where σ_i^2 *denotes the variance of* a_i *and we abbreviated*

$$B_d := \max_i |B_{ii}|, \qquad B_o := \max_{i \ne j} |B_{ij}|.$$

(iii) *Let* $a_1, \dots, a_N$ *and* $b_1, \dots, b_N$ *be independent random variables, each satisfying* (3.18). *Then there is a* $\nu > 0$, *depending only on* C *in* (3.18), *such that for all* ξ *satisfying* (2.4) *and* $B_{ij} \in \mathbb{C}$ *we have with* (ξ, ν)*-high probability*

$$\left|\sum_{i,j=1}^N a_i B_{ij} b_j\right| \le (\log N)^{2\xi} \left[\frac{B_d}{q^2} + \frac{B_o}{q} + \left(\frac{1}{N^2} \sum_{i \ne j} |B_{ij}|^2\right)^{1/2}\right]. \tag{3.22}$$

REMARK 3.9. Note that the estimates (3.19) and (3.20) are special cases of (3.17). The right-hand side of the large deviation bound (3.17) consists of two terms, which can be understood as follows. The first term gives the large deviation bound for the special case where A_i vanishes for all but one i; in this case it is immediate that $|A_i a_i| \leq (\log N)^{\xi} |A_i| q^{-\alpha}$ with (ξ, ν)-high probability. The second term is equal to the variance of $\sum_i A_i a_i$. In particular, (3.17) is optimal (up to factors of $\log N$). The estimates (3.19)–(3.22) can be interpreted similarly. [Note that the powers of q in the estimates (3.21)–(3.22) are not optimal; this is, however, of no consequence for later applications.]

For a family $F_1, \dots, F_N$ we introduce the notation

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

The following lemma contains the self-consistent resolvent equation on which our proof relies.

LEMMA 3.10. *We have the identity*

$$G_{ii} = \frac{1}{-z - m_{\mathrm{sc}} - ([v] - \Upsilon_i)}, \tag{3.23}$$

where

$$\Upsilon_i := h_{ii} - Z_i + \mathcal{A}_i$$

and

$$\mathcal{A}_i := \frac{1}{N} \sum_j \frac{G_{ij} G_{ji}}{G_{ii}}. \tag{3.24}$$

PROOF. The proof is a simple calculation using (3.13) and (3.12). □

3.2. *Basic estimates on the event* $\Omega(z)$.

DEFINITION 3.11. For $z \in D_L$ introduce the event

$$\Omega(z) := \{\Lambda_d(z) + \Lambda_o(z) \leq (\log N)^{-\xi}\} \tag{3.25}$$

and the control parameter

$$\Psi(z) := \sqrt{\frac{\Lambda(z) + \operatorname{Im} m_{\mathrm{sc}}(z)}{N\eta}}. \tag{3.26}$$

Note that $\Psi(z)$ is a random variable. Moreover, on $\Omega(z)$ we have $\Psi(z) \leq C(\log N)^{-4\xi}$ by (3.2).

Throughout the following we shall make use of the fundamental identity

$$\sum_j |G_{ij}|^2 = \sum_j \sum_\alpha \frac{\bar{u}_\alpha(i) u_\alpha(j)}{\lambda_\alpha - z} \sum_\beta \frac{u_\beta(i)\bar{u}_\beta(j)}{\lambda_\beta - \bar{z}} = \sum_\alpha \frac{|u_\alpha(i)|^2}{|\lambda_\alpha - z|^2} = \frac{1}{\eta} \operatorname{Im} G_{ii}. \tag{3.27}$$

A similar identity holds for $H^{(\mathbb{T})}$. Using the lower bound $|m_{\mathrm{sc}}(z)| \geq c$ from (3.8) and the definition (3.25), we find

$$c \leq |G_{ii}(z)| \leq C \tag{3.28}$$

on $\Omega(z)$. Using (3.12) repeatedly, we find that on $\Omega(z)$ we have

$$c \leq \big|G_{ii}^{(\mathbb{T})}(z)\big| \leq C \tag{3.29}$$

for $|\mathbb{T}| \leq 10$ (here 10 can be replaced with any fixed number). Similarly, we have on $\Omega(z)$ that

$$\max_{i \neq j} \big|G_{ij}^{(\mathbb{T})}(z)\big| \leq C\Lambda_o(z) \leq C(\log N)^{-\xi} \tag{3.30}$$

for $|\mathbb{T}| \leq 10$.

LEMMA 3.12. *Fixing $z = E + i\eta \in D_L$, we have for any i and $\mathbb{T} \subset \{1, \ldots, N\}$ satisfying $i \notin \mathbb{T}$ and $|\mathbb{T}| \leq 10$ that*

$$m^{(i\mathbb{T})}(z) = m^{(\mathbb{T})}(z) + O\left(\frac{1}{N\eta}\right) \tag{3.31}$$

holds in $\Omega(z)$.

PROOF. We use (3.12) to write

$$\frac{1}{N} \sum_j^{(i\mathbb{T})} G_{jj}^{(i\mathbb{T})} = \frac{1}{N} \sum_j^{(i\mathbb{T})} G_{jj}^{(\mathbb{T})} - \frac{1}{N} \sum_j^{(i\mathbb{T})} \frac{G_{ji}^{(\mathbb{T})} G_{ij}^{(\mathbb{T})}}{G_{ii}^{(\mathbb{T})}} = \frac{1}{N} \sum_j^{(\mathbb{T})} G_{jj}^{(\mathbb{T})} - \frac{1}{N} \sum_j^{(\mathbb{T})} \frac{G_{ji}^{(\mathbb{T})} G_{ij}^{(\mathbb{T})}}{G_{ii}^{(\mathbb{T})}}.$$

Therefore,

$$\frac{1}{N} \sum_j^{(i\mathbb{T})} G_{jj}^{(i\mathbb{T})} = \frac{1}{N} \sum_j^{(\mathbb{T})} G_{jj}^{(\mathbb{T})} + O\left(\frac{1}{N} \sum_j^{(\mathbb{T})} |G_{ij}^{(\mathbb{T})}|^2\right) = \frac{1}{N} \sum_j^{(\mathbb{T})} G_{jj}^{(\mathbb{T})} + O\left(\frac{1}{N\eta} \operatorname{Im} G_{ii}^{(\mathbb{T})}\right).$$

The claim now follows from (3.29). □

LEMMA 3.13. *For fixed $z \in D_L$ we have on $\Omega(z)$ with (ξ, ν)-high probability*

$$\Lambda_o(z) \le C\left(\frac{1}{q} + (\log N)^{2\xi}\Psi(z)\right), \tag{3.32}$$

$$\max_i |Z_i(z)| \le C\left(\frac{(\log N)^{\xi}}{q} + (\log N)^{2\xi}\Psi(z)\right). \tag{3.33}$$

PROOF. We start with (3.32). Let $i \neq j$. Using (3.13), (3.29), (3.30) and (3.12) we get on $\Omega(z)$ with (ξ, ν)-high probability

$$\begin{aligned} |G_{ij}| &\le C(|h_{ij}| + |Z_{ij}|) \le \frac{C}{q} + C\left|\sum_{k,l}^{(ij)} h_{ik} G_{kl}^{(ij)} h_{lj}\right| \\ &\le \frac{C}{q} + C(\log N)^{2\xi}\frac{\Lambda_o}{q} + C(\log N)^{2\xi}\left(\frac{1}{N^2}\sum_{k,l}^{(ij)}|G_{kl}^{(ij)}|^2\right)^{1/2}, \end{aligned} \tag{3.34}$$

where the last step follows using (3.22) and (2.6). Using (3.12) repeatedly and recalling (3.29), we find on $\Omega(z)$ that $G_{kk}^{(ij)} = G_{kk} + O(\Lambda_o^2)$. Thus, we get on $\Omega(z)$, by (3.27),

$$\frac{1}{N^2}\sum_{k,l}^{(ij)}|G_{kl}^{(ij)}|^2 = \frac{1}{N^2\eta}\sum_{k}^{(ij)} \operatorname{Im} G_{kk}^{(ij)} \le \frac{\operatorname{Im} m}{N\eta} + \frac{C\Lambda_o^2}{N\eta}. \tag{3.35}$$

Taking the maximum over $i \neq j$ in (3.34) therefore yields, on $\Omega(z)$ with (ξ, ν)-high probability,

$$\Lambda_o \le \frac{C}{q} + o(1)\Lambda_o + C(\log N)^{2\xi}\sqrt{\frac{\operatorname{Im} m}{N\eta}},$$

where we used (2.6) and the fact that $N\eta \ge (\log N)^{8\xi}$ by (3.2). This concludes the proof of (3.32).

In order to prove (3.33), we write, recalling the definition (3.15),

$$Z_i = \sum_k^{(i)}\left(|h_{kk}|^2 - \frac{1}{N}\right)G_{kk}^{(i)} + \sum_{k \neq l}^{(i)} h_{ik} G_{kl}^{(i)} h_{li}.$$

Using (3.20), (3.21) and (3.29), we therefore get, on $\Omega(z)$ with (ξ, ν)-high probability,

$$\begin{aligned} |Z_i| &\le \frac{C(\log N)^{\xi}}{q} + C(\log N)^{2\xi}\left[\frac{\Lambda_o}{q} + \left(\frac{1}{N^2}\sum_{k,l}^{(i)}|G_{kl}^{(i)}|^2\right)^{1/2}\right] \\ &\le \frac{C(\log N)^{\xi}}{q} + \frac{C(\log N)^{2\xi}\Lambda_o}{\sqrt{N\eta}} + C(\log N)^{2\xi}\sqrt{\frac{\operatorname{Im} m}{N\eta}}, \end{aligned}$$

similarly to above. Invoking (3.32) and recalling (3.2) finishes the proof. □

We may now estimate Λ_d in terms of Λ.

LEMMA 3.14. *Fix $z = E + i\eta \in D_L$. On $\Omega(z)$ we have with (ξ, ν)-high probability*

$$\max_i |G_{ii}(z) - m(z)| \le C\left(\frac{(\log N)^\xi}{q} + (\log N)^{2\xi}\Psi(z)\right). \tag{3.36}$$

PROOF. We use the resolvent equation (3.23). On $\Omega(z)$ we have $|\mathcal{A}_i| \le C\Lambda_o^2$ and $|h_{ii}| \le C/q$ with (ξ, ν)-high probability by Lemma 3.7. Thus, Lemma 3.13 yields on $\Omega(z)$ with (ξ, ν)-high probability

$$|\Upsilon_i| \le C\left(\frac{(\log N)^\xi}{q} + (\log N)^{2\xi}\Psi(z)\right) \ll 1. \tag{3.37}$$

From (3.23) we therefore get on $\Omega(z)$ with (ξ, ν)-high probability

$$|G_{ii} - G_{jj}| = |G_{ii}||G_{jj}||\Upsilon_i - \Upsilon_j| \le C\left(\frac{(\log N)^\xi}{q} + (\log N)^{2\xi}\Psi(z)\right). \tag{3.38}$$

Since $m = \frac{1}{N}\sum_j G_{jj}$, the proof is complete. □

Note that (3.36) implies

$$\Lambda_d(z) \le \Lambda(z) + C\left(\frac{(\log N)^\xi}{q} + (\log N)^{2\xi}\Psi(z)\right) \tag{3.39}$$

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

3.3. *Stability of the self-consistent equation of $[v]$ on $\Omega(z)$.* We now expand the self-consistent equation into a form in which the stability of the averaged quantity $[v]$ may be analyzed. Recall the definition $v_i := G_{ii} - m_{sc}$.

LEMMA 3.15. *Fix $z \in D_L$. Then we have on $\Omega(z)$ with (ξ, ν)-high probability*

$$\begin{aligned}(1 - m_{sc}^2)[v] &= m_{sc}^3[v]^2 + m_{sc}^2[Z] \\ &\quad + O\left(\frac{(\log N)^{2\xi+1}}{q^2} + (\log N)^{4\xi+1}\Psi^2 + \frac{\Lambda^2}{\log N}\right).\end{aligned} \tag{3.40}$$

PROOF. Recall that on $\Omega(z)$ we have $v_i = o(1)$. Moreover, (2.13) and (3.8) imply that $|m_{sc}(z) + z| = |m_{sc}(z)|^{-1} \ge c$ for $z \in D_L$. With (3.37) we may therefore expand (3.23) on $\Omega(z)$ up to second order to get, with (ξ, ν)-high probability,

$$v_i = m_{sc}^2([v] - \Upsilon_i) + m_{sc}^3([v] - \Upsilon_i)^2 + O([v] - \Upsilon_i)^3. \tag{3.41}$$

Averaging over i in (3.41) yields with (ξ,ν)-high probability

$$(1-m_{sc}^2)[v] = -m_{sc}^2[\Upsilon] + m_{sc}^3[v]^2 - 2m_{sc}^3[v][\Upsilon] + m_{sc}^3[\Upsilon^2] + O\Big([v] + \max_i |\Upsilon_i|\Big)^3.$$

Recall the definition (3.24) of $\mathcal{A}_i$. Using (3.19) and (3.29), we find on $\Omega(z)$ with (ξ,ν)-high probability

$$\begin{aligned}[\Upsilon] &= \frac{1}{N}\sum_i h_{ii} - [Z] + [\mathcal{A}] = -[Z] + O\bigg(\frac{(\log N)^\xi}{N} + \frac{1}{N^2}\sum_{i,j}|G_{ij}|^2\bigg)\\ &= -[Z] + O\bigg(\frac{(\log N)^\xi}{N} + \Psi^2\bigg),\end{aligned}$$

where in the last step we used (3.27). Moreover, recalling that $|[v]| = \Lambda$, we get by Young's inequality

$$-2m_{sc}^3[v][\Upsilon] = O\bigg(\frac{\Lambda^2}{\log N} + (\log N)|[\Upsilon]|^2\bigg).$$

Recalling (3.37), we therefore have

$$\begin{aligned}&(1-m_{sc}^2)[v]\\ &\quad= m_{sc}^3[v]^2 + m_{sc}^2[Z]\\ &\qquad + O\bigg(\frac{(\log N)^\xi}{N} + \Psi^2 + (\log N)|[\Upsilon]|^2 + \max_i|\Upsilon_i|^2 + \Lambda^3 + \frac{\Lambda^2}{\log N}\bigg)\\ &\quad= m_{sc}^3[v]^2 + m_{sc}^2[Z] + O\bigg(\frac{(\log N)^{2\xi+1}}{q^2} + (\log N)^{4\xi+1}\Psi^2 + \frac{\Lambda^2}{\log N}\bigg),\end{aligned}$$

where we used that on $\Omega(z)$ we have $\Lambda \le \Lambda_d \le (\log N)^{-\xi} \le (\log N)^{-1}$. $\square$

Note that, together with (3.33), Lemma 3.15 implies a weak self-consistent equation on $[v]$:

$$(1-m_{sc}^2)[v] = m_{sc}^3[v]^2 + O\bigg(\frac{\Lambda^2}{\log N}\bigg) + O\bigg(\frac{(\log N)^\xi}{q} + (\log N)^{2\xi}\Psi\bigg) \tag{3.42}$$

on $\Omega(z)$ with (ξ,ν)-high probability. Here we used (2.6) and (3.2). For the proof of the weak semicircle law, Theorem 3.1, we shall only use the weaker form (3.42) of the self-consistent equation.

3.4. *Initial estimates for large η.* In order to get the continuity argument of Section 3.6 started, we need some initial estimates on $\Lambda_d + \Lambda_o$ for large η. In other words, we need to prove that $\Omega(E+i\eta)$ is an event of high probability for $\eta \sim 1$.

LEMMA 3.16. *Let $\eta \geq 2$. Then for $z = E + \mathrm{i}\eta \in D_L$ we have*

$$\Lambda_d(z) + \Lambda_o(z) \leq \frac{C(\log N)^{\xi}}{q} + \frac{C(\log N)^{2\xi}}{\sqrt{N}} \leq C(\log N)^{-2\xi}$$

with (ξ, ν)-high probability.

PROOF. Fix $z = E + \mathrm{i}\eta \in D_L$ with $\eta \geq 2$. We shall repeatedly make use of the trivial estimates

$$|G_{ij}^{(\mathbb{T})}| \leq \frac{1}{\eta}, \qquad |m^{(\mathbb{T})}| \leq \frac{1}{\eta}, \qquad |m_{\mathrm{sc}}| \leq \frac{1}{\eta}, \tag{3.43}$$

where $\mathbb{T} \subset \{1, \ldots, N\}$ is arbitrary. These estimates follow immediately from the definitions of $G^{(\mathbb{T})}$ and m_{sc}.

We begin by estimating Λ_o. For $i \neq j$ we get, following the calculation in (3.34) and recalling (3.27), with (ξ, ν)-high probability,

$$|G_{ij}| \leq \frac{C}{q} + o(1)\Lambda_o + C(\log N)^{2\xi}\sqrt{\frac{\operatorname{Im} m^{(ij)}}{N\eta}} \leq \frac{C}{q} + o(1)\Lambda_o + \frac{C(\log N)^{2\xi}}{\sqrt{N}}.$$

Taking the maximum over $i \neq j$ yields with (ξ, ν)-high probability

$$\Lambda_o \leq \frac{C}{q} + \frac{C(\log N)^{2\xi}}{\sqrt{N}}.$$

What remains is an estimate on Λ_d. We begin by estimating with (ξ, ν)-high probability

$$|\Upsilon_i| \leq \frac{C}{q} + |Z_i| + |\mathcal{A}_i|.$$

In order to estimate $|\mathcal{A}_i|$, we observe that (3.13) implies

$$\frac{G_{ij}}{G_{ii}} = -G_{jj}^{(i)}(h_{ij} - Z_{ij}) \qquad (i \neq j).$$

Therefore, we have with (ξ, ν)-high probability

$$\begin{aligned} |\mathcal{A}_i| &\leq \frac{1}{N}|G_{ii}| + \frac{1}{N}\sum_j^{(i)} |G_{jj}^{(i)}||G_{ji}|(|h_{ij}| + |Z_{ij}|) \\ &\leq \frac{C}{N} + C\Lambda_o\left(\frac{1}{q} + \sup_{i \neq j}|Z_{ij}|\right) \leq \frac{C}{q}, \end{aligned} \tag{3.44}$$

where we used that with (ξ, ν)-high probability

$$|Z_{ij}| \leq (\log N)^{2\xi}\left[\frac{C}{q^2} + \frac{\Lambda_o}{q} + \frac{C}{\sqrt{N}}\right]$$

as follows from (3.22) and (3.27). Similarly, from (3.15) and using (3.20) and (3.21), we find with (ξ,ν)-high probability

$$|Z_i| \le \frac{C(\log N)^{\xi}}{q} + \frac{C(\log N)^{2\xi}}{\sqrt{N}}.$$

Thus we have proved that with (ξ,ν)-high probability $|\Upsilon_i| \le C(\log N)^{\xi} q^{-1} + C(\log N)^{2\xi} N^{-1/2}$.

Next, using (2.13), we write the self-consistent equation (3.23) in the form

$$v_i = \frac{[v] - \Upsilon_i}{(z + m_{\mathrm{sc}} + [v] - \Upsilon_i)(z + m_{\mathrm{sc}})}. \tag{3.45}$$

The denominator of (3.45) is with (ξ,ν)-high probability larger in absolute value than

$$\big(2 - 1 - O\big((\log N)^{\xi} q^{-1} + (\log N)^{2\xi} N^{-1/2}\big)\big)2 \ge 3/2,$$

since $|z + m_{\mathrm{sc}}| = |m_{\mathrm{sc}}|^{-1} \ge 2$ and $|[v]| \le 1$ by (3.43). Thus,

$$|v_i| \le \frac{\Lambda_d + O((\log N)^{\xi} q^{-1} + (\log N)^{2\xi} N^{-1/2})}{3/2},$$

which yields, after taking the maximum over i,

$$\Lambda_d \le \frac{\Lambda_d + O((\log N)^{\xi} q^{-1} + (\log N)^{2\xi} N^{-1/2})}{3/2}.$$

This completes the estimate of Λ_d, and hence the proof. □

3.5. *Dichotomy argument for* Λ. The following dichotomy argument serves as the basis for the continuity argument of Section 3.6.

We introduce the control parameters

$$\alpha := \left|\frac{1 - m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}\right|, \qquad \beta := \frac{(\log N)^{\xi}}{\sqrt{q}} + \frac{(\log N)^{4\xi/3}}{(N\eta)^{1/3}}, \tag{3.46}$$

where $\alpha = \alpha(z)$ and $\beta = \beta(z)$ depend on the spectral parameter z. For any $z \in D_L$ we have the bound $\beta \le (\log N)^{-\xi}$.

From Lemma 3.2 it also follows that there is a constant constant $K \ge 1$, depending only on Σ, such that

$$\frac{1}{K}\sqrt{\kappa + \eta} \le \alpha(z) \le K\sqrt{\kappa + \eta} \tag{3.47}$$

for any $z \in D_L$.

We shall fix E and vary η from 2 down to $(\log N)^L N^{-1}$. Since $\sqrt{\kappa + \eta}$ is increasing and $\beta(E + \mathrm{i}\eta)$ decreasing in η, we find that, for any $U > 1$, the equation $\sqrt{\kappa + \eta} = 2U^2 K \beta(E + \mathrm{i}\eta)$ has a unique solution η, which we denote by

$\widetilde{\eta} = \widetilde{\eta}(U, E)$ (recall that $\kappa = ||E| - 2|$ is independent of η). Moreover, it is easy to see that for any fixed U we have

$$\widetilde{\eta} \ll 1. \tag{3.48}$$

LEMMA 3.17 (Dichotomy). *There exists a constant U_0 such that, for any fixed $U \geq U_0$, there exists a constant $C_1(U)$, depending only on U, such that the following estimates hold for any $z = E + \mathrm{i}\eta \in D_L$:*

$$\Lambda(z) \leq U\beta(z) \quad \textit{or} \quad \Lambda(z) \geq \frac{\alpha(z)}{U} \qquad \textit{if } \eta \geq \widetilde{\eta}(U, E), \tag{3.49}$$

$$\Lambda(z) \leq C_1(U)\beta(z) \qquad \textit{if } \eta < \widetilde{\eta}(U, E) \tag{3.50}$$

on $\Omega(z)$ with (ξ, ν)-high probability and for sufficiently large N.

PROOF. Fix $z = E + \mathrm{i}\eta \in D_L$. From (3.42) and Lemma 3.2 we find

$$\frac{1 - m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}[v] = [v]^2 + O\left(\frac{\Lambda^2}{\log N}\right) + O\left(\frac{(\log N)^\xi}{q} + \sqrt{\beta^3\Lambda + \beta^3\alpha}\right)$$

with (ξ, ν)-high probability. The third term on the right-hand side is bounded by $C^*(\beta\Lambda + \alpha\beta + \beta^2)$ for some constant $C^* \geq 1$. We set $U_0 := 9(C^* + 1)$. We conclude that in $\Omega(z)$ we have with (ξ, ν)-high probability

$$\left|\frac{1 - m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}[v] - [v]^2\right| \leq O\left(\frac{\Lambda^2}{\log N}\right) + C^*(\beta\Lambda + \alpha\beta + \beta^2). \tag{3.51}$$

Depending on the size of β relative to α, which is determined by z, we shall estimate either $[v]$ or $[v]^2$ using (3.51). This gives rise to the two cases in Lemma 3.17.

Case 1: $\eta \geq \widetilde{\eta}$. From the definition of $\widetilde{\eta}$ and C^* we find that

$$\beta \leq \frac{\alpha}{2U^2} \leq \frac{\alpha}{2C^*} \leq \alpha. \tag{3.52}$$

Recalling that $\Lambda = |[v]|$, we therefore obtain from (3.51) with (ξ, ν)-high probability

$$\alpha\Lambda \leq 2\Lambda^2 + C^*(\beta\Lambda + \alpha\beta + \beta^2) \leq 2\Lambda^2 + \frac{\alpha\Lambda}{2} + 2C^*\alpha\beta,$$

which gives

$$\alpha\Lambda \leq 4\Lambda^2 + 4C^*\alpha\beta.$$

Thus, either $\alpha\Lambda/2 \leq 4\Lambda^2$ which implies $\Lambda \geq \alpha/8 \geq \alpha/U$, or $\alpha\Lambda/2 \leq 4C^*\alpha\beta$ which implies $\Lambda \leq 8C^*\beta \leq U\beta$. This proves (3.49).

Case 2: $\eta < \widetilde{\eta}$. In this case the definition of $\widetilde{\eta}$ yields $\alpha \leq 2U^2K^2\beta$. We express $|[v]|^2 = \Lambda^2$ from (3.51) and we get

$$\Lambda^2 \leq 2\alpha\Lambda + 2C^*(\beta\Lambda + \alpha\beta + \beta^2) \leq C'\beta\Lambda + C'\beta^2 \tag{3.53}$$

for some constant C' depending on U. Now (3.50) is an immediate consequence. □

3.6. *Continuity argument*: *Conclusion of the proof of Theorem* 3.1. We complete the proof of Theorem 3.1 using a continuity argument in η to go from $\eta = 2$ down to $\eta = N^{-1}(\log N)^L$. We focus first on proving (3.5). We use Lemma 3.16 for the initial estimate, and the dichotomy in Lemma 3.17 to propagate a strong estimate on Λ to smaller values of η.

Choose a decreasing finite sequence η_k, $k = 1, 2, \ldots, k_0$, satisfying $k_0 \le CN^8$, $|\eta_k - \eta_{k+1}| \le N^{-8}$, $\eta_1 = 2$, and $\eta_{k_0} = N^{-1}(\log N)^L$. We fix $E \in [-\Sigma, \Sigma]$ and set $z_k := E + \mathrm{i}\eta_k$. Throughout this section we fix a $U \ge U_0$ in Lemma 3.17, and recall the definition of $\widetilde{\eta}(U, E)$ from Section 3.5.

Consider first z_1. It is easy to see that, for large enough N, we have $\eta_1 \ge \widetilde{\eta}(U, E)$, for any $E \in [-\Sigma, \Sigma]$. Therefore, Lemmas 3.16 and 3.17 imply that both $\Omega(z_1)$ and

$$\Lambda(z_1) \le U\beta(z_1)$$

hold with (ξ, ν)-high probability. This estimate takes care of the initial point η_1. The next lemma extends this result to all $k \le k_0$.

LEMMA 3.18. *Define the event*

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

where

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

Then

$$\mathbb{P}(\Omega_k^c) \le 2k\mathrm{e}^{-\nu(\log N)^\xi}. \tag{3.54}$$

PROOF. We proceed by induction on k. The case $k = 1$ was just proved. Let us therefore assume that (3.54) holds for k. We need to estimate

$$\begin{aligned} \mathbb{P}(\Omega_{k+1}^c) &\le \mathbb{P}\big(\Omega_k \cap \Omega(z_{k+1}) \cap \Omega_{k+1}^c\big) + \mathbb{P}\big(\Omega_k \cap (\Omega(z_{k+1}))^c\big) + \mathbb{P}(\Omega_k^c) \\ &= B + A + \mathbb{P}(\Omega_k^c), \end{aligned} \tag{3.55}$$

where we defined

$$\begin{aligned} A &:= \mathbb{P}[\Omega_k \cap \{\Lambda_d(z_{k+1}) + \Lambda_o(z_{k+1}) > (\log N)^{-\xi}\}], \\ B &:= \mathbb{P}\big[\Omega_k \cap \Omega(z_{k+1}) \cap \{\Lambda(z_{k+1}) > C^{(k+1)}(U)\beta(z_{k+1})\}\big]. \end{aligned}$$

We begin by estimating A. For any i, j, we have

$$\begin{aligned} |G_{ij}(z_{k+1}) - G_{ij}(z_k)| &\le |z_{k+1} - z_k| \sup_{z \in D_L} \left|\frac{\partial G_{ij}(z)}{\partial z}\right| \\ &\le N^{-8} \sup_{z \in D_L} \frac{1}{(\operatorname{Im} z)^2} \le N^{-6}. \end{aligned} \tag{3.56}$$

Therefore, by (3.32) and (3.39), we have on Ω_k with (ξ,ν)-high probability

$$\begin{aligned}\Lambda_d(z_{k+1}) + \Lambda_o(z_{k+1}) &\le \Lambda_d(z_k) + \Lambda_o(z_k) + 2N^{-6} \\ &\le C\left(\frac{(\log N)^\xi}{q} + (\log N)^{2\xi}\Psi(z_k)\right) + \Lambda(z_k) \\ &\le C\beta(z_k) \ll (\log N)^{-\xi}.\end{aligned}$$

Thus, we find that $A \le \mathrm{e}^{-\nu(\log N)^\xi}$.

Next, we estimate B. Suppose first that $\eta_k \ge \widetilde{\eta}(U,E)$. Then, similarly to (3.56), we find $|\Lambda(z_{k+1}) - \Lambda(z_k)| \le N^{-6}$. Thus, we find on Ω_k with (ξ,ν)-high probability

$$\Lambda(z_{k+1}) \le \Lambda(z_k) + N^{-6} \le U\beta(z_k) + N^{-6} \le \frac{3U}{2}\beta(z_{k+1}). \tag{3.57}$$

Suppose now that $\eta_{k+1} \ge \widetilde{\eta}(U,E)$. Then from (3.57) and (3.52) we find $\Lambda(z_{k+1}) < \frac{\alpha(z_{k+1})}{U}$. Now the dichotomy of (3.49) yields on $\Omega_k \cap \Omega(z_{k+1})$ with (ξ,ν)-high probability that $\Lambda(z_{k+1}) \le U\beta(z_{k+1})$. On the other hand, if $\eta_{k+1} < \widetilde{\eta}(U,E)$, then (3.57) immediately implies $\Lambda(z_{k+1}) \le C_1(U)\beta(z_{k+1})$. This concludes the proof of $B \le \mathrm{e}^{-\nu(\log N)^\xi}$ if $\eta_k \ge \widetilde{\eta}(U,E)$.

Finally, suppose that $\eta_k < \widetilde{\eta}(U,E)$. Thus, we also have $\eta_{k+1} < \widetilde{\eta}(U,E)$. In this case we immediately get from (3.50) on $\Omega(z_{k+1})$ with (ξ,ν)-high probability $\Lambda(z_{k+1}) \le C_1(U)\beta(x_{k+1})$.

We have therefore proved, for all k, that $\mathbb{P}(\Omega_{k+1}^c) \le 2\mathrm{e}^{-\nu(\log N)^\xi} + \mathbb{P}(\Omega_k^c)$, and the claim follows. □

In order to complete the proof of Theorem 3.1, we invoke the following simple lattice argument which strengthens the result of Lemma 3.18 to a statement uniform in $z \in D_L$. The main ingredient is the Lipschitz continuity of the map $z \mapsto G_{ij}(z)$, with a Lipschitz constant bounded by $\eta^{-2} \le N^2$.

COROLLARY 3.19. *There is a constant C such that*

$$\mathbb{P}\left[\bigcup_{z\in D_L}(\Omega(z))^c\right] + \mathbb{P}\left[\bigcup_{z\in D_L}\{\Lambda(z) > C\beta(z)\}\right] \le \mathrm{e}^{-\nu(\log N)^\xi}. \tag{3.58}$$

PROOF. Take a lattice $\mathcal{L} \subset D_L$ such that $|\mathcal{L}| \le CN^6$ and for any $z \in D_L$ there is a $\tilde{z} \in \mathcal{L}$ satisfying $|z - \tilde{z}| \le N^{-3}$. From the definition of G it is easy to see that for $z, \tilde{z} \in D_L$

$$|G_{ij}(z) - G_{ij}(\tilde{z})| \le \eta^{-2}|z - \tilde{z}| \le \frac{1}{N}. \tag{3.59}$$

The same bound holds for $|m(z) - m(\tilde{z})|$. Moreover, Lemma 3.18 immediately yields

$$\mathbb{P}\Bigg[\bigcap_{\tilde{z}\in\mathcal{L}}\bigg\{\Lambda(\tilde{z}) \le \frac{C}{2}\beta(\tilde{z})\bigg\}\Bigg] \ge 1 - \mathrm{e}^{-\nu(\log N)^{\xi}} \tag{3.60}$$

for some C large enough and some $\nu > 0$. From (3.59), (3.60) and $N^{-1} \ll \beta(z)$ we get

$$\mathbb{P}\Bigg[\bigcup_{z\in D_L}\{\Lambda(z) > C\beta(z)\}\Bigg] \le \mathrm{e}^{-\nu(\log N)^{\xi}}.$$

The first term of (3.58) is estimated similarly. □

We have proved (3.5). In order to prove (3.3), we note that (3.5), (3.32) and (3.58) imply

$$\Lambda_o(z) \le \frac{C}{q} + \frac{C(\log N)^{2\xi}}{\sqrt{N\eta}}$$

with (ξ, ν)-high probability. Now a lattice argument analogous to Corollary 3.19 yields (3.3). The diagonal estimate (3.4) follows similarly using (3.36). This concludes the proof of Theorem 3.1.

4. Proof of Theorem 2.8. In the previous section we proved Theorem 3.1, which is weaker than the main result Theorem 2.8 (strong local semicircle law), but will be used as an a priori bound in the proof of Theorem 2.8. The key ingredient that allows us to strengthen Theorem 3.1 to Theorem 2.8 is the following lemma, which shows that $[Z]$, the average of the Z_i's, is much smaller than that of a typical Z_i. (Notice that in the proof of Theorem 3.1, to arrive at (3.42), $[Z]$ was estimated by the same quantity as each individual Z_i.) This lemma is analogous to Lemma 5.2 in [21] and Corollary 4.2 in [23], but we will present a new proof (in Section 5.3), which admits sparse matrix entries and effectively tracks the dependence of the exponent p. Our new proof is based on an abstract decoupling result, Theorem 5.6 below, which is useful in other contexts as well, such as for proving Proposition 7.11 below.

LEMMA 4.1. *Recall the notation* $[Z] = \frac{1}{N}\sum_i Z_i$. *Suppose that* ξ *satisfies* (2.4), $q \ge (\log N)^{5\xi}$ *and that there exists* $\widetilde{D} \subset D_L$ *with* $L \ge 14\xi$ *such that we have with* (ξ, ν)*-high probability*

$$\Lambda(z) \le \gamma(z) \qquad \text{for } z \in \widetilde{D}, \tag{4.1}$$

where γ is a deterministic function satisfying $\gamma(z) \le (\log N)^{-\xi}$. Then we have with $(\xi - 2, \nu)$-high probability

$$|[Z](z)| \le (\log N)^{14\xi}\left(\frac{1}{q^2} + \frac{1}{(N\eta)^2} + (\log N)^{4\xi}\frac{\operatorname{Im} m_{\mathrm{sc}}(z) + \gamma(z)}{N\eta}\right) \quad \text{for } z \in \widetilde{D}. \tag{4.2}$$

In particular, by (3.40), we have with $(\xi - 2, \nu)$-high probability

$$\begin{aligned} &\left|\frac{1 - m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}[v] - [v]^2\right| \\ &\quad \le C\frac{\Lambda^2}{\log N} + C(\log N)^{14\xi}\left(\frac{1}{q^2} + \frac{1}{(N\eta)^2} + (\log N)^{4\xi}\frac{\operatorname{Im} m_{\mathrm{sc}} + \gamma}{N\eta}\right) \end{aligned} \tag{4.3}$$

for any value of the spectral parameter $z \in \widetilde{D}$.

The proof of Lemma 4.1 is given in Section 5. In this section we use it to prove Theorem 2.8 and to derive an estimate on $\|H\|$ (Lemma 4.4).

The basic idea behind the proof of Theorem 2.8 using Lemma 4.1 is to iterate (4.2) in order to obtain successively better estimates for Λ. Each step of the iteration improves the power $1 - \tau$ of the control parameter $(q^{-1} + (N\eta)^{-1})^{1-\tau}$. The iteration is started with the weak local semicircle law, Theorem 3.1, which yields $1 - \tau = 1/3$. At each step of the iteration, τ is halved at the expense of reducing the parameter ξ to $\xi - 2$, thus reducing the probability on which the estimate holds. This iteration procedure is repeated an order $\log\log N$ times, which allows us effectively to reach $\tau = 0$.

The iteration step is based on the following lemma, which is entirely deterministic.

LEMMA 4.2. *Let $1 \le \xi_1 \le \xi_2$ and $q > 1$. Let $0 < \tau < 1$ and $L > 1$. Suppose that there is a number $\gamma(z)$ satisfying*

$$\gamma(z) \le (\log N)^{19\xi_2}\left(\frac{1}{q} + \frac{1}{N\eta}\right)^{1-\tau} \qquad \text{for } z \in D_L \tag{4.4}$$

such that (4.1) holds with $\widetilde{D} := D_L$. We also assume that

$$\begin{aligned} &\left|\frac{1 - m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}[v] - [v]^2\right| \\ &\quad \le C\frac{\Lambda^2}{\log N} + C(\log N)^{14\xi_1}\left(\frac{1}{q^2} + \frac{1}{(N\eta)^2} + (\log N)^{4\xi_1}\frac{\alpha + \gamma}{N\eta}\right) \\ &\qquad \text{for } z \in D_L, \end{aligned} \tag{4.5}$$

where α was defined in (3.46). *Finally, we assume that if $\eta \sim 1$, then*

$$\Lambda(z) \ll 1. \tag{4.6}$$

Then we have

$$\Lambda(z) \le (\log N)^{19\xi_2} \left(\frac{1}{q} + \frac{1}{N\eta}\right)^{1-\tau/2} \tag{4.7}$$

for $z \in D_L$ and large enough N.

PROOF. The proof is based on a dichotomy argument. Define

$$\alpha_0(z) := (\log N)^{(18+3/4)\xi_2} \left(\frac{1}{q} + \frac{1}{N\eta}\right)^{1-\tau/2}. \tag{4.8}$$

We consider two cases.

Case 1: $\alpha \le 10\alpha_0$. Using the estimate (4.4), we find

$$\frac{1}{q^2} + \frac{1}{(N\eta)^2} + (\log N)^{4\xi_1} \frac{\gamma}{N\eta} \le 2(\log N)^{23\xi_2} \left(\frac{1}{q} + \frac{1}{N\eta}\right)^{2-\tau}. \tag{4.9}$$

Now in (4.5) we may absorb the term $\Lambda^2/\log N$ into the term $|[v]|^2$ on the left-hand side, at the expense of a constant 2. Then we complete the square on the left-hand side and take the square root of the resulting equation; this yields

$$\Lambda \le 4\alpha + C(\log N)^{37\xi_2/2} \left(\frac{1}{q} + \frac{1}{N\eta}\right)^{1-\tau/2} + C(\log N)^{9\xi_1} \sqrt{\frac{\alpha}{N\eta}}, \tag{4.10}$$

where we used (4.9). Now (4.7) follows from (4.10).

Case 2: $\alpha \ge 10\alpha_0$. Let us assume that $\Lambda \le \alpha/2$. Then in (4.5) the terms $[v]^2$ and Λ^2 can be absorbed into the term $\alpha|[v]|$, so that we get

$$\begin{aligned} \Lambda \le{}& C\frac{(\log N)^{14\xi_1}}{\alpha} \left(\frac{1}{q} + \frac{1}{N\eta}\right)^2 + C(\log N)^{18\xi_1} \frac{\gamma}{N\eta\alpha} \\ &+ C(\log N)^{18\xi_1} \frac{1}{N\eta}. \end{aligned} \tag{4.11}$$

By the definitions of γ and α_0, we have

$$C(\log N)^{18\xi_1} \frac{\gamma}{N\eta\alpha} \le \frac{\alpha_0}{(\log N)^{1/4}} \le \frac{10\alpha}{(\log N)^{1/4}} \tag{4.12}$$

and the first term in the right-hand side of (4.11) is bounded by $\alpha/\log N$ thanks to (4.8). The last term can be estimated similarly. Hence, (4.11) implies that $\Lambda \le \alpha/4$ provided that $\Lambda \le \alpha/2$.

In other words, if $\alpha \ge 10\alpha_0$, then either $\Lambda > \alpha/2$ or $\Lambda \le \alpha/4$. Using the continuity of $\Lambda(z)$ and $\alpha = \alpha(z)$ in $\eta = \operatorname{Im} z$, and the assumption

$$\Lambda(z) \ll 1 = O(\alpha)$$

for $\eta \sim 1$, we get $\Lambda \le \alpha/4$ on the whole D_L. Together with (4.11), we obtain (4.7). □

PROOF OF THEOREM 2.8. The main work is to prove Theorem 2.8 for spectral parameters $z \in D_L$, where

$$L := 120\xi. \tag{4.13}$$

Once this is done, the extension to all $z \in D$ is relatively straightforward, and is given at the end of the proof. Recall the definitions (3.26) of Ψ and (3.25) of $\Omega(z)$. It is clear that if D is replaced everywhere by D_L, then (2.17) follows from (2.16), (3.32) and (3.36). Therefore, we only need to prove (2.16).

We begin by introducing

$$\widetilde{\xi} := 2(\log\log N/\log 2) + \xi. \tag{4.14}$$

By the assumptions (2.15) and (4.13), we have $\widetilde{\xi} \le 3\xi/2 \le A_0 \log\log N$, $L \ge 60\widetilde{\xi}$, and $q \ge (\log N)^{60\widetilde{\xi}}$. To prove (2.16) with D replaced by D_L, it therefore suffices to establish

$$\bigcap_{z\in D_L}\left\{|m(z) - m_{sc}(z)| \le (\log N)^{20\widetilde{\xi}}\left(\min\left\{\frac{(\log N)^{20\widetilde{\xi}}}{\sqrt{\kappa_E+\eta}}\frac{1}{q^2}, \frac{1}{q}\right\} + \frac{1}{N\eta}\right)\right\} \tag{4.15}$$

with (ξ, ν)-high probability.

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

$$\begin{aligned}\Lambda &\le (\log N)^{2\widetilde{\xi}}\left(\frac{1}{q} + \frac{1}{N\eta}\right)^{1/3} \\ &\le (\log N)^{19\widetilde{\xi}}\left(\frac{1}{q} + \frac{1}{N\eta}\right)^{1-2/3} \qquad \text{for } z \in D_L\end{aligned} \tag{4.16}$$

with $(\widetilde{\xi}, \nu)$-high probability. Thus, (4.1) holds with

$$\gamma(z) := (\log N)^{19\widetilde{\xi}}\left(\frac{1}{q} + \frac{1}{N\eta}\right)^{1/3}. \tag{4.17}$$

With $L \ge 60\widetilde{\xi}$ and $q \ge (\log N)^{60\widetilde{\xi}}$, we also have $\gamma \le (\log N)^{-\widetilde{\xi}}$. Thus, Lemma 4.1 implies that, with $\widetilde{\xi}$ replacing ξ and $\widetilde{D} = D_L$, the statement

$$\begin{aligned}&\left|\frac{1-m_{sc}^2}{m_{sc}^3}[v] - [v]^2\right| \\ &\quad \le C\frac{\Lambda^2}{\log N} + C(\log N)^{14\widetilde{\xi}}\left(\frac{1}{q^2} + \frac{1}{(N\eta)^2} + (\log N)^{4\widetilde{\xi}}\frac{\operatorname{Im} m_{sc} + \gamma}{N\eta}\right) \\ &\qquad \text{for } z \in D_L\end{aligned} \tag{4.18}$$

holds with $(\widetilde{\xi}-2,\nu)$-high probability. This implies (4.5), with the choice $\xi_1=\widetilde{\xi}$ in (4.5), since $\operatorname{Im} m_{\mathrm{sc}}\le C\alpha$. Moreover, γ satisfies (4.4) with $\xi_2=\widetilde{\xi}$ and $\tau=2/3$. We also find that Λ satisfies (4.6), since $\Lambda\le\gamma\le(\log N)^{-\widetilde{\xi}}$ [see (4.16)]. We may therefore apply Lemma 4.2 with $\xi_1=\xi_2=\widetilde{\xi}$ to get that

$$\Lambda\le(\log N)^{19\widetilde{\xi}}\left(\frac{1}{q}+\frac{1}{N\eta}\right)^{1-1/3}\qquad\text{for } z\in D_L \tag{4.19}$$

holds with $(\widetilde{\xi}-2,\nu)$-high probability. We now repeat this process M times, each iteration yielding a stronger bound on Λ which holds with a smaller probability. After M iterations we get that

$$\Lambda\le(\log N)^{19\widetilde{\xi}}\left(\frac{1}{q}+\frac{1}{N\eta}\right)^{1-2(1/2)^M/3}\qquad\text{for } z\in D_L \tag{4.20}$$

holds with $(\widetilde{\xi}-2M,\nu)$-high probability.

To clarify the iteration, we spell out the details of the second step. We start from (4.19) and define γ as the right-hand side of (4.19),

$$\gamma(z):=(\log N)^{19\widetilde{\xi}}\left(\frac{1}{q}+\frac{1}{N\eta}\right)^{1-1/3}. \tag{4.21}$$

Thus, Lemma 4.1, with $\widetilde{\xi}-2$ replacing ξ, implies that

$$\begin{aligned}\left|\frac{1-m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}[v]-[v]^2\right| &\le C\frac{\Lambda^2}{\log N}\\ &\quad+C(\log N)^{14(\widetilde{\xi}-2)}\left(\frac{1}{q^2}+\frac{1}{(N\eta)^2}+(\log N)^{4(\widetilde{\xi}-2)}\frac{\operatorname{Im} m_{\mathrm{sc}}+\gamma}{N\eta}\right)\\ &\qquad\text{for } z\in D_L\end{aligned} \tag{4.22}$$

holds with $(\widetilde{\xi}-4,\nu)$-high probability. We now apply Lemma 4.2 with $\xi_1=\widetilde{\xi}-2$, $\xi_2=\widetilde{\xi}$ and $\tau=1/3$. [Similarly, in the kth step we set $\xi_1=\widetilde{\xi}-2(k-1)$, $\xi_2=\widetilde{\xi}$, and $\tau=(2/3)(1/2)^{k-1}$.] This shows that

$$\Lambda\le(\log N)^{19\widetilde{\xi}}\left(\frac{1}{q}+\frac{1}{N\eta}\right)^{1-1/6}\qquad\text{for } z\in D_L \tag{4.23}$$

holds with $(\widetilde{\xi}-4,\nu)$-high probability. This is (4.20) for $M=2$.

Now we return to (4.20) and choose $M:=[\log\log N/\log 2]-1$ (where $[\cdot]$ denotes the integer part). Using $q^{-1}+(N\eta)^{-1}\ge cN^{-1/2}$ [by (2.6)], we get

$$\left(\frac{1}{q}+\frac{1}{N\eta}\right)^{-2(1/2)^M/3}\le C\le(\log N)^{\widetilde{\xi}/2}.$$

Thus,

$$\Lambda \le (\log N)^{39\widetilde{\xi}/2}\left(\frac{1}{q}+\frac{1}{N\eta}\right) \qquad \text{for } z \in D_L \tag{4.24}$$

holds with $(\xi+2,\nu)$-high probability. Recalling (3.47), we find that (4.24) implies (4.15), unless

$$(\log N)^{-39\widetilde{\xi}/2}\alpha \ge \frac{1}{q} \ge \frac{1}{N\eta}. \tag{4.25}$$

Let us therefore assume that (4.25) holds. Then it remains to prove that with (ξ,ν)-high probability

$$\Lambda \le (\log N)^{40\widetilde{\xi}}\frac{1}{\alpha q^2}+(\log N)^{20\widetilde{\xi}}\frac{1}{N\eta} \qquad \text{for } z \in D_L. \tag{4.26}$$

Defining γ as the right-hand side of (4.24), we use Lemma 4.1, with $\xi+2$ replacing ξ, to get

$$\left|\frac{1-m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}[v]-[v]^2\right| \le C\frac{\Lambda^2}{\log N}+C(\log N)^{18\widetilde{\xi}}\left(\frac{1}{q^2}+\frac{\alpha}{N\eta}\right) \tag{4.27}$$

with (ξ,ν)-high probability, where we used (4.25) and $|\operatorname{Im} m_{\mathrm{sc}}(z)| \le C\alpha(z)$. We can estimate the term $[v]^2$ by (4.24) and (4.25), so that

$$\alpha\Lambda = \alpha|[v]| \le C(\log N)^{39\widetilde{\xi}}\frac{1}{q^2}+C(\log N)^{18\widetilde{\xi}}\frac{\alpha}{N\eta}. \tag{4.28}$$

This yields (4.26) and hence completes the proof of (2.16) with D replaced with D_L. (Recall the simple lattice argument of Corollary 3.19.)

What remains is to extend (2.16) and (2.17) from $z \in D_L$ to $z \in D$. Let us therefore assume that $z = E+\mathrm{i}\eta \in D$ with $0<\eta\le\widetilde{\eta} := (\log N)^L N^{-1}$. For any $i,j=1,\dots,N$ we get the bound

$$|G_{ij}(E+\mathrm{i}\eta)| = \left|\sum_\alpha \frac{\bar{u}_\alpha(i)u_\alpha(j)}{\lambda_\alpha - z}\right| \le \max_l \sum_\alpha \frac{|u_\alpha(l)|^2}{|\lambda_\alpha - z|}.$$

We define the dyadic decomposition of the eigenvalues

$$U_0 := \{\alpha : |\lambda_\alpha - E| < \eta\}, \qquad U_k := \{\alpha : 2^{k-1}\eta \le |\lambda_\alpha - E| < 2^k\eta\} \qquad (k\ge 1).$$

This yields

$$\begin{aligned}\sum_\alpha \frac{|u_\alpha(l)|^2}{|\lambda_\alpha - z|} &= \sum_{k\ge 0}\sum_{\alpha\in U_k}\frac{|u_\alpha(l)|^2}{|\lambda_\alpha - z|} \le C\sum_{k\ge 0}\sum_{\alpha\in U_k}\operatorname{Im}\frac{|u_\alpha(l)|^2}{\lambda_\alpha - E - \mathrm{i}2^k\eta}\\ &\le C\sum_{k\ge 0}\operatorname{Im} G_{ll}(E+\mathrm{i}2^k\eta).\end{aligned}$$

Next, we break the summation over k into three pieces delimited by $k_1 := \max\{k : 2^k\eta < \widetilde{\eta}\}$ and $k_2 := \max\{k : 2^k\eta < 3\}$. By spectral decomposition, it is easy to see that the function $y \mapsto y \operatorname{Im} G_{ll}(E + \mathrm{i}y)$ is monotone increasing. Therefore, we get

$$\begin{aligned}
\sum_{k \geq 0} \operatorname{Im} G_{ll}(E + \mathrm{i}2^k\eta) &\leq \sum_{k=0}^{k_1} \frac{\widetilde{\eta}}{2^k\eta} \operatorname{Im} G_{ll}(E + \mathrm{i}\widetilde{\eta}) \\
&\quad + \sum_{k=k_1+1}^{k_2} \operatorname{Im} G_{ll}(E + \mathrm{i}2^k\eta) + \sum_{k=k_2+1}^{\infty} \frac{1}{\eta 2^k} \\
&\leq \frac{(\log N)^{C\xi}}{N\eta} + C(k_2 - k_1) + C \\
&\leq \frac{(\log N)^{C\xi}}{N\eta}
\end{aligned}$$

with (ξ, ν)-high probability, where in the second step we used (2.17) for $z \in D_L$. Therefore, we have proved that

$$\max_{i,j} |G_{ij}(E + \mathrm{i}\eta)| \leq \frac{(\log N)^{C\xi}}{N\eta}$$

with (ξ, ν)-high probability. This concludes the proof of Theorem 2.8. □

4.1. *Estimate of* $\|H\|$. In this section we derive an upper bound on the norm of H. A standard application of the moment method yields the following weak bound on $\|H\|$. Its proof is given in the Appendix.

LEMMA 4.3. *Suppose that H satisfies Definition 2.1, that ξ satisfies (2.4) and that q satisfies (2.6). Then with (ξ, ν)-high probability we have*

$$\|H\| \leq 2 + (\log N)^{\xi} q^{-1/2}. \tag{4.29}$$

Using the local semicircle law, Theorem 2.8, we may prove a much stronger bound on $\|H\|$. Lemma 4.3 will be used as an a priori bound in the proof of Lemma 4.4.

LEMMA 4.4. *Suppose that H satisfies Definition 2.1, and that ξ and q satisfy (2.15). Then with (ξ, ν)-high probability we have*

$$\|H\| \leq 2 + (\log N)^{C\xi}(q^{-2} + N^{-2/3}). \tag{4.30}$$

PROOF. We only consider the largest eigenvalue $\lambda_N = \max_\alpha \lambda_\alpha$; the smallest eigenvalue λ_1 is handled similarly. Set $L = 120\xi$. Using (2.16) with $\xi + 2$ replacing ξ, we get with $(\xi + 2, \nu)$-high probability

$$\Lambda(z) \le (\log N)^{41\xi}\left(\frac{1}{q} + \frac{1}{N\eta}\right). \tag{4.31}$$

Then applying Lemma 4.1 with

$$\gamma(z) := (\log N)^{41\xi}\left(\frac{1}{q} + \frac{1}{N\eta}\right) \tag{4.32}$$

and $\xi + 2$ replacing ξ, we have with (ξ, ν)-high probability

$$\begin{aligned} &\left|\frac{1 - m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}[v] - [v]^2\right| \\ &\quad \le C\frac{\Lambda^2}{\log N} + C(\log N)^{C_1\xi}\left(\frac{1}{q^2} + \frac{1}{(N\eta)^2} + \frac{\operatorname{Im} m_{\mathrm{sc}}}{N\eta}\right) \\ &\qquad \text{for } z \in D_L, \end{aligned} \tag{4.33}$$

where C_1 is a sufficiently large constant. Now if $E > 2$ and $\kappa \ge \eta$, then Lemma 3.2 and (3.47) yield

$$\operatorname{Im} m_{\mathrm{sc}} \sim \frac{\eta}{\sqrt{\kappa}}, \qquad \alpha \sim \sqrt{\kappa}. \tag{4.34}$$

Inserting (4.34) into (4.33), we find with (ξ, ν)-high probability

$$\left|\frac{1 - m_{\mathrm{sc}}^2}{m_{\mathrm{sc}}^3}[v] - [v]^2\right| \le C\frac{\Lambda^2}{\log N} + C(\log N)^{C_1\xi}\left(\frac{1}{q^2} + \frac{1}{(N\eta)^2} + \frac{1}{N\sqrt{\kappa}}\right). \tag{4.35}$$

Next, for any fixed $C_1 > 0$, we can find a large enough constant $C_2 > 2C_1$ such that if E satisfies

$$2 + (\log N)^{C_2\xi}(q^{-2} + N^{-2/3}) \le E \le 3, \tag{4.36}$$

then

$$\min\{N^{-1/2}\kappa^{1/4}, N^{-1}\kappa^{1/2}q^2, \kappa\} \ge (\log N)^{C_1\xi+2}N^{-1}\kappa^{-1/2}. \tag{4.37}$$

(Here $\kappa = \kappa_E = E - 2$.) From now on we assume that E satisfies (4.36). We define

$$\eta = \eta_E := (\log N)^{C_1\xi+1}N^{-1}\kappa^{-1/2}. \tag{4.38}$$

Note that η depends on E via κ. From (4.37) we have

$$\kappa \ge \eta. \tag{4.39}$$

Using (4.37), (4.38) and (4.34), we get

$$\frac{1}{N\eta} \gg \frac{\eta}{\sqrt{\kappa}} \sim \operatorname{Im} m_{\mathrm{sc}}(E + i\eta). \tag{4.40}$$

Similarly, using (4.37), we have

$$\frac{1}{N\eta} \geq \frac{1}{q^2\sqrt{\kappa}}. \tag{4.41}$$

Next, with the lower bound $\alpha \geq \sqrt{\kappa}/K$ from (3.47) and (4.39), we find, using (4.31), that

$$\alpha \geq c(\log N)^{C_1\xi+1}\left(\frac{1}{q} + \frac{1}{N\eta}\right) \gg \Lambda \tag{4.42}$$

with (ξ, ν)-high probability, where we used (4.36) to obtain the first term q^{-1} on the right-hand side and we used $N\eta\sqrt{\kappa} = (\log N)^{C_1\xi+1}$ [see the definition (4.38) of η] for the second term. Now we can assume

$$q \geq (\log N)^{C_3\xi} \tag{4.43}$$

for some large $C_3 > 0$ [otherwise (4.30) holds for some constant C by Lemma 4.3]. We have $E + \mathrm{i}\eta \in D_L$ [recall that E satisfies (4.36)]. Using (4.42), we can neglect the terms Λ^2 and $[v]^2$ in (4.35) to get, with (ξ, ν)-high probability,

$$\Lambda \leq C(\log N)^{C_1\xi}\left(\frac{1}{\alpha q^2} + \frac{1}{\alpha(N\eta)^2} + \frac{1}{\alpha N\sqrt{\kappa}}\right). \tag{4.44}$$

Since $\alpha \geq K\sqrt{\kappa}$, the last term is bounded by

$$\frac{1}{\alpha N\sqrt{\kappa}} \leq (\log N)^{-C_1\xi-1}\frac{\eta}{\sqrt{\kappa}} \leq (\log N)^{-C_1\xi-1}\frac{1}{N\eta},$$

where we have used (4.40). The first term on the right-hand side of (4.44) can be estimated similarly using (4.41) and (4.37). Finally, the middle term on the right-hand side of (4.44) can be estimated by using (4.42). Putting everything together, we obtain, for any E satisfying (4.36), that

$$\Lambda(z) \ll \frac{1}{N\eta} \qquad \text{for } z = E + \mathrm{i}\eta \in D_L \tag{4.45}$$

with (ξ, ν)-high probability. Furthermore, with (4.34) and (4.40), we obtain that for any E in (4.36)

$$\operatorname{Im} m(z) \leq \operatorname{Im} m_{\mathrm{sc}}(z) + \Lambda(z) \ll \frac{1}{N\eta} \qquad \text{for } z = E + \mathrm{i}\eta \in D_L \tag{4.46}$$

with (ξ, ν)-high probability. Since

$$\operatorname{Im} m(z) = \frac{1}{N}\sum_\alpha \frac{\eta}{(\lambda_\alpha - E)^2 + \eta^2}, \tag{4.47}$$

we have

$$\operatorname{Im} m(z) \geq \frac{c}{N\eta}$$

if there is an eigenvalue in $[E-\eta, E+\eta]$. Then (4.47) and (4.46) imply that, for any E satisfying (4.36), there is no eigenvalue in $[E-\eta, E+\eta]$ with (ξ, ν)-high probability. The regime $E \geq 3$ is covered by Lemma 4.3. This completes the proof. □

5. Abstract decoupling lemma and applications. In this section we prove an abstract decoupling lemma which is independent of the random matrix model. We shall apply this abstract result to random matrices in Sections 5.2, 5.3 and 7.4.

5.1. *Abstract decoupling lemma.* Throughout this section we use the letters A and B to denote abstract random variables. Note that A in this context has nothing to do with the matrix A from Definition 2.2. We work on the probability space generated by the $N \times N$ random matrices H. Let $(A^{[\mathbb{U}]})$ be a family of random variables indexed by subsets $\mathbb{U} \subset \{1, \dots, N\}$, and denote $A := A^{[\varnothing]}$. For $\mathbb{U} \subset \mathbb{S} \subset \{1, \dots, N\}$ we define the random variable

$$A^{\mathbb{S},\mathbb{U}} := \sum_{\mathbb{T}\subset\mathbb{U}} (-1)^{|\mathbb{T}|} A^{[(\mathbb{S}\setminus\mathbb{U})\cup\mathbb{T}]} = (-1)^{|\mathbb{S}\setminus\mathbb{U}|} \sum_{\mathbb{V}:\, \mathbb{S}\setminus\mathbb{U}\subset\mathbb{V}\subset\mathbb{S}} (-1)^{|\mathbb{V}|} A^{[\mathbb{V}]}. \tag{5.1}$$

LEMMA 5.1 (Resolution of dependence). *For any $\mathbb{S}$ we have*

$$A = \sum_{\mathbb{U}\subset\mathbb{S}} A^{\mathbb{S},\mathbb{U}}. \tag{5.2}$$

PROOF. The proof is a standard inclusion-exclusion argument. □

DEFINITION 5.2. Let $A := A(H)$ be a random variable. Then we define the new random variable $A^{(\mathbb{T})}$ through

$$A^{(\mathbb{T})}(H) := A(\pi_{\mathbb{T}}(H)), \tag{5.3}$$

where $\pi_{\mathbb{T}}$ was defined in (3.10).

REMARK 5.3. Note that the operation $(\cdot)^{(\mathbb{T})}$ is compatible with algebraic operations in the sense that

$$(A+B)^{(\mathbb{T})} = A^{(\mathbb{T})} + B^{(\mathbb{T})}, \qquad (AB)^{(\mathbb{T})} = A^{(\mathbb{T})} B^{(\mathbb{T})}. \tag{5.4}$$

Since $\pi_{\mathbb{U}} \circ \pi_{\mathbb{V}} = \pi_{\mathbb{U}\cup\mathbb{V}}$, we also have $(A^{(\mathbb{U})})^{(\mathbb{V})} = A^{(\mathbb{U}\cup\mathbb{V})}$.

REMARK 5.4. The matrices $H^{(\mathbb{T})}$ and $G^{(\mathbb{T})}$ defined through (5.3) are $N \times N$ matrices. We adopt this convention only in this section. This is in contrast to Definition 3.3, where the same notation was used for the $(N - |\mathbb{T}|) \times (N - |\mathbb{T}|)$ minors of the same matrices. This slight abuse of notation will not cause ambiguity, however, because we shall only consider matrix elements $H_{ij}^{(\mathbb{T})}$ and $G_{ij}^{(\mathbb{T})}$ for $i, j \notin \mathbb{T}$; for these matrix elements the two definitions coincide.

DEFINITION 5.5. We say that a random variable A is *independent* of the set $\mathbb{U} \subset \{1, \dots, N\}$ if $A = A^{(\mathbb{U})}$ [or, equivalently, if A is independent of the family $(h_{ij} : i \in \mathbb{U} \text{ or } j \in \mathbb{U})$].

We shortly explain the idea behind these definitions. In many applications we choose $A^{[\mathbb{U}]} := A^{(\mathbb{U})}$, so that $A^{[\mathbb{U}]}$ is independent of $\mathbb{U}$. In this case, the decomposition (5.2) can be interpreted as follows. We first fix a reference set $\mathbb{S}$. From (5.1) it is clear that $A^{\mathbb{S},\mathbb{U}}$ is independent of $\mathbb{S} \setminus \mathbb{U}$, that is, it depends only on the set $\mathbb{U}$ (among the variables in $\mathbb{S}$). Therefore, (5.2) can be viewed as a resolution of dependence of A on subsets of $\mathbb{S}$. We shall see that when we apply this decomposition to resolvent matrix elements, that is, set $A = G_{ij}$, then $G_{ij}^{\mathbb{S},\mathbb{U}}$ will be comparable in size with a product of at least $|\mathbb{U}| + 1$ off-diagonal resolvent matrix elements, which are small with high probability. Hence, in this case, the decomposition (5.2) is effectively a graded resolution with a trade-off between dependence and size. A larger $\mathbb{U}$ means that $G_{ij}^{\mathbb{S},\mathbb{U}}$ is smaller, but it depends on more variables. For smaller $\mathbb{U}$'s we will exploit that $G^{\mathbb{S},\mathbb{U}}$ is independent of more variables.

The purpose of this graded decoupling is to obtain large deviation estimates on the average $[\mathcal{Z}] := \frac{1}{N} \sum_i \mathcal{Z}_i$ of N weakly dependent centered random variables $\mathcal{Z}_i$. The precise result is given in Theorem 5.6 below. Before stating it, we outline the main ideas.

In our applications, the covariances between different variables Z_i are too large to be controlled in terms of their variances and, hence, standard methods for sums of weakly dependent random variables relying on such ideas do not apply. Instead, the weak dependence will be expressed in terms of the smallness of $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ for large $\mathbb{U}$; the size of $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ reflects how strongly $\mathcal{Z}_i$ depends on the set $\mathbb{U}$. The basic strategy is a high-moment estimate

$$\mathbb{E}\mathbf{1}(\Xi)|[\mathcal{Z}]|^p = \frac{1}{N^p} \sum_{i_1,\dots,i_p} \mathbb{E}(\mathbf{1}(\Xi)\mathcal{Z}_{i_1} \cdots \bar{\mathcal{Z}}_{i_p})$$

on some high-probability event Ξ, whereby each term $\mathcal{Z}_{i_j}$ is expanded according to the graded expansion of (5.2). The right-hand side is controlled using the two following facts: (i) $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ is small for large $\mathbb{U}$ (weak dependence of $\mathcal{Z}_i$ on $\mathbb{U}$). (ii) The expectation vanishes if all factors are independent. Note that this graded expansion differs from the conventional martingale-type arguments used to establish central limit theorems for correlated random variables.

The basic idea of a graded expansion to control large deviations of sums of weakly dependent random variables was introduced in Lemma 5.2 of [21] in the context of Wigner matrices. This result considers the special case $\mathcal{Z}_i = Z_i$ [as defined in (3.15)] and uses expansions in full rows and columns to detect dependencies. For the applications in [21], only large but N-independent powers p were considered. Hence, in [21] it was not necessary to keep track of the p-dependence or the probability of Ξ.

A new proof was given in Lemma 4.1 of [23], where the p-dependence and the probability of Ξ were tracked precisely. This proof relied on an expansion in terms of individual matrix elements and not full rows and columns. Thus, the expansion was more economical, but its combinatorial structure was considerably more involved.

In this paper we present an abstract generalization of the row and column expansion method of [21]. It is formulated for an arbitrary family of random variables $\mathcal{Z}_1, \dots, \mathcal{Z}_N$. As input, it needs bounds on the terms of the graded expansion of $\mathcal{Z}_i$. The abstract formulation thus streamlines the argument by dissociating two unrelated steps of the proof: (i) the moment estimate using the graded expansion (a probabilistic estimate given in Theorem 5.6) and (ii) controlling the size of the graded terms for a concrete application (in the case of resolvent matrix elements, a deterministic, almost entirely algebraic, argument given in Section 5.2).

For our purposes, this increased generality is needed for two reasons. First, it allows for an efficient control of the strong fluctuations associated with sparse matrix entries. Second, we use it to control the average of not only Z_i (Lemma 5.13) but also quantities like (7.27) with a different algebraic structure. In the special case $\mathcal{Z}_i = Z_i$ and $q = N^{1/2}$ (Wigner matrix), our result reduces to that of Lemma 4.1 in [23].

THEOREM 5.6 (Abstract decoupling lemma). *Let $\mathcal{Z}_1, \dots, \mathcal{Z}_N$ be random variables and recall the notation*

$$[\mathcal{Z}] = \frac{1}{N} \sum_{i=1}^{N} \mathcal{Z}_i .$$

Let Ξ be an event and p an even integer. Suppose that there exists a family of random variables $(\mathcal{Z}_i^{[\mathbb{U}]})_{i,\mathbb{U}}$ indexed by $i \in \{1, \dots, N\}$ and $\mathbb{U} \subset \{1, \dots, N\}$ satisfying $i \notin \mathbb{U}$, such that $\mathcal{Z}_i^{[\varnothing]} = \mathcal{Z}_i$ and the following assumptions hold with some constant C:

(i) *Recall the partial expectation $\mathbb{E}_i$ from Definition 3.6. For $i \notin \mathbb{U}$ we have that $\mathcal{Z}_i^{[\mathbb{U}]}$ is independent of $\mathbb{U}$ and*

$$\mathbb{E}_i \mathcal{Z}_i^{[\mathbb{U}]} = 0. \tag{5.5}$$

(ii) (L^r*-norm in* Ξ). *For any $\mathbb{U}$, $\mathbb{S}$ with $\mathbb{U} \subset \mathbb{S}$ and $i \notin \mathbb{S}$ we consider $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ defined by (5.1) from the family $\mathcal{Z}_i^{[\mathbb{U}]}$. Then for any numbers $r \leq p$ with $|\mathbb{S}| \leq p$ we have*

$$\mathbb{E}(\mathbf{1}(\Xi) |\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}|^r) \leq (Y(CXu)^u)^r \qquad \text{with } u := |\mathbb{U}| + 1, \tag{5.6}$$

where X and Y are deterministic and X satisfies

$$X \le \frac{1}{p^5 \log N}. \tag{5.7}$$

(iii) (*rough bound on the L^2-norm in $[\Xi]_i$*). *Define*

$$[\Xi]_i := (\pi_i^{-1} \circ \pi_i)(\Xi). \tag{5.8}$$

For any $\mathbb{U}$, $\mathbb{S}$ satisfying $\mathbb{U} \subset \mathbb{S}$, $i \notin \mathbb{S}$, and $|\mathbb{S}| \le p$ we have

$$\mathbb{E}(\mathbf{1}([\Xi]_i)|\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}|^2) \le N^{Cp}. \tag{5.9}$$

(iv) (*rough bound on $\mathcal{Z}_i$*). *For any $\mathbb{U}$ we have*

$$\mathbf{1}(\Xi)\big|\mathcal{Z}_i^{[\mathbb{U}]}\big| \le Y N^C. \tag{5.10}$$

(v) (Ξ *has high probability*). *We require that*

$$\mathbb{P}[\Xi^c] \le \mathrm{e}^{-c(\log N)^{3/2} p}. \tag{5.11}$$

Then, *under the assumptions* (i)–(v), *we have*

$$\mathbb{P}\big(\mathbf{1}(\Xi)|[\mathcal{Z}]| \ge p^{12} Y(X^2 + N^{-1})\big) \le \frac{C^p}{p^p} \tag{5.12}$$

for some $C > 0$ and sufficiently large N. The constant in (5.12) *depends on the constants in* (5.6), (5.10) *and* (5.11).

The key assumptions in Theorem 5.6 are (i) and (ii); the key (small) parameter is X. Assumption (i) simply ensures that all terms of the graded expansion of $\mathcal{Z}_i$ have zero expectation. Assumption (ii) defines the decay of $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ in the size of $\mathbb{U}$; roughly, it states that

$$|\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}| \lesssim X^{|\mathbb{U}|+1}$$

in the sense of high moments. This is in accordance with the principle outlined above that terms of the graded expansion which depend on many variables have a small size, while those which are independent of many variables may be larger. The parameter Y is trivial in our applications, where we shall take it to be a logarithmic factor. In Lemma 4.1 of [23], the role of X was played by the parameter Ψ defined in (3.26).

PROOF OF THEOREM 5.6. We find

$$\mathbb{E}(\mathbf{1}(\Xi)|\mathcal{Z}|^p) = N^{-p} \sum_{\alpha_1,\alpha_2,\dots,\alpha_p=1}^{N} \mathbb{E}\Bigg(\mathbf{1}(\Xi) \prod_{j=1}^{p} \mathcal{Z}_{\alpha_j}^{\#}\Bigg), \tag{5.13}$$

where # stands for either nothing or complex conjugation. Let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)$ and define $\mathbb{S} \equiv \mathbb{S}(\boldsymbol{\alpha}) := \{\alpha_1, \alpha_2, \dots, \alpha_p\}$. Then we have

$$\mathbb{E}(\mathbf{1}(\Xi)|\mathcal{Z}|^p) \le N^{-p} p^p \sum_{s=1}^{p} N^s \max_{\boldsymbol{\alpha} : |\mathbb{S}(\boldsymbol{\alpha})|=s} \left| \mathbb{E}\left(\mathbf{1}(\Xi) \prod_{j=1}^{p} \mathcal{Z}_{\alpha_j}^{\#}\right)\right|. \tag{5.14}$$

Abbreviating $\mathbb{S}_j := \mathbb{S} \setminus \{\alpha_j\}$, we find from (5.2) that

$$\mathcal{Z}_{\alpha_j} = \sum_{\mathbb{U}'_j \subset \mathbb{S}_j} \mathcal{Z}_{\alpha_j}^{\mathbb{S}_j, \mathbb{U}'_j}. \tag{5.15}$$

Thus, (5.14) implies

$$\begin{aligned} &\mathbb{E}(\mathbf{1}(\Xi)|\mathcal{Z}|^p) \\ &\quad \le N^{-p} p^p \sum_{s=1}^{p} N^s \max_{\boldsymbol{\alpha} : |\mathbb{S}(\boldsymbol{\alpha})|=s} \left| \mathbb{E}\left(\mathbf{1}(\Xi) \sum_{\mathbb{U}'_1 \subset \mathbb{S}_1} \cdots \sum_{\mathbb{U}'_p \subset \mathbb{S}_p} \prod_{j=1}^{p} [\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j, \mathbb{U}'_j}\right)\right| \end{aligned} \tag{5.16}$$

(abbreviating $\overline{A^{\mathbb{S},\mathbb{U}}} = \bar{A}^{\mathbb{S},\mathbb{U}}$). Writing $\mathbb{U}_j := \mathbb{U}'_j \cup \{\alpha_j\}$, we have

$$\begin{aligned} &\mathbb{E}(\mathbf{1}(\Xi)|\mathcal{Z}|^p) \\ &\quad \le \left(\frac{p}{N}\right)^p \sum_{s=1}^{p} \sum_{n=1}^{sp} N^s s^n n^p \max\left\{ \left| \mathbb{E}\left(\mathbf{1}(\Xi) \prod_{j=1}^{p} [\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j, \mathbb{U}'_j}\right)\right| : |\mathbb{S}(\boldsymbol{\alpha})| = s, \right. \\ &\qquad\qquad \left. \mathbb{U}'_j \subset \mathbb{S}_j, \sum_{j=1}^{p} |\mathbb{U}_j| = n \right\}. \end{aligned} \tag{5.17}$$

Now we claim that

$$\begin{aligned} &N^s s^n n^p \max\left\{ \left| \mathbb{E}\left(\mathbf{1}(\Xi) \prod_{j=1}^{p} [\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j, \mathbb{U}'_j}\right)\right| : |\mathbb{S}(\boldsymbol{\alpha})| = s, \mathbb{U}'_j \subset \mathbb{S}_j, \sum_{j=1}^{p} |\mathbb{U}_j| = n \right\} \\ &\quad \le \left(C N p^{10} Y (X^2 + N^{-1})\right)^p \end{aligned} \tag{5.18}$$

for some $C > 0$. Then inserting (5.18) into (5.17), we find

$$\mathbb{E}(\mathbf{1}(\Xi)|\mathcal{Z}|^p) \le \left(C p^{11} Y (X^2 + N^{-1})\right)^p, \tag{5.19}$$

which implies (5.12) by Markov's inequality.

It only remains to prove (5.18). We consider two cases: $n \ge 2s$ and $n \le 2s - 1$.

We begin by proving (5.18) for the case $n \ge 2s$. Using Hölder's inequality, we find

$$\left| \mathbb{E}\left(\mathbf{1}(\Xi) \prod_{j=1}^{p} [\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j, \mathbb{U}'_j}\right)\right| \le \left(\prod_{j=1}^{p} \mathbb{E}(\mathbf{1}(\Xi) |[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j, \mathbb{U}'_j}|^p) \right)^{1/p}. \tag{5.20}$$

Applying (5.6) to the right-hand side, we obtain that

$$\left(\prod_{j=1}^{p} \mathbb{E}(\mathbf{1}(\Xi)|[\mathcal{Z}^{\#}_{\alpha_j}]^{\mathbb{S}_j,\mathbb{U}'_j}|^p)\right)^{1/p} \le Y^p (CnX)^n \tag{5.21}$$

since $\sum_j (|\mathbb{U}'_j|+1) = \sum_j |\mathbb{U}'_j| = n$. Combining (5.20), (5.21) and the factor $n \ge p$, we have bounded the left-hand side of (5.18) as follows:

$$\begin{aligned}
N^s s^n n^p \max&\left\{\left|\mathbb{E}\left(\mathbf{1}(\Xi)\prod_{j=1}^{p}[\mathcal{Z}^{\#}_{\alpha_j}]^{\mathbb{S}_j,\mathbb{U}'_j}\right)\right| : |\mathbb{S}(\boldsymbol{\alpha})| = s, \mathbb{U}'_j \subset \mathbb{S}_j, \sum_{j=1}^{p}|\mathbb{U}'_j| = n\right\}\\
&\le N^s Y^p (Cn^2Xs)^n\\
&\le N^s Y^p (Cn^2Xs)^{2s}\\
&\le \bigl(CNp^{10}Y(X^2+N^{-1})\bigr)^p,
\end{aligned}$$

where in the second inequality we used

$$Cn^2Xs \le CXs^3p^2 \le CXp^5 \ll 1$$

[see (5.7)] and $n \ge 2s$, and in the third inequality $s \le p$ and $n \le sp$. This completes the proof of (5.18) for the case $n \ge 2s$.

Now we prove (5.18) for the case $n \le 2s-1$. Fix sets $\mathbb{U}'_j$ with $\sum_j |\mathbb{U}_j| = n$, where we recall that $\mathbb{U}_j := \mathbb{U}'_j \cup \{\alpha_j\}$ and $|\mathbb{U}_j| = |\mathbb{U}'_j| + 1$. By definition of $\mathbb{U}_j$, we have $\alpha_j \in \mathbb{U}_j$ for all j. Since $n \le 2s-1$, we therefore find that there exists a k such that $\alpha_k \in \mathbb{U}_k$ and $\alpha_k \notin \mathbb{U}_j$ for $j \ne k$. By the definitions (5.1) and (5.5), $[\mathcal{Z}^{\#}_{\alpha_j}]^{\mathbb{S}_j,\mathbb{U}'_j}$ is independent of $\mathbb{S}_j \setminus \mathbb{U}'_j$, that is, of $\mathbb{S} \setminus \mathbb{U}_j$. We conclude that

$$\prod_{j\ne k}^{p}[\mathcal{Z}^{\#}_{\alpha_j}]^{\mathbb{S}_j,\mathbb{U}'_j} \tag{5.22}$$

is independent of $\{\alpha_k\}$. Therefore,

$$\begin{aligned}
&\mathbb{E}\left(\mathbf{1}([\Xi]_{\alpha_k})\prod_{j=1}^{p}[\mathcal{Z}^{\#}_{\alpha_j}]^{\mathbb{S}_j,\mathbb{U}'_j}\right)\\
&\quad = \mathbb{E}\left(\left[\prod_{j\ne k}^{p}[\mathcal{Z}^{\#}_{\alpha_j}]^{\mathbb{S},\mathbb{U}_j}\right]\mathbf{1}([\Xi]_{\alpha_k})\mathbb{E}_{\alpha_k}[\mathcal{Z}^{\#}_{\alpha_k}]^{\mathbb{S}_k,\mathbb{U}'_k}\right) = 0.
\end{aligned} \tag{5.23}$$

Thus,

$$\mathbb{E}\left(\mathbf{1}(\Xi)\prod_{j=1}^{p}[\mathcal{Z}^{\#}_{\alpha_j}]^{\mathbb{S}_j,\mathbb{U}'_j}\right) = -\mathbb{E}\left(\mathbf{1}([\Xi]_{\alpha_k}\setminus\Xi)\prod_{j=1}^{p}[\mathcal{Z}^{\#}_{\alpha_j}]^{\mathbb{S}_j,\mathbb{U}'_j}\right), \tag{5.24}$$

which yields

$$\begin{aligned}
&\left|\mathbb{E}\left(\mathbf{1}(\Xi)\prod_{j=1}^{p}[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}\right)\right| \\
&\quad\le \mathbb{E}\left(\mathbf{1}([\Xi]_{\alpha_k}\setminus\Xi)\prod_{j=1}^{p}|[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}|\right) \qquad (5.25)\\
&\quad\le \left\|\mathbf{1}([\Xi]_{\alpha_k}\setminus\Xi)\prod_{j\neq k}^{p}|[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}|\right\|_\infty \mathbb{E}\big(\mathbf{1}([\Xi]_{\alpha_k}\setminus\Xi)|[\mathcal{Z}_{\alpha_k}^{\#}]^{\mathbb{S}_k,\mathbb{U}_k'}|\big).
\end{aligned}$$

Since (5.22) is independent of α_k, we get

$$\begin{aligned}
\left\|\mathbf{1}([\Xi]_{\alpha_k}\setminus\Xi)\prod_{j\neq k}^{p}|[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}|\right\|_\infty &\le \left\|\mathbf{1}([\Xi]_{\alpha_k})\prod_{j\neq k}^{p}|[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}|\right\|_\infty \\
&= \left\|\mathbf{1}(\Xi)\prod_{j\neq k}^{p}|[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}|\right\|_\infty. \qquad (5.26)
\end{aligned}$$

Using the definition of $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ in (5.1) and (5.10), we have

$$|\mathbf{1}(\Xi)[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}| \le YN^C 2^{|\mathbb{U}_j|} \qquad (5.27)$$

and

$$\left|\mathbf{1}(\Xi)\prod_{j\neq k}^{p}[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}\right| \le (YN^C)^{p-1}2^n \le (YN^C)^{p-1}2^{2p}, \qquad (5.28)$$

where we used $s\le p$ and $n\le 2s$ in the last inequality. Combining (5.25), (5.26) and (5.28), we get

$$\left|\mathbb{E}\left(\mathbf{1}(\Xi)\prod_{j=1}^{p}[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}\right)\right| \le (YN^C)^{p-1}2^{2p}\big(\mathbb{E}\mathbf{1}([\Xi]_{\alpha_k}\setminus\Xi)|[\mathcal{Z}_{\alpha_k}^{\#}]^{\mathbb{S}_k,\mathbb{U}_k'}|\big). \qquad (5.29)$$

Applying Schwarz's inequality on the right-hand side, we find

$$\begin{aligned}
&\left|\mathbb{E}\left(\mathbf{1}(\Xi)\prod_{j=1}^{p}[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}\right)\right| \qquad (5.30)\\
&\quad\le (YN^C)^{p-1}2^{2p}\big(\mathbb{P}([\Xi]_{\alpha_k}\setminus\Xi)\big)^{1/2}\big(\mathbb{E}\mathbf{1}([\Xi]_{\alpha_k})|[\mathcal{Z}_{\alpha_k}^{\#}]^{\mathbb{S}_k,\mathbb{U}_k'}|^2\big)^{1/2}.
\end{aligned}$$

Using (5.7), (5.11), (5.9) and that $n\le 2s-1\le 2p$, we get for any $\widetilde{C}>0$

$$\left|\mathbb{E}\left(\mathbf{1}(\Xi)\prod_{j=1}^{p}[\mathcal{Z}_{\alpha_j}^{\#}]^{\mathbb{S}_j,\mathbb{U}_j'}\right)\right| \le (YN^C)^p 2^{2p}\mathbb{P}(\Xi^c)^{1/2} \le Y^p N^{-\widetilde{C}p}. \qquad (5.31)$$

Since $s\le p$, the proof of (5.18) in the case $n\le 2s-1$ is complete. □

5.2. *Decomposition of* G_{ij}. In order to apply Theorem 5.6 to estimate $[Z]$, we need to derive bounds, and hence formulas, for the decomposition $G_{ij}^{\mathbb{S},\mathbb{U}}$ of resolvent matrix elements G_{ij}. As usual, G refers to the resolvent of H at a fixed spectral parameter z (which is suppressed in the notation), that is, $G = G(H)$ is viewed as a function of H. The main result of this section is the bound (5.73) below.

Note that the results in this subsection are entirely deterministic.

LEMMA 5.7. *Let* $z = E + i\eta \in D$, *where* $D \subset \mathbb{C}$ *is some compact domain. Let* $\mathbb{U} \subset \{1, 2, \dots, N\}$ *and*

$$|\mathbb{U}| \le \frac{1}{(\Lambda_o + \Lambda_d)\log N}. \tag{5.32}$$

Then for any $i, j \notin \mathbb{U}$, *we have*

$$\big|G_{ij}^{(\mathbb{U})} - m_{\mathrm{sc}}\delta_{ij}\big| \le C\big(\mathbf{1}(i = j)\Lambda_d + \Lambda_o\big). \tag{5.33}$$

In particular, if $\Lambda_d + \Lambda_o \le (\log N)^{-1}$, *then*

$$\inf_{i \notin \mathbb{U}} \big|G_{ii}^{(\mathbb{U})}\big| \ge c. \tag{5.34}$$

Here the constants c *and* C *depend only on* D.

PROOF. Define

$$B_m := \max\big\{\big|G_{ij}^{(\mathbb{V})} - \delta_{ij}G_{ii}\big| : i, j \notin \mathbb{V}, |\mathbb{V}| = m\big\}. \tag{5.35}$$

In the case $m = 0$, (5.33) follows from the definitions of Λ_o and Λ_d. The estimate (5.34) follows from (5.33), noting that $|m_{\mathrm{sc}}(z)| \ge c$ on a compact domain $z \in D$ with c depending on D. Next, from (3.12) we get

$$G_{ij}^{(k\mathbb{T})} = G_{ij}^{(\mathbb{T})} - \frac{G_{ik}^{(\mathbb{T})}G_{kj}^{(\mathbb{T})}}{G_{kk}^{(\mathbb{T})}} \qquad \text{where } i, j \notin \{k\} \cup \mathbb{T} \text{ and } k \notin \mathbb{T}. \tag{5.36}$$

Assuming (5.34) for $|\mathbb{U}| = m$, we therefore obtain

$$B_{m+1} \le B_m + C_0 B_m^2 \tag{5.37}$$

for some constant $C_0 > 0$ independent of m. This implies that

$$B_{m+1} \le C_0 \sum_{k=0}^{m} B_k^2 + B_0. \tag{5.38}$$

By induction on m one obtains $B_m \le 2B_0$ as long as $C_0 m B_0 \le 1/2$. $\square$

In order to state the next result, we introduce a class of rational functions in resolvent matrix elements. Fix two sets $\mathbb{U} \subset \mathbb{S}$ satisfying $\mathbb{U} \ne \varnothing$. For fixed $n \in \mathbb{N}$ let the following be given:

(i) a sequence of integers $(i_r)_{r=1}^{n+1}$ satisfying $i_k \neq i_{k+1}$ for $1 \leq k \leq n$;

(ii) a collection of sets $(\mathbb{U}_\alpha)_{\alpha=1}^{n}$ satisfying $i_\alpha, i_{\alpha+1} \notin \mathbb{U}_\alpha$ as well as $\mathbb{S} \setminus \mathbb{U} \subset \mathbb{U}_\alpha \subset \mathbb{S}$ for $1 \leq \alpha \leq n$;

(iii) a collection of sets $(\mathbb{T}_\beta)_{\beta=2}^{n}$ satisfying $i_\beta \notin \mathbb{T}_\beta$ as well as $\mathbb{S} \setminus \mathbb{U} \subset \mathbb{T}_\beta \subset \mathbb{S}$ for $2 \leq \beta \leq n$.

Then we define the random variable, parametrized by $(i_r)_{r=1}^{n+1}, (\mathbb{U}_\alpha)_{\alpha=1}^{n}, (\mathbb{T}_\beta)_{\beta=2}^{n}$,

$$F\big((i_r)_{r=1}^{n+1}, (\mathbb{U}_\alpha)_{\alpha=1}^{n}, (\mathbb{T}_\beta)_{\beta=2}^{n}\big) := \frac{P}{Q}, \tag{5.39}$$

where

$$P = \prod_{\alpha=1}^{n} G_{i_\alpha, i_{\alpha+1}}^{(\mathbb{U}_\alpha)}, \qquad Q = \prod_{\beta=2}^{n} G_{i_\beta, i_\beta}^{(\mathbb{T}_\beta)}.$$

Note that F depends on the randomness via the resolvent matrix elements. All matrix elements are off-diagonal in the numerator and diagonal in the denominator; n counts the number of off-diagonal elements in the numerator. The sequence of indices of these matrix elements is consecutive if P/Q is written as an alternating product of off-diagonal elements from the numerator P and reciprocals of diagonal elements from the denominator Q, that is, in the form

$$\frac{P}{Q} = G_{i_1 i_2}^{(\mathbb{U}_1)} \big[G_{i_2 i_2}^{(\mathbb{T}_2)}\big]^{-1} G_{i_2 i_3}^{(\mathbb{U}_2)} \big[G_{i_3 i_3}^{(\mathbb{T}_3)}\big]^{-1} \cdots G_{i_n i_{n+1}}^{(\mathbb{U}_n)}. \tag{5.40}$$

DEFINITION 5.8. For $\mathbb{U} \subset \{1, \ldots, N\}$ and $i, j \notin \mathbb{U}$ define $G_{ij}^{[\mathbb{U}]} := G_{ij}^{(\mathbb{U})}$. For $i, j \notin \mathbb{S}$ and $\mathbb{U} \subset \mathbb{S}$ define $G_{ij}^{\mathbb{S},\mathbb{U}}$ through (5.1).

LEMMA 5.9. *Let $\mathbb{S} \subset \{1, \ldots, N\}$ and $i, j \notin \mathbb{S}$. Then*

$$G_{ij}^{\mathbb{S},\varnothing} = G_{ij}^{(\mathbb{S})}. \tag{5.41}$$

If $\varnothing \neq \mathbb{U} \subset \mathbb{S}$, then $G_{ij}^{\mathbb{S},\mathbb{U}}$ can be written as

$$G_{ij}^{\mathbb{S},\mathbb{U}} = \sum_{n=|\mathbb{U}|+1}^{2|\mathbb{U}|} F_n, \qquad F_n = \sum_{k=1}^{K_n} F_{n,k}, \tag{5.42}$$

where $\sum_{n=|\mathbb{U}|+1}^{2|\mathbb{U}|} K_n \leq 4^{|\mathbb{U}|}|\mathbb{U}|!$ and each $F_{n,k}$ is of the form (5.39) (*with a possible minus sign*), *with $i_2, \ldots, i_n \in \mathbb{U}$, $i_1 = i$, $i_{n+1} = j$, and with some appropriately chosen sets $(\mathbb{U}_\alpha)_{\alpha=1}^{n}$ and $(\mathbb{T}_\beta)_{\beta=2}^{n}$ which may be different for each $F_{n,k}$.*

Note: the index n in F_n and $F_{n,k}$ refers to the number of off-diagonal elements appearing in the rational functions (5.39), while k is just a counting index.

PROOF OF LEMMA 5.9. First, (5.41) follows from (5.1).

It remains to prove (5.42) in the case $\mathbb{U} \neq \varnothing$. Using Definition 5.8 and Remark 5.3, one readily sees that, for a set $\mathbb{T}$ satisfying $\mathbb{T} \cap \mathbb{U} = \varnothing$ and $i, j \notin \mathbb{S} \cup \mathbb{T}$, we have

$$(G_{ij}^{\mathbb{S},\mathbb{U}})^{(\mathbb{T})} = G_{ij}^{\mathbb{S}\cup\mathbb{T},\mathbb{U}}. \tag{5.43}$$

Thus, if $a \in \mathbb{U} \subset \mathbb{S}$, we get from (5.1) and (5.43) that

$$\begin{aligned} G_{ij}^{\mathbb{S},\mathbb{U}} &= G_{ij}^{\mathbb{S}\setminus\{a\},\mathbb{U}\setminus\{a\}} - G_{ij}^{\mathbb{S},\mathbb{U}\setminus\{a\}} \\ &= G_{ij}^{\mathbb{S}\setminus\{a\},\mathbb{U}\setminus\{a\}} - (G_{ij}^{\mathbb{S}\setminus\{a\},\mathbb{U}\setminus\{a\}})^{(a)} \qquad \text{for } i, j \notin \mathbb{S}. \end{aligned} \tag{5.44}$$

In the special case $|\mathbb{U}| = 1$, writing $\mathbb{U} = \{a\}$, we have

$$G_{ij}^{\mathbb{S},\mathbb{U}} = G_{ij}^{\mathbb{S},\{a\}} = G_{ij}^{(\mathbb{S}\setminus\{a\})} - G_{ij}^{(\mathbb{S})}. \tag{5.45}$$

Using (3.12), we obtain (5.42) for the case $|\mathbb{U}| = 1$, that is,

$$G_{ij}^{\mathbb{S},\{a\}} = G_{ij}^{(\mathbb{S}\setminus\{a\})} - G_{ij}^{(\mathbb{S})} = \frac{G_{ia}^{(\mathbb{S}\setminus\{a\})} G_{aj}^{(\mathbb{S}\setminus\{a\})}}{G_{aa}^{(\mathbb{S}\setminus\{a\})}}. \tag{5.46}$$

For a general set $\mathbb{U}$ with $|\mathbb{U}| \geq 2$, using (5.44), we can write $G^{\mathbb{S},\mathbb{U}}$ iteratively as $F - F^{(a)}$, where F itself is of the form $E - E^{(b)}$ for some appropriate E. For example, for $a, b \in \mathbb{U}$ we have

$$\begin{aligned} G_{ij}^{\mathbb{S},\mathbb{U}} &= G_{ij}^{\mathbb{S}\setminus\{a\},\mathbb{U}\setminus\{a\}} - (G_{ij}^{\mathbb{S}\setminus\{a\},\mathbb{U}\setminus\{a\}})^{(a)} \\ &= G_{ij}^{\mathbb{S}\setminus\{ab\},\mathbb{U}\setminus\{ab\}} - (G_{ij}^{\mathbb{S}\setminus\{ab\},\mathbb{U}\setminus\{ab\}})^{(b)} \\ &\quad - (G_{ij}^{\mathbb{S}\setminus\{ab\},\mathbb{U}\setminus\{ab\}} - (G_{ij}^{\mathbb{S}\setminus\{ab\},\mathbb{U}\setminus\{ab\}})^{(b)})^{(a)}. \end{aligned}$$

Recall $F^{(a)} = F \circ \pi_a$ from Definition 5.2. Then to prove (5.42) in the case $\mathbb{U}$ with $|\mathbb{U}| \geq 2$, we use induction on $|\mathbb{U}|$. The key step is Lemma 5.10 below, which contains the required properties of $F - F^{(a)}$. Its proof will be given later.

LEMMA 5.10. *Let F be of the form* (5.39). *We assume that*

$$\left| \left(\bigcup_{\alpha=1}^{n} \mathbb{U}_\alpha \right) \cup \left(\bigcup_{\beta=2}^{n} \mathbb{T}_\beta \right) \right| \leq \frac{1}{(\Lambda_o + \Lambda_d) \log N} - 1. \tag{5.47}$$

If

$$s \notin \{i_1, i_2, i_3, \ldots, i_{n+1}\} \cup \left(\bigcup_{\alpha=1}^{n} \mathbb{U}_\alpha \right) \cup \left(\bigcup_{\beta=1}^{n-1} \mathbb{T}_\beta \right), \tag{5.48}$$

then $F - F^{(s)}$ is equal to the sum (with signs $\pm$) of $2n-1$ terms of the form (5.39),

$$F - F^{(s)} = \sum_{l=1}^{n} F_l^A\big((\tilde{i}_{l,r}^A)_{r=1}^{\tilde{n}^A+1}, (\widetilde{\mathbb{U}}_{l,\alpha}^A)_{\alpha=1}^{\tilde{n}^A}, (\widetilde{\mathbb{T}}_{l,\beta}^A)_{\beta=2}^{\tilde{n}^A}\big) + \sum_{l=1}^{n-1} F_l^B\big((\tilde{i}_{l,r}^B)_{r=1}^{\tilde{n}^B+1}, (\widetilde{\mathbb{U}}_{l,\alpha}^B)_{\alpha=1}^{\tilde{n}^B}, (\widetilde{\mathbb{T}}_{l,\beta}^B)_{\beta=2}^{\tilde{n}^B}\big), \tag{5.49}$$

where the new arguments, carrying a tilde, satisfy the following relations:

(i)

$$\tilde{n}^A = n+1 \quad \textit{and} \quad \tilde{n}^B = n+2. \tag{5.50}$$

(ii) *For $1 \le l \le n$, the family $(\tilde{i}_{l,r}^A)$ is given by*

$$(\tilde{i}_{l,1}^A, \tilde{i}_{l,2}^A, \tilde{i}_{l,3}^A, \dots, \tilde{i}_{l,n+2}^A) := (i_1, i_2, \dots, i_l, s, i_{l+1}, \dots, i_{n+1}). \tag{5.51}$$

For $1 \le l \le n-1$, the family $(\tilde{i}_{l,r}^B)$ is given by

$$(\tilde{i}_{l,1}^B, \tilde{i}_{l,2}^B, \tilde{i}_{l,3}^B, \dots, \tilde{i}_{l,n+3}^B) := (i_1, i_2, \dots, i_l, i_{l+1}, s, i_{l+1}, i_{l+2}, \dots, i_{n+1}). \tag{5.52}$$

(iii) *All sets $\widetilde{\mathbb{U}}_{l,\alpha}^A$, $\widetilde{\mathbb{T}}_{l,\beta}^A$, $\widetilde{\mathbb{U}}_{l,\alpha}^B$ and $\widetilde{\mathbb{T}}_{l,\beta}^B$ appearing in (5.49) are subsets of*

$$\left(\bigcup_{\alpha=1}^{n} \mathbb{U}_\alpha\right) \cup \left(\bigcup_{\beta=2}^{n} \mathbb{T}_\beta\right) \cup \{s\}. \tag{5.53}$$

Now we return to complete the proof for Lemma 5.9. Using (5.44), we get for $s \in \mathbb{U}$ and $i, j \notin \mathbb{S}$ that

$$G_{ij}^{\mathbb{S},\mathbb{U}} = G_{ij}^{\mathbb{S}\setminus\{s\},\mathbb{U}\setminus\{s\}} - \big(G_{ij}^{\mathbb{S}\setminus\{s\},\mathbb{U}\setminus\{s\}}\big)^{(s)}. \tag{5.54}$$

Using induction on $|\mathbb{U}|$ and applying the decomposition (5.42) to $G_{ij}^{\mathbb{S}\setminus\{s\},\mathbb{U}\setminus\{s\}}$, we get

$$G_{ij}^{\mathbb{S}\setminus\{s\},\mathbb{U}\setminus\{s\}} = \sum_{n=|\mathbb{U}|}^{2|\mathbb{U}|-2} F_n, \qquad F_n = \sum_{k=1}^{K_n'} F_{n,k}, \tag{5.55}$$

where $\sum_{n=|\mathbb{U}|}^{2|\mathbb{U}|-2} K_n' \le 4^{|\mathbb{U}|-1}(|\mathbb{U}|-1)!$ and each $F_{n,k}$ is of the form (5.39) (with a possible minus sign) with $i_2, \dots, i_n \in \mathbb{U} \setminus \{s\}$, $i_1 = i$, $i_{n+1} = j$, and with some appropriately chosen sets $(\mathbb{U}_\alpha)_{\alpha=1}^n$, $(\mathbb{T}_\beta)_{\beta=2}^n$ satisfying

$$\mathbb{S} \setminus \mathbb{U} \subset \mathbb{U}_\alpha, \qquad \mathbb{T}_\beta \subset \mathbb{S} \setminus \{s\}, \qquad 1 \le \alpha \le n, 2 \le \beta \le n.$$

Now from (5.54) we get

$$G_{ij}^{\mathbb{S},\mathbb{U}} = \sum_{n=|\mathbb{U}|}^{2|\mathbb{U}|-2} \sum_{k=1}^{K_n'} \big(F_{n,k} - (F_{n,k})^{(s)}\big). \tag{5.56}$$

Moreover, using (5.49), we get

$$F_{n,k} - (F_{n,k})^{(s)} = \sum_{l=1}^{n} F^A_{n,k,l} + \sum_{l=1}^{n-1} F^B_{n,k,l}, \tag{5.57}$$

where each $F^A_{n,k,l}$ and $F^B_{n,k,l}$ is of the form (5.39) (with a possible minus sign) with $i_1 = i$, $i_{m+1} = j$, where $m = n+2$ for $F^A_{n,k,l}$ and $m = n+3$ for $F^B_{n,k,l}$, and the other indices belong to $\mathbb{U}$. Here the sets $(\widetilde{\mathbb{U}}_\alpha)_{\alpha=1}^m$ and $(\widetilde{\mathbb{T}}_\beta)_{\beta=2}^m$ satisfy

$$\mathbb{S} \setminus \mathbb{U} \subset \widetilde{\mathbb{U}}_\alpha, \qquad \widetilde{\mathbb{T}}_\beta \subset \mathbb{S}, \qquad 1 \le \alpha \le m, 2 \le \beta \le m.$$

Furthermore, with (5.50), the number of off-diagonal elements in the numerators of $F^A_{n,k,l}$ and $F^B_{n,k,l}$ are $n+1$ and $n+2$, respectively. Hence, together with (5.56), we obtain

$$G_{ij}^{\mathbb{S},\mathbb{U}} = \sum_{n=|\mathbb{U}|}^{2|\mathbb{U}|-2} \sum_{k=1}^{K'_n} \left(\sum_{l=1}^{n} F^A_{n,k,l} + \sum_{l=1}^{n-1} F^B_{n,k,l} \right).$$

With the assumption of $\sum_{n=|\mathbb{U}|}^{2|\mathbb{U}|-2} K'_n \le 4^{|\mathbb{U}|-1}(|\mathbb{U}|-1)!$ for the summation bounds in (5.55), we know that $G_{ij}^{\mathbb{S},\mathbb{U}}$ can be written in the form (5.42) with $\sum_{n=|\mathbb{U}|+1}^{2|\mathbb{U}|} K_n \le 4^{|\mathbb{U}|}|\mathbb{U}|!$. This completes the proof of Lemma 5.9. □

PROOF OF LEMMA 5.10. Using (3.12), it is easy to derive the following two identities for $s \notin \mathbb{U}$:

$$G_{ij}^{(\mathbb{U})} = G_{ij}^{(\mathbb{U}s)} + \frac{G_{is}^{(\mathbb{U})} G_{sj}^{(\mathbb{U})}}{G_{ss}^{(\mathbb{U})}} \qquad \text{for } i, j \notin \mathbb{U} \cup \{s\}, \tag{5.58}$$

$$\frac{1}{G_{kk}^{(\mathbb{U})}} = \frac{1}{G_{kk}^{(\mathbb{U}s)}} + \frac{G_{ks}^{(\mathbb{U})} G_{sk}^{(\mathbb{U})}}{G_{kk}^{(\mathbb{U}s)} G_{ss}^{(\mathbb{U})} G_{kk}^{(\mathbb{U})}} \qquad \text{for } k \notin \mathbb{U} \cup \{s\}. \tag{5.59}$$

Now (5.58) implies that Lemma 5.10 holds in the case $n = 1$. We shall first prove it for the case $n = 2$, and then give the proof of the general case. If $n = 2$, then by assumption F has the form

$$F = \frac{G_{ij}^{(\mathbb{U})} G_{jk}^{(\mathbb{V})}}{G_{jj}^{(\mathbb{T})}} \tag{5.60}$$

with some sets $\mathbb{U}$, $\mathbb{V}$, $\mathbb{T}$ and indices i, j, k. For $s \notin \mathbb{U} \cup \mathbb{V} \cup \mathbb{T} \cup \{ijk\}$ we get from (5.58) that

$$F = \frac{G_{ij}^{(\mathbb{U}s)} G_{jk}^{(\mathbb{V})}}{G_{jj}^{(\mathbb{T})}} + \frac{G_{is}^{(\mathbb{U})} G_{sj}^{(\mathbb{U})} G_{jk}^{(\mathbb{V})}}{G_{ss}^{(\mathbb{U})} G_{jj}^{(\mathbb{T})}}. \tag{5.61}$$

Next, using (5.59) on the first term, we obtain

$$F = \frac{G_{ij}^{(\mathbb{U}s)} G_{jk}^{(\mathbb{V})}}{G_{jj}^{(\mathbb{T}s)}} + \frac{G_{ij}^{(\mathbb{U}s)} G_{js}^{(\mathbb{T})} G_{sj}^{(\mathbb{T})} G_{jk}^{(\mathbb{V})}}{G_{jj}^{(\mathbb{T}s)} G_{ss}^{(\mathbb{T})} G_{jj}^{(\mathbb{T})}} + \frac{G_{is}^{(\mathbb{U})} G_{sj}^{(\mathbb{U})} G_{jk}^{(\mathbb{V})}}{G_{ss}^{(\mathbb{U})} G_{jj}^{(\mathbb{T})}}. \tag{5.62}$$

Using (5.58) again on the first term, we have

$$F = F^{(s)} + \frac{G_{ij}^{(\mathbb{U}s)} G_{js}^{(\mathbb{V})} G_{sk}^{(\mathbb{V})}}{G_{jj}^{(\mathbb{T}s)} G_{ss}^{(\mathbb{V})}} + \frac{G_{ij}^{(\mathbb{U}s)} G_{js}^{(\mathbb{T})} G_{sj}^{(\mathbb{T})} G_{jk}^{(\mathbb{V})}}{G_{jj}^{(\mathbb{T}s)} G_{ss}^{(\mathbb{T})} G_{jj}^{(\mathbb{T})}} + \frac{G_{is}^{(\mathbb{U})} G_{sj}^{(\mathbb{U})} G_{jk}^{(\mathbb{V})}}{G_{ss}^{(\mathbb{U})} G_{jj}^{(\mathbb{T})}}. \tag{5.63}$$

One can easily check that the last three terms are of the form (5.39), and the indices satisfy (5.50)–(5.53). This completes the proof for Lemma 5.10 in the case $n = 2$.

Now we consider the case of a general n. Inserting (5.58) and (5.59) into each term in (5.39), we have

$$F\big((i_r)_{r=1}^{n+1}, (\mathbb{U}_\alpha)_{\alpha=1}^{n}, (\mathbb{T}_\beta)_{\beta=2}^{n}\big) = \frac{P}{Q}, \tag{5.64}$$

where

$$P = \prod_{\alpha=1}^{n} G_{i_\alpha,i_{\alpha+1}}^{(\mathbb{U}_\alpha)} = \prod_{\alpha=1}^{n} \bigg(G_{i_\alpha,i_{\alpha+1}}^{(\mathbb{U}_\alpha s)} + \frac{G_{i_\alpha,s}^{(\mathbb{U}_\alpha)} G_{s i_{\alpha+1}}^{(\mathbb{U}_\alpha)}}{G_{ss}^{(\mathbb{U}_\alpha)}} \bigg)$$

and

$$Q^{-1} = \prod_{\beta=2}^{n} \big(G_{i_\beta,i_\beta}^{(\mathbb{T}_\beta)}\big)^{-1} = \prod_{\beta=2}^{n} \bigg(\frac{1}{G_{i_\beta,i_\beta}^{(\mathbb{T}_\beta s)}} + \frac{G_{i_\beta s}^{(\mathbb{T}_\beta)} G_{s i_\beta}^{(\mathbb{T}_\beta)}}{G_{i_\beta i_\beta}^{(\mathbb{T}_\beta)} G_{ss}^{(\mathbb{T}_\beta)} G_{i_\beta i_\beta}^{(\mathbb{T}_\beta s)}} \bigg).$$

On the other hand,

$$\big(F\big((i_r)_{r=1}^{n+1}, (\mathbb{U}_\alpha)_{\alpha=1}^{n}, (\mathbb{T}_\beta)_{\beta=2}^{n}\big)\big)^{(s)} = \frac{P^{(s)}}{Q^{(s)}}, \tag{5.65}$$

where

$$P^{(s)} = \prod_{\alpha=1}^{n} G_{i_\alpha,i_{\alpha+1}}^{(\mathbb{U}_\alpha s)} \quad \text{and} \quad \big(Q^{(s)}\big)^{-1} = \prod_{\beta=2}^{n} \big(G_{i_\beta,i_\beta}^{(\mathbb{T}_\beta s)}\big)^{-1}.$$

For $m \in \mathbb{N}$, we write, using (5.58) and (5.59),

$$G_{i_m,i_{m+1}}^{(\mathbb{U}_m)} = A_{2m-1} + B_{2m-1}, \tag{5.66}$$

where

$$A_{2m-1} := G_{i_m,i_{m+1}}^{(\mathbb{U}_m s)} \quad \text{and} \quad B_{2m-1} := \frac{G_{i_m,s}^{(\mathbb{U}_m)} G_{s,i_{m+1}}^{(\mathbb{U}_m)}}{G_{ss}^{(\mathbb{U}_m)}}. \tag{5.67}$$

Similarly, we write

$$(G^{(\mathbb{T}_{m+1})}_{i_{m+1},i_{m+1}})^{-1} = A_{2m} + B_{2m}, \tag{5.68}$$

where

$$A_{2m} := (G^{(\mathbb{T}_{m+1}s)}_{i_{m+1},i_{m+1}})^{-1} \quad \text{and} \quad B_{2m} := \frac{G^{(\mathbb{T}_{m+1})}_{i_{m+1}s} G^{(\mathbb{T}_{m+1})}_{si_{m+1}}}{G^{(\mathbb{T}_{m+1})}_{i_{m+1}i_{m+1}} G^{(\mathbb{T}_{m+1})}_{ss} G^{(\mathbb{T}_{m+1}s)}_{i_{m+1}i_{m+1}}}. \tag{5.69}$$

Then

$$F = \prod_{m=1}^{2n-1} (A_m + B_m), \qquad F^{(s)} = \prod_{m=1}^{2n-1} A_m.$$

To complete the proof, we use the identity

$$\prod_{m=1}^{2n-1} (A_m + B_m) - \prod_{m=1}^{2n-1} A_m = \sum_{m=1}^{2n-1} \left(\prod_{j=1}^{m-1} A_j\right) B_m \left(\prod_{j=m+1}^{2n-1} (A_j + B_j)\right). \tag{5.70}$$

It is easy to check that, for any term of the form

$$\left(\prod_{j=1}^{m-1} A_j\right) B_m \left(\prod_{j=m+1}^{2n-1} (A_j + B_j)\right)$$

in the sum (5.70), the desired properties (5.49)–(5.53) hold. □

We may now easily obtain the following bound on $G^{\mathbb{S},\mathbb{U}}_{ij}$.

LEMMA 5.11. *Let* $\mathbb{U} \subset \mathbb{S} \subset \{1, 2, \ldots, N\}$ *and*

$$|\mathbb{S}| \le \frac{1}{(\Lambda_o + \Lambda_d) \log N}. \tag{5.71}$$

Then

$$|G^{\mathbb{S},\varnothing}_{ij} - m_{sc}\delta_{ij}| \le C\big(\mathbf{1}(i = j)\Lambda_d + \Lambda_o\big). \tag{5.72}$$

If in addition $\mathbb{U} \ne \varnothing$ *and* $i, j \notin \mathbb{S}$, *then*

$$|G^{\mathbb{S},\mathbb{U}}_{ij}| \le (C|\mathbb{U}|\Lambda_o)^{|\mathbb{U}|+1} \tag{5.73}$$

and

$$|(1/G_{ii})^{\mathbb{S},\mathbb{U}}| \le (C|\mathbb{U}|\Lambda_o)^{|\mathbb{U}|+1}. \tag{5.74}$$

PROOF. The estimate (5.72) follows (5.41) and (5.33). In order to prove (5.73), we apply Lemma 5.9 to each $G_{ij}^{\mathbb{S},\mathbb{U}}$, and get

$$G_{ij}^{\mathbb{S},\mathbb{U}} = \sum_{n=|\mathbb{U}|+1}^{2|\mathbb{U}|} F_n, \qquad F_n = \sum_{k=1}^{K_n} F_{n,k}, \tag{5.75}$$

where $\sum_{n=|\mathbb{U}|+1}^{2|\mathbb{U}|} K_n \le 4^{|\mathbb{U}|}|\mathbb{U}|!$. Here each $F_{n,k}$ is of the form (5.39) (with a possible minus sign), where n counts the number of off-diagonal elements in the numerator; the indices satisfy $i_2, \dots, i_n \in \mathbb{U}$, $i_1 = i$, $i_{n+1} = j$. Note that the factors P in (5.39) are the product of off-diagonal terms and the factors Q the product of diagonal terms. Applying (5.33) and (5.34) on the off-diagonal and diagonal terms in P and Q, we get

$$F_{n,k} \le \frac{(C\Lambda_o)^n}{c^{n-1}} \le (C\Lambda_o)^n. \tag{5.76}$$

Together with $\sum_{n=|\mathbb{U}|+1}^{2|\mathbb{U}|} K_n \le 4^{|\mathbb{U}|}|\mathbb{U}|!$, this implies (5.73).

In order to prove (5.74), we observe that, similarly to Lemma 5.9, we have

$$|(1/G_{ii})^{\mathbb{S},\mathbb{U}}| \le (C|\mathbb{U}|)^{|\mathbb{U}|+1} \frac{(\max_{k,j\notin\mathbb{T},\mathbb{T}\subset\mathbb{S}} |G_{kj}^{(\mathbb{T})}|)^{|\mathbb{U}|+1}}{(\min_{j\notin\mathbb{T},\mathbb{T}\subset\mathbb{S}} |G_{jj}^{(\mathbb{T})}|)^{|\mathbb{U}|+2}} \tag{5.77}$$

provided that

$$\max_{k,j\notin\mathbb{T},\mathbb{T}\subset\mathbb{S}} |G_{kj}^{(\mathbb{T})}| \le \min_{j\notin\mathbb{T},\mathbb{T}\subset\mathbb{S}} |G_{jj}^{(\mathbb{T})}|.$$

Hence, (5.74) follows. □

5.3. *Proof of Lemma* 4.1. Observe first that (4.3) follows immediately from (4.2) and Lemma 3.15. It therefore remains to prove (4.2).

We define the event Ξ by requiring that on it (4.1) and the following two events hold:

(i) For every $z \in \widetilde{D}$ we have

$$\begin{aligned} \Lambda_o(z) &\le C\left(\frac{1}{q} + (\log N)^{2\xi}\Psi(z)\right) \\ &\le C\left(\frac{1}{q} + (\log N)^{2\xi}\sqrt{\frac{\operatorname{Im} m_{sc}(z) + \gamma(z)}{N\eta}}\right). \end{aligned} \tag{5.78}$$

(ii) For every $z \in \widetilde{D}$ we have

$$\begin{aligned} \max_i |G_{ii}(z) - m(z)| &\le C\left(\frac{(\log N)^{\xi}}{q} + (\log N)^{2\xi}\Psi(z)\right) \\ &\le C\left(\frac{(\log N)^{\xi}}{q} + (\log N)^{2\xi}\sqrt{\frac{\operatorname{Im} m_{sc}(z) + \gamma(z)}{N\eta}}\right). \end{aligned} \tag{5.79}$$

Now Theorem 3.1, Lemmas 3.13 and 3.14, as well as (4.1) and $\widetilde{D} \subset D_L$ imply that Ξ holds with $(\xi - 1/2, \nu)$-high probability. Note that here we reduced the ξ to $\xi - 1/2$ to account for the intersection of three events of (ξ, ν)-high probability. It is crucial that ν remain constant in this step, as in some applications, such as Theorem 2.8, it is iterated.

We write Z_i as

$$Z_i = \sum_k^{(i)} \biggl(h_{ik}^2 - \frac{1}{N}\biggr) G_{kk}^{(i)} + \sum_{k \neq l}^{(i)} h_{ik} G_{kl}^{(i)} h_{li}. \tag{5.80}$$

Lemma 4.1 follows from the next two lemmas. As before, we shall consistently omit the spectral parameter $z \in \widetilde{D}$ from the notation in the following arguments.

LEMMA 5.12. *On Ξ we have with (ξ, ν)-high probability*

$$\begin{aligned} &\biggl| \mathbf{1}(\Xi) \frac{1}{N} \sum_i \sum_k^{(i)} \biggl(h_{ik}^2 - \frac{1}{N}\biggr) G_{kk}^{(i)} \biggr| \\ &\qquad \le (\log N)^{4\xi} \biggl(\frac{1}{q^2} + \frac{1}{(N\eta)^2} + \frac{\operatorname{Im} m_{\mathrm{sc}} + \gamma}{N\eta} \biggr). \end{aligned} \tag{5.81}$$

LEMMA 5.13. *On Ξ we have with $(\xi - 2, \nu)$-high probability*

$$\begin{aligned} &\biggl| \mathbf{1}(\Xi) \frac{1}{N} \sum_i \sum_{k \neq l}^{(i)} h_{ik} G_{kl}^{(i)} h_{li} \biggr| \\ &\qquad \le (\log N)^{14\xi} \biggl(\frac{1}{q^2} + \frac{1}{(N\eta)^2} + (\log N)^{4\xi} \frac{\operatorname{Im} m_{\mathrm{sc}} + \gamma}{N\eta} \biggr). \end{aligned} \tag{5.82}$$

PROOF OF LEMMA 5.12. We split the sum inside the absolute value on the left-hand side of (5.81) as

$$\begin{aligned} &\frac{1}{N} \sum_{i \neq k} \biggl(h_{ik}^2 - \frac{1}{N}\biggr) m + \frac{1}{N} \sum_{i \neq k} \biggl(h_{ik}^2 - \frac{1}{N}\biggr) \bigl(m^{(i)} - m\bigr) \\ &\qquad + \frac{1}{N} \sum_{i \neq k} \biggl(h_{ik}^2 - \frac{1}{N}\biggr) \bigl(G_{kk}^{(i)} - m^{(i)}\bigr). \end{aligned} \tag{5.83}$$

In order to estimate the first term of (5.83), we use the estimate (3.17) [with $(h_{ik}^2 - N^{-1})$ playing the role of a_i, and setting $A_i = N^{-1}$, $\alpha = 2$, $\beta = -2$ and $\gamma = 1$] to get, with (ξ, ν)-high probability,

$$\biggl| \frac{1}{N} \sum_{i \neq k} \biggl(h_{ik}^2 - \frac{1}{N}\biggr) \biggr| \le (\log N)^{\xi} \frac{1}{N^{1/2} q}. \tag{5.84}$$

Therefore,

$$\begin{aligned}
\left|\mathbf{1}(\Xi)\left(\frac{1}{N}\sum_{i\neq k}\left(h_{ik}^2-\frac{1}{N}\right)\right)m\right| &\le (\log N)^{\xi}\left|\mathbf{1}(\Xi)\frac{m}{N^{1/2}q}\right| \\
&\le C(\log N)^{\xi}\frac{1}{N^{1/2}q}.
\end{aligned} \tag{5.85}$$

Similarly, in order to estimate the second term of (5.83), we fix i and sum over k, which yields with (ξ,ν)-high probability

$$\left|\max_i \sum_k^{(i)}\left(h_{ik}^2-\frac{1}{N}\right)\right| \le (\log N)^{\xi} q^{-1}, \tag{5.86}$$

where the sum over k was estimated by (3.17). This yields with (ξ,ν)-high probability

$$\begin{aligned}
&\left|\mathbf{1}(\Xi)\frac{1}{N}\sum_i\sum_k^{(i)}\left(h_{ik}^2-\frac{1}{N}\right)\left(m^{(i)}-m\right)\right| \\
&\quad\le \frac{1}{N}\sum_i\left|\mathbf{1}(\Xi)(\log N)^{\xi}q^{-1}\left(m^{(i)}-m\right)\right|.
\end{aligned} \tag{5.87}$$

Using (3.31), we have in Ξ

$$\left|m^{(i)}-m\right| = \left|-\frac{1}{N}\sum_j\frac{G_{ji}G_{ij}}{G_{ii}}\right| \le O\left(\frac{\operatorname{Im} G_{ii}}{\eta}\right).$$

Thus, we get with (ξ,ν)-high probability

$$\left|\mathbf{1}(\Xi)\frac{1}{N}\sum_{i\neq k}\left(h_{ik}^2-\frac{1}{N}\right)\left(m^{(i)}-m\right)\right| \le C(\log N)^{\xi}\frac{\operatorname{Im} m_{sc}+\gamma}{qN\eta}. \tag{5.88}$$

Finally, we estimate the third term of (5.83). First, with (3.12) and $|m_{sc}|\ge c$, we note that if $\Lambda_d \ll 1$, then

$$\left|G_{ij}-G_{ij}^{(k)}\right| \le C\Lambda_o^2 \qquad \text{for } i,j\neq k. \tag{5.89}$$

Together with (5.79) we get with (ξ,ν)-high probability

$$\max_{k\neq i}\left|\left(G_{kk}^{(i)}-m^{(i)}\right)\right| \le C(\log N)^{2\xi}\left(\frac{1}{q}+\sqrt{\frac{\operatorname{Im} m_{sc}+\gamma}{N\eta}}\right). \tag{5.90}$$

Then we use (3.19) [with $(h_{ik}^2-N^{-1})$ playing the role of a_k and $G_{kk}^{(i)}-m^{(i)}$ playing the role of A_k, and setting $\alpha=2$, $\beta=-2$ and $\gamma=1$] to get, with (ξ,ν)-high probability,

$$\max_i\left|\sum_k^{(i)}\left(h_{ik}^2-\frac{1}{N}\right)\left(G_{kk}^{(i)}-m^{(i)}\right)\right| \le (\log N)^{4\xi}\left(\frac{1}{q}+\sqrt{\frac{\operatorname{Im} m_{sc}+\gamma}{N\eta}}\right)q^{-1}. \tag{5.91}$$

Hence, we have with (ξ,ν)-high probability

$$\begin{aligned} &\left|\mathbf{1}(\Xi)\frac{1}{N}\sum_i\Biggl(\sum_k^{(i)}\Bigl(h_{ik}^2-\frac{1}{N}\Bigr)(G_{kk}^{(i)}-m^{(i)})\Biggr)\right| \\ &\qquad\le(\log N)^{4\xi}\left(\frac{1}{q}+\sqrt{\frac{\operatorname{Im}m_{\mathrm{sc}}+\gamma}{N\eta}}\right)q^{-1}. \end{aligned}\tag{5.92}$$

Note that, in applying (3.19), we used that the family $\{h_{ik}^2-N^{-1}\}_k$ is independent of the family $\{G_{kk}^{(i)}-m^{(i)}\}_k$. Combining (5.85), (5.88) and (5.92), we obtain (5.81). □

PROOF OF LEMMA 5.13. We shall apply Theorem 5.6 to the quantities

$$\mathcal{Z}_i:=\sum_{k\neq l}^{(i)}h_{ik}G_{kl}^{(i)}h_{li},\qquad \mathcal{Z}_i^{[\mathbb{V}]}:=\mathbf{1}(i\notin\mathbb{V})\sum_{k\neq l}^{(i\mathbb{V})}h_{ik}G_{kl}^{(i\mathbb{V})}h_{li},\tag{5.93}$$

and define Ξ as in the beginning of Section 5.3, that is, Ξ is defined by requiring that (4.1) and (5.78)–(5.79) hold. Recall that the collection of random variables $Z_i^{[\mathbb{V}]}$ generates random variables $Z_i^{\mathbb{S},\mathbb{U}}$ for any $\mathbb{U}\subset\mathbb{S}$ by (5.1). Let

$$p:=(\log N)^{\xi-3/2}.\tag{5.94}$$

Next, choose

$$\begin{aligned} X&:=\frac{1}{q}+(\log N)^{2\xi}\sqrt{\frac{\operatorname{Im}m_{\mathrm{sc}}+\gamma}{N\eta}}, \\ Y&:=(\log N)^{2\xi}. \end{aligned}\tag{5.95}$$

[In other words, X is defined as the right-hand side of (5.78) up to a constant.] We now derive a bound which implies both (5.6) and (5.9), that is, we establish the assumptions (ii) and (iii) of Theorem 5.6. To this end, we shall prove the stronger statement that, for $i\notin\mathbb{S}$, $r\le p$ and any sets $\mathbb{U}\subset\mathbb{S}$ with $|\mathbb{S}|\le p$, we have

$$\mathbb{E}(\mathbf{1}([\Xi]_i)|\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}|^r)\le(Y(CXu)^u)^r\qquad\text{for }u=|\mathbb{U}|+1.\tag{5.96}$$

Using the assumptions of Lemma 4.1, we have in $\widetilde{D}$ that

$$q\ge(\log N)^{5\xi},\qquad N\eta\ge(\log N)^{14\xi},\qquad \gamma\le(\log N)^{-\xi}.\tag{5.97}$$

It is therefore easy to check that $\mathcal{Z}_i$ and Ξ satisfy the assumptions (i), (iv) and (v) of Theorem 5.6. Thus, the conclusion of Theorem 5.6, (5.12), implies the claim (5.82).

It remains to prove (5.96). By the definition of $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ in (5.1) and (5.93), for $i \notin \mathbb{S}$, we have

$$
\begin{aligned}
\mathcal{Z}_i^{\mathbb{S},\mathbb{U}} &= (-1)^{|\mathbb{S}\backslash\mathbb{U}|} \sum_{\mathbb{V}:\,\mathbb{S}\backslash\mathbb{U}\subset\mathbb{V}\subset\mathbb{S}} (-1)^{|\mathbb{V}|} \sum_{k\neq l}^{(i\mathbb{V})} h_{ik} G_{kl}^{(i\mathbb{V})} h_{li} \\
&= (-1)^{|\mathbb{S}\backslash\mathbb{U}|} \sum_{k\neq l}^{(i\cup\mathbb{S}\backslash\mathbb{U})} \sum_{\mathbb{V}:\,\mathbb{S}\backslash\mathbb{U}\subset\mathbb{V}\subset\mathbb{S}\backslash\{k,l\}} (-1)^{|\mathbb{V}|} h_{ik} G_{kl}^{(i\mathbb{V})} h_{li} \\
&= \sum_{k\neq l}^{(i\cup\mathbb{S}\backslash\mathbb{U})} h_{ik}h_{li} \sum_{\mathbb{V}:\,\mathbb{S}\backslash\mathbb{U}\subset\mathbb{V}\subset\mathbb{S}\backslash\{k,l\}} (-1)^{|\mathbb{S}\backslash\mathbb{U}|}(-1)^{|\mathbb{V}|} G_{kl}^{(i\mathbb{V})} \\
&= \sum_{k\neq l}^{(i\cup\mathbb{S}\backslash\mathbb{U})} h_{ik}h_{li} [G_{kl}]^{(\mathbb{S}i)\backslash\{k,l\},\mathbb{U}\backslash\{k,l\}},
\end{aligned}
\tag{5.98}
$$

where in the last equality we used the definition of $G^{\mathbb{S},\mathbb{U}}$, Definition 5.8. Thus, we may write

$$
\mathcal{Z}_i^{\mathbb{S},\mathbb{U}} = A_1 + A_2 + A_3 + A_4, \tag{5.99}
$$

where

$$
\begin{aligned}
A_1 &:= \sum_{k\in\mathbb{U}} \sum_{l\in\mathbb{U}\backslash\{k\}} h_{ik}h_{li}[G_{kl}]^{(\mathbb{S}i)\backslash\{k,l\},\mathbb{U}\backslash\{k,l\}}, \\
A_2 &:= \sum_{k\in\mathbb{U}} \sum_{l}^{(\mathbb{S}ik)} h_{ik}h_{li}[G_{kl}]^{(\mathbb{S}i)\backslash\{k\},\mathbb{U}\backslash\{k\}}, \\
A_3 &:= \sum_{l\in\mathbb{U}} \sum_{k}^{(\mathbb{S}il)} h_{ik}h_{li}[G_{kl}]^{(\mathbb{S}i)\backslash\{l\},\mathbb{U}\backslash\{l\}}, \\
A_4 &:= \sum_{k\neq l}^{(\mathbb{S}i)} h_{ik}h_{li}[G_{kl}]^{(\mathbb{S}i),\mathbb{U}}.
\end{aligned}
$$

Now we have

$$
\begin{aligned}
\mathbb{E}(\mathbf{1}([\Xi]_i)|\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}|^r) &= \mathbb{E}\big(\mathbf{1}([\Xi]_i)|A_1 + A_2 + A_3 + A_4|^r\big) \\
&\leq 4^r \sum_{j=1}^{4} \mathbb{E}(\mathbf{1}([\Xi]_i)|A_j|^r),
\end{aligned}
\tag{5.100}
$$

and we are going to bound $\mathbb{E}(\mathbf{1}([\Xi]_i)|A_j|^r)$ for each $j = 1, 2, 3, 4$. Using the assumption (4.1), that is,

$$
\Lambda \leq \gamma \leq (\log N)^{-\xi}, \tag{5.101}
$$

(5.78) and (5.79), we get $\Lambda_o + \Lambda_d \le C(\log N)^{-\xi}$, which implies the assumption (5.71) of Lemma 5.11.

Throughout the following we set $u := |\mathbb{U}| + 1$. We begin by estimating the contribution of A_1. Observe that if $i \neq k, l$, $i \in \mathbb{A}$ and $i \notin \mathbb{B}$, then $[G_{kl}]^{\mathbb{A},\mathbb{B}}$ is independent of the ith row and column of H. (The same argument will be repeatedly used in the rest of the proof below.) Thus, we have

$$\|\mathbf{1}([\Xi]_i)[G_{kl}]^{(\mathbb{S}i)\setminus\{k,l\},\mathbb{U}\setminus\{k,l\}}\|_\infty = \|\mathbf{1}(\Xi)[G_{kl}]^{(\mathbb{S}i)\setminus\{k,l\},\mathbb{U}\setminus\{k,l\}}\|_\infty \le (C|\mathbb{U}|X)^{|\mathbb{U}|-1},$$

where in the second step we used (5.73) and $\Lambda_o \le CX$ on Ξ. Thus, we find, using $|\mathbb{U}| \le |\mathbb{S}| \le p = (\log N)^{\xi-3/2}$ and $q^{-1} \le X$, that

$$\begin{aligned}\mathbb{E}\mathbf{1}([\Xi]_i)|A_1|^r &\le (\log N)^{2\xi} \max_{i,k,l} \mathbb{E}|h_{ik}|^r|h_{li}|^r\big((C|\mathbb{U}|X)^{|\mathbb{U}|-1}\big)^r \\ &\le (\log N)^{2\xi} q^{-2r}\big((C|\mathbb{U}|X)^{|\mathbb{U}|-1}\big)^r \\ &\le (Y(CXu)^u)^r.\end{aligned} \tag{5.102}$$

In order to bound the contribution of A_2, we estimate, as above,

$$\|\mathbf{1}([\Xi]_i)[G_{kl}]^{(\mathbb{S}i)\setminus\{k\},\mathbb{U}\setminus\{k\}}\|_\infty = \|\mathbf{1}(\Xi)[G_{kl}]^{(\mathbb{S}i)\setminus\{k\},\mathbb{U}\setminus\{k\}}\|_\infty \le (C|\mathbb{U}|X)^{|\mathbb{U}|},$$

where in the last step we used (5.73) and $\Lambda_o \le CX$ on Ξ. Thus, we may apply the moment estimate (A.4) from the Appendix with

$$B_{kl} := \mathbf{1}(k \in \mathbb{U})\mathbf{1}(l \notin \mathbb{S} \cup \{i\})\mathbf{1}([\Xi]_i)[G_{kl}]^{(\mathbb{S}i)\setminus\{k\},\mathbb{U}\setminus\{k\}}.$$

This yields

$$\begin{aligned}\mathbb{E}\mathbf{1}([\Xi]_i)|A_2|^r &\le (Cr)^{2r}\bigg(\bigg(\frac{1}{q} + \bigg(\frac{1}{N^2}(\log N)^{\xi} N\bigg)^{1/2}\bigg)(C|\mathbb{U}|X)^{|\mathbb{U}|}\bigg)^r \\ &\le (Cr)^{2r}\big(X(C|\mathbb{U}|X)^{|\mathbb{U}|}\big)^r,\end{aligned}$$

where we used that $|\mathbb{U}| \le (\log N)^{\xi}$, that the B_{kl} defined above are independent of the randomness in the ith column of H, and that

$$\frac{1}{q} + \frac{(\log N)^{\xi/2}}{\sqrt{N}} \le \frac{1}{q} + (\log N)^{2\xi}\sqrt{\frac{\operatorname{Im} m_{\mathrm{sc}}}{N\eta}} \le X$$

as follows from $\operatorname{Im} m_{\mathrm{sc}} \ge \sqrt{\eta}$ and $\eta \le 3$. Thus, we get

$$\mathbb{E}\mathbf{1}([\Xi]_i)|A_2|^r \le (Y(CXu)^u)^r. \tag{5.103}$$

(Recall that $u = |\mathbb{U}| + 1$.)

Exchanging k and l in the above estimate of A_2, we obtain

$$\mathbb{E}\mathbf{1}([\Xi]_i)|A_3|^r \le (Y(CXu)^u)^r. \tag{5.104}$$

Finally, we estimate the contribution of A_4. As above, we estimate

$$\|\mathbf{1}([\Xi]_i)[G_{kl}]^{(\mathbb{S}i),\mathbb{U}}\|_\infty = \|\mathbf{1}(\Xi)[G_{kl}]^{(\mathbb{S}i),\mathbb{U}}\|_\infty \le (C|\mathbb{U}|X)^{|\mathbb{U}|+1}$$

by (5.73) and $\Lambda_o \leq CX$ on Ξ. We may now apply the moment estimate (A.4) the Appendix with

$$B_{kl} := \mathbf{1}(k, l \notin \mathbb{S} \cup \{i\}) \mathbf{1}([\Xi]_i) [G_{kl}]^{(\mathbb{S}i),\mathbb{U}}.$$

This yields

$$\mathbb{E}\mathbf{1}([\Xi]_i)|A_2|^r \leq \bigl(r^2 (C|\mathbb{U}|X)^{|\mathbb{U}|+1}\bigr)^r,$$

where we used that the B_{kl} are independent of the randomness in the ith column of H. This gives

$$\mathbb{E}\mathbf{1}([\Xi]_i)|A_2|^r \leq (Y(CXu)^u)^r. \tag{5.105}$$

Combining (5.102), (5.103), (5.104) and (5.105), we obtain (5.96). This completes the proof. □

6. The largest eigenvalue of A.

6.1. *Eigenvalue interlacing.* We now concentrate on the spectrum of A. We begin by proving the following interlacing property. Recall that $\lambda_1 \leq \cdots \leq \lambda_N$ denote the eigenvalues of H and $\mu_1 \leq \cdots \leq \mu_N =: \mu_{\max}$ the eigenvalues of A. The associated eigenvectors of H are denoted by $\mathbf{u}_1, \ldots, \mathbf{u}_N$, and those of A by $\mathbf{v}_1, \ldots, \mathbf{v}_N =: \mathbf{v}_{\max}$. Also, we set $\widetilde{G}(z) := (A - z)^{-1}$.

LEMMA 6.1. *The eigenvalues of H and A are interlaced*,

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \cdots \leq \mu_{N-1} \leq \lambda_N \leq \mu_N. \tag{6.1}$$

PROOF. We use the identity

$$\langle \mathbf{e}, \widetilde{G}\mathbf{e}\rangle^{-1} = f + \langle \mathbf{e}, G\mathbf{e}\rangle^{-1}, \tag{6.2}$$

which follows by taking $\langle \mathbf{e}, \cdot \mathbf{e}\rangle$ in

$$\widetilde{G}(z)(A - z)G(z) = \widetilde{G}(z)(H - z)G(z) + f\widetilde{G}(z)|\mathbf{e}\rangle\langle\mathbf{e}|G(z).$$

From (6.2) we get

$$\left(\sum_\alpha \frac{|\langle \mathbf{v}_\alpha, \mathbf{e}\rangle|^2}{\mu_\alpha - z}\right)^{-1} = f + \left(\sum_\alpha \frac{|\langle \mathbf{u}_\alpha, \mathbf{e}\rangle|^2}{\lambda_\alpha - z}\right)^{-1}. \tag{6.3}$$

It is easy to see that the left-hand side of (6.3) defines a function of $z \in \mathbb{R}$ with $N - 1$ singularities and N zeros, which is smooth and decreasing away from the singularities. Moreover, its zeros are the eigenvalues of A. The interlacing property now follows from the fact that z is an eigenvalue of H if and only if the right-hand side of (6.3) is equal to f. □

6.2. *The laws of* $\mu_{\max}$ *and* $v_{\max}$. In this section we establish the basic properties of $\mu_{\max}$ and $\mathbf{v}_{\max}$. We make the assumption that $f \geq 1 + \varepsilon_0$ uniformly in N [see (6.4) below], which is necessary to guarantee that μ_{N-1} and μ_N are separated by a gap of order one. Note that in [26] it was proved, in the case where H is a Hermitian Wigner matrix, that if $f \leq 1$ no such gap exists.

The following result collects the main properties of $\mu_{\max}$ and $\mathbf{v}_{\max}$ for the rank-one perturbation $A = H + f|\mathbf{e}\rangle\langle\mathbf{e}|$ of the sparse matrix H. The most important technical result is (6.9). It states that, for large f, the eigenvector $\mathbf{v}_{\max}$ is almost parallel to the perturbation $\mathbf{e}$. Consequently, $\mathbf{e}$ is almost orthogonal to the eigenvectors $\mathbf{v}_\alpha$ for $\alpha = 1, \ldots, N-1$ (Corollary 6.7). As it turns out, this near orthogonality is the key input for establishing the local semicircle law for A in Section 7. We refer to the discussion at the beginning of Section 7.1 for more details on the use of Corollary 6.7.

THEOREM 6.2. *Suppose that A satisfies Definition 2.2 and that in addition to (2.8) we have the lower bound*

$$f \geq 1 + \varepsilon_0 \tag{6.4}$$

for some constant $\varepsilon_0 > 0$.

Then we have with (ξ, ν)-high probability

$$\mu_{\max} = f + \frac{1}{f} + o(1). \tag{6.5}$$

In particular, there is a constant c, depending on ε_0, such that with (ξ, ν)-high probability we have

$$\mu_{\max} \geq 2 + c. \tag{6.6}$$

Also, we have

$$\mathbb{E}\mu_{\max} = f + \frac{1}{f} + O\left(\frac{1}{f^3} + \frac{1}{f^2 q} + \frac{1}{fN}\right) \tag{6.7}$$

as well as, with (ξ, ν)-high probability,

$$\mu_{\max} = f + \frac{1}{f} + O\left(\frac{1}{f^3} + \frac{1}{f^2 q} + \frac{(\log N)^\xi}{\sqrt{N}}\right). \tag{6.8}$$

Note that (6.7) and (6.8) locate $\mu_{\max}$ more precisely than (6.5) in the large-f regime.

Moreover, the phase of $\mathbf{v}_{\max}$ can be chosen so that we have with (ξ, ν)-high probability

$$\langle\mathbf{v}_{\max}, \mathbf{e}\rangle = 1 - \frac{1}{2f^2} + O\left(\frac{1}{f^3} + \frac{(\log N)^{2\xi}}{f\sqrt{N}}\right). \tag{6.9}$$

Finally, there is a constant C_0 *such that if*

$$f \geq C_0 (\log N)^{2\xi} \quad \textit{and} \quad \xi \geq 2, \tag{6.10}$$

then we have with (ξ, ν)*-high probability*

$$\mu_{\max} = \mathbb{E}\mu_{\max} + \frac{1}{N} \sum_{i,j} h_{ij} + O\left(\frac{(\log N)^{2\xi}}{f\sqrt{N}} \right). \tag{6.11}$$

In particular, if (6.10) *holds, we have* (*by the central limit theorem*)

$$\sqrt{\frac{N}{2}} (\mu_{\max} - \mathbb{E}\mu_{\max}) \longrightarrow \mathcal{N}(0, 1) \tag{6.12}$$

in distribution, where $\mathcal{N}(0, 1)$ *denotes a standard normal random variable.*

REMARK 6.3. In analogy to Definition 3.3, we define $A^{(\mathbb{T})}$ as the $(N - |\mathbb{T}|) \times (N - |\mathbb{T}|)$ minor of A obtained by removing all columns of A indexed by $i \in \mathbb{T}$; here $\mathbb{T} \subset \{1, \dots, N\}$. If A satisfies Definition 2.2, then so does $(N/(N - |\mathbb{T}|))^{1/2} A^{(\mathbb{T})}$. Therefore, all results of this section also hold for $A^{(\mathbb{T})}$ provided $|\mathbb{T}| \leq 10$. (Here 10 can be any fixed number.) Throughout Sections 6 and 7 we abbreviate $\mu_{\max}^{(\mathbb{T})} := \mu_{N-|\mathbb{T}|}^{(\mathbb{T})}$ and $\mathbf{v}_{\max}^{(\mathbb{T})} := \mathbf{v}_{N-|\mathbb{T}|}^{(\mathbb{T})}$.

REMARK 6.4. Statistical properties of the k largest eigenvalues of a random Wigner matrix with a large rank-k perturbation have been studied in [3, 5, 26, 32]. Theorem 6.2 collects analogous results for the more singular case of sparse matrices. We restrict our attention to the special case where the perturbation is $f|\mathbf{e}\rangle\langle\mathbf{e}|$. (Note that in [3, 5, 32] the authors allow quite general finite-rank perturbations of Wigner matrices.)

The rest of this section is devoted to the proof of Theorem 6.2. It is based on the following standard observation. Let μ be an eigenvalue of A with associated normalized eigenvector $\mathbf{v}$. This means that

$$(\mu - H)\mathbf{v} = f\langle\mathbf{e}, \mathbf{v}\rangle\mathbf{e}.$$

Suppose now that μ is not an eigenvalue of H. Thus, we can choose $\mathbf{v}$ and $K > 0$ such that

$$\mathbf{v} = K(\mu - H)^{-1}\mathbf{e}, \tag{6.13}$$

$$1 = f\langle\mathbf{e}, (\mu - H)^{-1}\mathbf{e}\rangle. \tag{6.14}$$

Using the spectral decomposition of H, we rewrite (6.14) as

$$\frac{1}{f} = \sum_{\alpha} \frac{|\langle\mathbf{e}, \mathbf{u}_\alpha\rangle|^2}{\mu - \lambda_\alpha}. \tag{6.15}$$

It is easy to see that (6.15) has a unique solution, $\mu_{\max}$, greater than λ_N. Moreover, (6.15) readily yields $\mu_{\max} - \lambda_N \le f \le \mu_{\max} - \lambda_1$, that is,

$$\mu_{\max} \in [f + \lambda_1, f + \lambda_N]. \tag{6.16}$$

Our proof is based on the series expansions

$$\mu_{\max} = f \sum_{k \ge 0} \langle \mathbf{e}, (H/\mu_{\max})^k \mathbf{e} \rangle, \tag{6.17}$$

$$\mathbf{v}_{\max} = K \sum_{k \ge 0} (H/\mu_{\max})^k \mathbf{e}. \tag{6.18}$$

Note that the expansions (6.17) and (6.18) can be interpreted as perturbative corrections around the matrix $f|\mathbf{e}\rangle\langle\mathbf{e}|$.

In order to control the expansions (6.17) and (6.18), we shall need the following large deviation bound, proved in the Appendix.

LEMMA 6.5. *Let* $1 \le k \le \log N$. *Then*

$$|\langle \mathbf{e}, H^k \mathbf{e} \rangle - \mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle| \le C \frac{(\log N)^{k\xi}}{N^{1/2}} \tag{6.19}$$

with (ξ, ν)-*high probability provided that* $1 \le q \le C N^{1/2}$.

PROOF OF (6.5). The key observation is that

$$\left| \mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle - \int x^k \varrho_{\mathrm{sc}}(x)\,\mathrm{d}x \right| \le \frac{C(k)}{q} \tag{6.20}$$

for some constant $C(k)$ depending on k. Indeed, a standard application of the moment method (see, e.g., [28], Section 1.2) shows that $\mathbb{E}\langle \mathbf{e}, H^{2n} \mathbf{e} \rangle = C_n + O_n(q^{-2})$, where $C_n := \frac{1}{n+1}\binom{2n}{n} = \int x^{2n} \varrho_{\mathrm{sc}}(x)\,\mathrm{d}x$ is the nth Catalan number. If k is odd, one finds by a similar moment estimate that $\mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle = O_k(q^{-1})$. We omit the details.

For the following we work on the event of (ξ, ν)-high probability on which (4.29) holds. We consider solutions μ of (6.14) in the interval $I := [2 + \varepsilon_0^2/20, \infty)$. By monotonicity of the right-hand side of (6.15) in I, we know that (6.14) has at most one solution in I. For any $k_0 \in \mathbb{N}$, using (6.19) and (6.20) we may expand (6.14) in I [see (6.17)] as

$$\begin{aligned} \mu = {} & f \sum_{k=0}^{k_0} \int \left(\frac{x}{\mu} \right)^k \varrho_{\mathrm{sc}}(x)\,\mathrm{d}x \\ & + O\left(f \sum_{k > k_0} \left(\frac{\|H\|}{\mu} \right)^k + \frac{C(k_0)}{q} + \frac{k_0 (\log N)^{\xi k_0}}{\sqrt{N}} \right) \end{aligned}$$

$$= -f\mu m_{\mathrm{sc}}(\mu) + O\biggl(f\sum_{k>k_0}\biggl(\frac{\|H\|}{2+\varepsilon_0^2/20}\biggr)^k + f\sum_{k>k_0}\biggl(\frac{2}{2+\varepsilon_0^2/20}\biggr)^k + \frac{C(k_0)}{q} + \frac{k_0(\log N)^{\xi k_0}}{\sqrt{N}}\biggr),$$

where the first term comes from extending the sum over k to infinity and using that

$$\sum_{k=0}^{\infty}\int\biggl(\frac{x}{\mu}\biggr)^k \varrho_{\mathrm{sc}}(x)\,\mathrm{d}x = \mu\int\frac{\varrho_{\mathrm{sc}}(x)\,\mathrm{d}x}{\mu - x}$$

for $\mu > 2$. It is easy to see that the second term is $o(1)$ by an appropriate choice of $k_0(N)$. Thus, we have proved that, for $\mu \in I$, the equation (6.14) reads $m_{\mathrm{sc}}(\mu) = -f^{-1} + r(\mu)$, where $r(\mu) \to 0$ as $N \to \infty$ uniformly in μ.

Next, the function $\mu \mapsto m_{\mathrm{sc}}(\mu)$ is continuous and monotone increasing on $(2,\infty)$, with range $(-1,0)$. Let $\bar{\mu}$ be the unique solution of $m_{\mathrm{sc}}(\bar{\mu}) = -f^{-1}$. (Note that here we need the assumption $f > 1$.) Using (2.13), we find that $\bar{\mu} = f + f^{-1} \geq 2 + \varepsilon_0^2/10$. We therefore find that, for N large enough, the equation $m_{\mathrm{sc}}(\mu) = -f^{-1} + r(\mu)$ [which is equivalent to (6.14) on I] has a unique solution $\mu \in I$ which satisfies $\mu = \bar{\mu} + o(1)$. Since μ is the only solution of (6.14) in I, we must have $\mu = \mu_{\max}$. □

Note that (6.5) remains valid if $\mathbf{e}$ in (2.7) is replaced with any ℓ^2-normalized vector. It is a simple matter to check that (6.20) is valid for arbitrary vectors $\mathbf{e}$. Moreover, Lemma 6.5 remains correct for arbitrary $\mathbf{e}$ provided one replaces $N^{-1/2}$ on the right-hand side of (6.19) with q^{-1}. We omit the details, as we shall not need this result.

From (6.6) and (4.29) we find that with (ξ,ν)-high probability

$$\frac{\|H\|}{\mu_{\max}} \leq 1 - c \tag{6.21}$$

for some constant $c > 0$. In particular, (6.17) and (6.18) converge with (ξ,ν)-high probability.

PROOF OF (6.8). From (6.17) and Lemma 4.3 we find $\mu_{\max} = f(1 + r(f))$ with (ξ,ν)-high probability, where $\lim_{f\to\infty} r(f) = 0$. Together with the simple identities

$$\mathbb{E}\langle\mathbf{e}, H\mathbf{e}\rangle = 0, \qquad \mathbb{E}\langle\mathbf{e}, H^2\mathbf{e}\rangle = 1, \tag{6.22}$$

(6.17), (6.21) and Lemma 6.5 yield with (ξ,ν)-high probability

$$\mu_{\max} = f + \frac{1}{f} + \frac{\mathbb{E}\langle\mathbf{e}, H^3\mathbf{e}\rangle}{f^2} + O\biggl(\frac{1}{f^3}\biggr) + O\biggl(\frac{(\log N)^{\xi}}{\sqrt{N}}\biggr). \tag{6.23}$$

By explicit computation we find that

$$\mathbb{E}\langle \mathbf{e}, H^3\mathbf{e}\rangle = O(q^{-1}). \tag{6.24}$$

Thus, (6.8) follows. □

PROOF OF (6.7). From (6.17) and (6.21) we get with (ξ, ν)-high probability

$$\mu_{\max} = f + \frac{f}{\mu_{\max}}\langle \mathbf{e}, H\mathbf{e}\rangle + \frac{f}{\mu_{\max}^2}\langle \mathbf{e}, H^2\mathbf{e}\rangle + O\left(\frac{1}{f^3} + \frac{1}{f^2 q}\right), \tag{6.25}$$

where we used (6.24). Iterating (6.25) yields with (ξ, ν)-high probability

$$\begin{aligned} \mu_{\max} &= f + \langle \mathbf{e}, H\mathbf{e}\rangle - \langle \mathbf{e}, H\mathbf{e}\rangle^2/f + \langle \mathbf{e}, H\mathbf{e}\rangle^3/f^2 - \langle \mathbf{e}, H\mathbf{e}\rangle\langle \mathbf{e}, H^2\mathbf{e}\rangle/f^2 \\ &\quad + \langle \mathbf{e}, H^2\mathbf{e}\rangle/f - 2\langle \mathbf{e}, H\mathbf{e}\rangle\langle \mathbf{e}, H^2\mathbf{e}\rangle/f^2 + O\left(\frac{1}{f^3} + \frac{1}{f^2 q}\right), \end{aligned} \tag{6.26}$$

where we used Lemma 4.3. In order to complete the proof of (6.7), we use the rough estimate $\mathbb{E}\mu_{\max}^2 \le \mathbb{E}\operatorname{Tr} A^2 \le CNf^2 + N \le N^C$, by (2.8). Recalling (6.19), we also get

$$|\mathbb{E}\langle \mathbf{e}, H\mathbf{e}\rangle^2| \le \frac{C}{N}, \qquad |\mathbb{E}\langle \mathbf{e}, H\mathbf{e}\rangle^3| \le \frac{C}{Nq}, \qquad |\mathbb{E}\langle \mathbf{e}, H\mathbf{e}\rangle\langle \mathbf{e}, H^2\mathbf{e}\rangle| \le \frac{C}{Nq}$$

by explicit calculation using (2.5). Now taking the expectation in (6.26), using (6.22) yields (6.7). □

PROOF OF (6.9). We compute the normalization constant K in (6.18) from

$$\begin{aligned} K^{-2} &= \sum_{k,k' \ge 0} \mu_{\max}^{-k-k'}\langle \mathbf{e}, H^{k+k'}\mathbf{e}\rangle \\ &= 1 + 2\mu_{\max}^{-1}\langle \mathbf{e}, H\mathbf{e}\rangle + 3\mu_{\max}^{-2}\langle \mathbf{e}, H^2\mathbf{e}\rangle + O(\mu_{\max}^{-3}) \\ &= 1 + \frac{3}{f^2} + O\left(\frac{1}{f^3} + \frac{(\log N)^{2\xi}}{f\sqrt{N}}\right) \end{aligned} \tag{6.27}$$

with (ξ, ν)-high probability, where we used Lemmas 6.5 and 4.3, as well as (6.8) and (6.21). Now (6.9) is an easy consequence of (6.18), (6.8) and Lemmas 6.5 and 4.3. □

What remains is to prove (6.11).

PROOF OF (6.11). We assume (6.10), and, in particular, $\mu_{\max} \ge \frac{C_0}{2}(\log N)^{2\xi}$ by (6.5). Thus, from Lemma 4.3 we get with (ξ, ν)-high probability

$$\frac{\|H\|}{\mu_{\max}} \le \frac{6}{C_0(\log N)^{2\xi}}.$$

From (6.17) and (6.8) we therefore get with (ξ, ν)-high probability

$$\mu_{\max} = f \sum_{k=0}^{c_0 \log N} \frac{\langle \mathbf{e}, H^k \mathbf{e} \rangle}{\mu_{\max}^k} + O(\mathrm{e}^{-c_0 \log N \log\log N}),$$

where $c_0 \le 1$ is a positive constant to be chosen later. Thus, we find with (ξ, ν)-high probability

$$\mu_{\max} = f \sum_{k=0}^{c_0 \log N} \frac{\mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle}{\mu_{\max}^k} + \frac{f}{\mu_{\max}} \langle \mathbf{e}, H\mathbf{e} \rangle + f \sum_{k=2}^{c_0 \log N} \frac{\langle \mathbf{e}, H^k \mathbf{e} \rangle - \mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle}{\mu_{\max}^k} + O(\mathrm{e}^{-c_0 \log N \log\log N}).$$

Therefore, we get, for any $0 < c_0 \le 1$ and using (6.8) and Lemma 6.5, that with (ξ, ν)-high probability we have

$$\mu_{\max} = f \sum_{k=0}^{c_0 \log N} \frac{\mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle}{\mu_{\max}^k} + \frac{f}{\mu_{\max}} \langle \mathbf{e}, H\mathbf{e} \rangle + O\left(\frac{(\log N)^{2\xi}}{f\sqrt{N}} \right).$$

Here the constant in $O(\cdot)$ depends on c_0.

Next, Lemma 4.3 yields

$$|\mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle| \le (5/2)^k + N^{Ck} \mathrm{e}^{-\nu(\log N)^\xi} \le 3^k \tag{6.28}$$

for $k \le (\nu/C)(\log N)^{\xi - 1}$. [Here we used Schwarz's inequality and the trivial estimate $\mathbb{E}\langle \mathbf{e}, H\mathbf{e} \rangle \le N^C$ to estimate the contribution of the low-probability event on which (4.29) does not hold.] By the assumption (6.10) on ξ, (6.28) holds for $k \le c_0 \log N$ for c_0 small enough. It is therefore easy to see that the equation

$$\bar{\mu} = f \sum_{k=0}^{c_0 \log N} \frac{\mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle}{\bar{\mu}^k}$$

has a unique solution $\bar{\mu} > 0$, which satisfies $\bar{\mu} = f + O(f^{-1})$. Writing $\mu_{\max} = \bar{\mu} + \zeta$, we get with (ξ, ν)-high probability

$$\zeta = \frac{f}{\mu_{\max}} \langle \mathbf{e}, H\mathbf{e} \rangle + f \sum_{k=0}^{c_0 \log N} \frac{\mathbb{E}\langle \mathbf{e}, H^k \mathbf{e} \rangle}{\bar{\mu}^k} \left[\left(1 + \frac{\zeta}{\bar{\mu}} \right)^{-k} - 1 \right] + O\left(\frac{(\log N)^{2\xi}}{f\sqrt{N}} \right). \tag{6.29}$$

Next, by (6.8) we find $\zeta = O(f^{-1})$ with (ξ, ν)-high probability. Moreover, (6.22) and Lemma 6.5 imply that $\langle \mathbf{e}, H\mathbf{e} \rangle = O((\log N)^\xi N^{-1/2})$ with (ξ, ν)-high probability, and that the sum in (6.29) starts at $k = 2$. This yields the expression with (ξ, ν)-high probability

$$\zeta = \langle \mathbf{e}, H\mathbf{e} \rangle + \frac{1}{f} \sum_{l \ge 1} a_l \zeta^l + O\left(\frac{(\log N)^{2\xi}}{f\sqrt{N}} \right)$$

for some coefficients $a_l = O(1)$, by (6.28). We conclude that with (ξ,ν)-high probability we have

$$\zeta = \langle \mathbf{e}, H\mathbf{e}\rangle\bigl(1+O(f^{-1})\bigr) + O\biggl(\frac{(\log N)^{2\xi}}{f\sqrt{N}}\biggr) = \langle \mathbf{e}, H\mathbf{e}\rangle + O\biggl(\frac{(\log N)^{2\xi}}{f\sqrt{N}}\biggr),$$

where we used that $|\langle \mathbf{e}, H\mathbf{e}\rangle| \le (\log N)^{\xi} N^{-1/2}$ with (ξ,ν)-high probability.

Summarizing, we have proved that with (ξ,ν)-high probability we have

$$\mu_{\max} = \bar{\mu} + \frac{1}{N}\sum_{i,j} h_{ij} + R, \tag{6.30}$$

where $|R| \le O(\frac{(\log N)^{2\xi}}{f\sqrt{N}})$. Using $\mathbb{E}\mu_{\max}^2 \le N^C$, we therefore get

$$\mathbb{E}|R| \le O\biggl(\frac{(\log N)^{2\xi}}{f\sqrt{N}}\biggr),$$

and (6.11) follows by taking the expectation in (6.30). □

This concludes the proof of Theorem 6.2.

For future reference, we record two simple corollaries which we shall use in Section 7 to control the matrix elements of $\widetilde{G}$.

COROLLARY 6.6. *Suppose that A satisfies Definition 2.2. Then we have with (ξ,ν)-high probability*

$$|\mu_\alpha| \le \max_\beta |\lambda_\beta| = \|H\| \le 2 + (\log N)^{\xi} q^{-1/2} \qquad \textit{for } \alpha = 1,\dots,N-1. \tag{6.31}$$

PROOF. Use (6.1) and Lemma 4.3. □

COROLLARY 6.7. *Suppose that A satisfies Definition 2.2 and that, in addition, $f \le C_0 N^{1/2}$. Then we have with (ξ,ν)-high probability*

$$\sum_{\alpha \ne N} |\langle \mathbf{v}_\alpha, \mathbf{e}\rangle|^2 = O(f^{-2}). \tag{6.32}$$

PROOF. The statement is trivial unless $f \ge 1 + \varepsilon_0$, in which case we use (6.9) and $\|\mathbf{e}\| = 1$. □

7. Control of $\widetilde{G}$: Proofs of Theorems 2.9 and 2.16. In this section we adopt the convention that if $F = F(H)$ is any function of H, then $F(A)$ is denoted by $\widetilde{F}$, that is, we use the tilde $\widetilde{(\cdot)}$ to indicate quantities defined in terms of $A = H + f|\mathbf{e}\rangle\langle\mathbf{e}|$. Thus, for example, we have

$$A = \widetilde{H}, \qquad \mu_\alpha = \widetilde{\lambda}_\alpha, \qquad \mathbf{v}_\alpha = \widetilde{\mathbf{u}}_\alpha,$$

$$\widetilde{G}(z) := (A - z)^{-1}, \qquad \widetilde{m}(z) := \frac{1}{N}\sum \widetilde{G}_{ii}(z)$$

and

$$\widetilde{\Lambda}_o := \max_{i \neq j} |\widetilde{G}_{ij}|, \qquad \widetilde{\Lambda}_d := \max_i |\widetilde{G}_{ii} - m_{\mathrm{sc}}|, \tag{7.1}$$
$$\widetilde{\Lambda} := |\widetilde{m} - m_{\mathrm{sc}}|, \qquad \widetilde{v}_i := \widetilde{G}_{ii} - m_{\mathrm{sc}}.$$

Note that $\widetilde{G}$ and $\widetilde{m}$ were already introduced in (2.19).

We begin by using the interlacing property (6.1) to derive a bound on Λ. Recall the convention that if $F = F(H)$ is any function of H, then $\widetilde{F}$ is defined as $F(A)$, where, we recall, $A = H + f|\mathbf{e}\rangle\langle\mathbf{e}|$.

LEMMA 7.1. *Let A satisfy Definition 2.2. Then for any $z \in D_L$ we have*

$$|\widetilde{\Lambda}(z) - \Lambda(z)| \leq \frac{\pi}{N\eta}.$$

PROOF. Define the empirical density $\widetilde{\varrho}(x) := \frac{1}{N}\sum_\alpha \delta(x - \mu_\alpha)$. Thus, the integrated empirical density defined in (2.23) can be written as $\widetilde{\mathfrak{n}}(E) = \int_{-\infty}^E \widetilde{\varrho}(x)\,\mathrm{d}x$. Similarly, define the quantities ϱ and $\mathfrak{n}$ in terms of the eigenvalues $\lambda_1, \dots, \lambda_N$ of H. Using integration by parts, we find

$$\widetilde{\Lambda}(z) - \Lambda(z) = \int \frac{\widetilde{\varrho}(x) - \varrho(x)}{x - z}\,\mathrm{d}x = -\int \frac{\widetilde{\mathfrak{n}}(x) - \mathfrak{n}(x)}{(x - z)^2}\,\mathrm{d}x.$$

By (6.1) we have $|\widetilde{\mathfrak{n}}(x) - \mathfrak{n}(x)| \leq N^{-1}$ for all x. Thus, we find

$$|\widetilde{\Lambda}(z) - \Lambda(z)| \leq \frac{1}{N}\int \frac{1}{|x - z|^2}\,\mathrm{d}x = \frac{\pi}{N\eta}. \qquad \square$$

We note that the claim (2.20) of Theorem 2.9 is now an immediate consequence of Lemma 7.1 and the strong local semicircle law (2.16) for H.

The rest of this section is devoted to the proof of the estimate (2.22) for the matrix elements of $\widetilde{G}$. From now on we consistently assume the upper bound (2.21) on f.

7.1. *Basic estimates on the good events.* In this section we control the individual matrix elements $\widetilde{G}_{ij}$ in terms of $\widetilde{\Lambda}$, which in turn will be estimated using Lemma 7.1. Our basic strategy is similar to that of Section 3, but, owing to the nonvanishing expectation of a_{ij}, the self-consistent equation for $\widetilde{G}_{ii}$ has several additional error terms as compared to Lemma 3.10; see Lemma 7.2 and Proposition 7.6 below. The most dangerous of these error terms is estimated in Lemma 7.5 below. We will use the spectral decomposition of $\widetilde{H}$, combined with bounds on $\langle\mathbf{e}, \mathbf{v}_\alpha\rangle$ and $\|\mathbf{v}_\alpha\|_\infty$. The former quantities are estimated using Corollary 6.7, while the latter are estimated by bootstrapping. The spectral decomposition requires simultaneous control of all eigenvectors, whose associated eigenvalues are distributed throughout the spectrum. Since bounds on $\|\mathbf{v}_\alpha\|_\infty$ (delocalization bounds) may be derived from a priori bounds on $\widetilde{\Lambda}_d(z)$ for $\operatorname{Re} z$ being near the corresponding eigenvalue, we will therefore need bounds on $\widetilde{\Lambda}_d(z)$ that are uniform for all $z \in D_L$ with a fixed imaginary part. Hence, the bootstrapping now occurs simultaneously for all $E \in [-\Sigma, \Sigma]$ (see Definition 7.3 below).

We use the following self-consistent equation for $\widetilde{G}$, whose proof is an elementary calculation using (3.12) and (3.13) applied to $\widetilde{G}$; see also Lemma 3.10.

LEMMA 7.2. *We have the identity*

$$\widetilde{G}_{ii} = \frac{1}{-z - m_{\mathrm{sc}} - ([\widetilde{v}] - \widetilde{\Upsilon}_i)}, \tag{7.2}$$

where

$$\widetilde{\Upsilon}_i := h_{ii} - \widetilde{Z}_i + \widetilde{\mathcal{A}}_i$$

and

$$\begin{aligned} \widetilde{Z}_i &:= \mathbb{IE}_i \sum_{k,l}^{(i)} a_{ik} \widetilde{G}_{kl}^{(i)} a_{li}, \\ \widetilde{\mathcal{A}}_i &:= \frac{f}{N} - \frac{f^2}{N} \frac{N-1}{N} \langle \mathbf{e}, \widetilde{G}^{(i)} \mathbf{e} \rangle + \frac{1}{N} \sum_j \frac{\widetilde{G}_{ij} \widetilde{G}_{ji}}{\widetilde{G}_{ii}}. \end{aligned} \tag{7.3}$$

Recall that in expressions such as (7.3) the vector $\mathbf{e}$ stands for $\mathbf{e}_{N-1}$; see (2.2).

DEFINITION 7.3. For $N^{-1}(\log N)^L \le \eta \le 3$ introduce the set $D(\eta) := \{z \in D_L : \operatorname{Im} z = \eta\}$. We define the event

$$\widetilde{\Omega}(\eta) := \Big\{ \sup_{z \in D(\eta)} \big(\widetilde{\Lambda}_d(z) + \widetilde{\Lambda}_o(z) \big) \le (\log N)^{-\xi} \Big\}. \tag{7.4}$$

Recall the definition of $A^{(\mathbb{T})}$ from Remark 6.3. Similarly to Lemma 3.12, we have the following result for the matrix A.

LEMMA 7.4. *Fixing $z = E + i\eta \in D_L$, we have for any i and $\mathbb{T} \subset \{1, \dots, N\}$ satisfying $i \notin \mathbb{T}$ and $|\mathbb{T}| \le 10$ that*

$$\widetilde{m}^{(i\mathbb{T})}(z) = \widetilde{m}^{(\mathbb{T})}(z) + O\left(\frac{1}{N\eta}\right) \tag{7.5}$$

holds in $\widetilde{\Omega}(\eta)$.

The following lemma is crucial in dealing with error terms arising from the nonvanishing expectation of a_{ij}. Recall that, when indexing the eigenvalues and eigenvectors of $A^{(\mathbb{T})}$, we defined $\alpha_{\max} := N - |\mathbb{T}|$.

LEMMA 7.5. *Fixing $z = E + \mathrm{i}\eta \in D_L$, we have for any $\mathbb{T} \subset \{1, \ldots, N\}$ satisfying $|\mathbb{T}| \le 10$ and for any $i \in \mathbb{T}$ that*

$$\left|\sum_{k,l}^{(\mathbb{T})} \frac{f}{N} \widetilde{G}_{kl}^{(\mathbb{T})} h_{li}\right| \le C(\log N)^{\xi}\left(\frac{1}{q} + \sqrt{\frac{\operatorname{Im}\widetilde{m}}{N\eta}} + \frac{1}{N\eta}\right) \tag{7.6}$$

on $\widetilde{\Omega}(\eta)$ with (ξ, ν)-high probability.

PROOF. For technical reasons, it is convenient to avoid the situation where $\mu_{\max}$ is close to Σ. In order to ensure this, we may if necessary increase Σ slightly and hence assume that $f \le \Sigma - 3$ or $f \ge \Sigma + 3$. We start by proving the following delocalization bound. Define

$$R := \max_{|\mathbb{T}| \le 10} \max_{\alpha \ne \alpha_{\max}} \max_j |v_\alpha^{(\mathbb{T})}(j)|, \qquad R_{\max} := \max_{|\mathbb{T}| \le 10} \max_j |v_{\max}^{(\mathbb{T})}(j)|. \tag{7.7}$$

First we claim that on $\widetilde{\Omega}(\eta)$ we have with (ξ, ν)-high probability

$$R \le C\sqrt{\eta} \tag{7.8}$$

and, assuming $f \le \Sigma - 3$, we have with (ξ, ν)-high probability

$$R_{\max} \le C\sqrt{\eta}. \tag{7.9}$$

In order to prove (7.8) and (7.9), we note that on $\widetilde{\Omega}(\eta)$ we have, in analogy to (3.29),

$$c \le |\widetilde{G}_{jj}^{(\mathbb{T})}(z)| \le C \tag{7.10}$$

for all $z \in D_L$ such that $\operatorname{Im} z = \eta$ and N large enough. From (6.31) we find that $z := \mu_\alpha^{(\mathbb{T})} + \mathrm{i}\eta \in D_L$ with (ξ, ν)-high probability for $\alpha \ne \alpha_{\max}$; see Remark 6.3. Thus, we get with (ξ, ν)-high probability

$$C \ge \operatorname{Im} \widetilde{G}_{jj}^{(\mathbb{T})}(\mu_\alpha^{(\mathbb{T})} + \mathrm{i}\eta) = \sum_\beta \frac{\eta |v_\beta^{(\mathbb{T})}(j)|^2}{(\mu_\beta^{(\mathbb{T})} - \mu_\alpha^{(\mathbb{T})})^2 + \eta^2} \ge \frac{|v_\alpha^{(\mathbb{T})}(j)|^2}{\eta}.$$

This concludes the proof of (7.8). Next, if $f \le \Sigma - 3$, then by (6.5) and Lemma 4.3 we have $\mu_{\max}^{(\mathbb{T})} \in [-\Sigma, \Sigma]$ with (ξ, ν)-high probability. Thus, we get (7.9) just like above.

Having established (7.8) and (7.9), we may now estimate the left-hand side of (7.6), using the spectral decomposition of $\widetilde{G}^{(\mathbb{T})}$, by

$$\begin{aligned} &\frac{f}{\sqrt{N - |\mathbb{T}|}} \left| \frac{\langle \mathbf{e}, \mathbf{v}_{\max}^{(\mathbb{T})} \rangle}{\mu_{\max}^{(\mathbb{T})} - z} \sum_l^{(\mathbb{T})} v_{\max}^{(\mathbb{T})}(l) h_{li} \right| \\ &\quad + \frac{f}{\sqrt{N - |\mathbb{T}|}} \left| \sum_{\alpha \ne \alpha_{\max}} \frac{\langle \mathbf{e}, \mathbf{v}_\alpha^{(\mathbb{T})} \rangle}{\mu_\alpha^{(\mathbb{T})} - z} \sum_l^{(\mathbb{T})} v_\alpha^{(\mathbb{T})}(l) h_{li} \right|. \end{aligned} \tag{7.11}$$

By the delocalization bound (7.8) and the large deviation estimate (3.19), we find for $\alpha \neq \alpha_{\max}$ on $\widetilde{\Omega}(\eta)$ with (ξ, ν)-high probability

$$\left| \sum_{l}^{(\mathbb{T})} v_{\alpha}^{(\mathbb{T})}(l) h_{li} \right| \leq (\log N)^{\xi} \left(\frac{R}{q} + \frac{1}{\sqrt{N}} \right).$$

Similarly, we have

$$\left| \sum_{l}^{(\mathbb{T})} v_{\max}^{(\mathbb{T})}(l) h_{li} \right| \leq (\log N)^{\xi} \left(\frac{R_{\max}}{q} + \frac{1}{\sqrt{N}} \right).$$

Next, we estimate the first term of (7.11). If $f \leq \Sigma - 3$, then $f \leq C$, and the first term of (7.11) is bounded, with (ξ, ν)-high probability, by

$$\frac{C}{\sqrt{N\eta}} (\log N)^{\xi} \left(\frac{\sqrt{\eta}}{q} + \frac{1}{\sqrt{N}} \right) \leq C (\log N)^{\xi} \left(\frac{1}{q} + \frac{1}{N\eta} \right).$$

If $f \geq \Sigma + 3$, then by (6.5) and (6.8) we get $|\mu_{\max}^{(\mathbb{T})} - z| \geq cf$ with (ξ, ν)-high probability. Thus, the first term of (7.11) is bounded with (ξ, ν)-high probability by

$$\frac{Cf}{\sqrt{N}} \frac{(\log N)^{\xi}}{f} \left(\frac{1}{q} + \frac{1}{\sqrt{N}} \right) \leq \frac{C (\log N)^{\xi}}{\sqrt{N} q},$$

where we used the trivial bound $R_{\max} \leq 1$. We therefore get that the left-hand side of (7.6) is bounded with (ξ, ν)-high probability by

$$\begin{aligned} &C (\log N)^{\xi} \left(\frac{1}{q} + \frac{1}{N\eta} \right) + C (\log N)^{\xi} \frac{f}{\sqrt{N}} \left(\frac{R}{q} + \frac{1}{\sqrt{N}} \right) \sum_{\alpha \neq \alpha_{\max}} \frac{|\langle \mathbf{e}, \mathbf{v}_{\alpha}^{(\mathbb{T})} \rangle|}{|\mu_{\alpha}^{(\mathbb{T})} - z|} \\ &\quad \leq C (\log N)^{\xi} \left(\frac{1}{q} + \frac{1}{N\eta} \right) \\ &\qquad + C (\log N)^{\xi} \frac{f}{\sqrt{N}} \left(\frac{R}{q} + \frac{1}{\sqrt{N}} \right) \left(\sum_{\alpha \neq \alpha_{\max}} |\langle \mathbf{e}, \mathbf{v}_{\alpha}^{(\mathbb{T})} \rangle|^2 \right)^{1/2} \\ &\qquad \times \left(\sum_{\alpha} \frac{1}{|\mu_{\alpha}^{(\mathbb{T})} - z|^2} \right)^{1/2}. \end{aligned}$$

By (6.32) this becomes

$$\begin{aligned} &C (\log N)^{\xi} \left(\frac{1}{q} + \frac{1}{N\eta} \right) \\ &\quad + C (\log N)^{\xi} \frac{1}{\sqrt{N}} \left(\frac{R}{q} + \frac{1}{\sqrt{N}} \right) \left(\sum_{\alpha} \frac{1}{|\mu_{\alpha}^{(\mathbb{T})} - z|^2} \right)^{1/2}. \end{aligned} \tag{7.12}$$

Using (7.5), we get

$$\sum_\alpha \frac{1}{|\mu_\alpha^{(\mathbb{T})} - z|^2} = \frac{1}{\eta} \operatorname{Im} \sum_\alpha \frac{1}{\mu_\alpha^{(\mathbb{T})} - z} = \frac{1}{\eta} \operatorname{Im} \operatorname{Tr} \widetilde{G}^{(\mathbb{T})} = \frac{N}{\eta} \operatorname{Im} \widetilde{m} + O\left(\frac{1}{\eta^2}\right),$$

which therefore yields the bound

$$C(\log N)^\xi \left(\frac{1}{q} + \frac{1}{N\eta}\right) + C(\log N)^\xi \left(\frac{\sqrt{\eta}}{\sqrt{N}q} + \frac{1}{N}\right)\left(\sqrt{\frac{N}{\eta} \operatorname{Im} \widetilde{m}} + \frac{1}{\eta}\right)$$

on $\widetilde{\Omega}(\eta)$ with (ξ, ν)-high probability. Here we used (7.8). The claim follows. □

For the following statements it is convenient to abbreviate

$$\Phi(z) := \frac{(\log N)^\xi}{q} + (\log N)^{2\xi} \left(\sqrt{\frac{\operatorname{Im} \widetilde{m}(z)}{N\eta}} + \frac{1}{N\eta}\right). \tag{7.13}$$

PROPOSITION 7.6. *Assume* (2.21). *Then for* $z = E + i\eta \in D_L$ *we have*

$$\widetilde{\Lambda}_o(z) \le C\Phi(z), \tag{7.14}$$

$$\max_i |\widetilde{Z}_i(z)| \le C\Phi(z), \tag{7.15}$$

$$\max_i |\widetilde{\mathcal{A}}_i(z)| \le \frac{C}{N\eta} \tag{7.16}$$

in $\widetilde{\Omega}(\eta)$ *with* (ξ, ν)*-high probability.*

PROOF. We start with (7.14). Let $i \neq j$. Using (3.13) for A instead of H, and writing $a_{ij} = f/N + h_{ij}$, we get with (ξ, ν)-high probability

$$\begin{aligned} C^{-1}|\widetilde{G}_{ij}| \le \frac{1}{q} &+ \left|\sum_{k,l}^{(ij)} h_{ik} \widetilde{G}_{kl}^{(ij)} h_{lj}\right| + \left|\sum_{k,l}^{(ij)} \frac{f}{N} \widetilde{G}_{kl}^{(ij)} h_{lj}\right| \\ &+ \left|\sum_{k,l}^{(ij)} h_{ik} \widetilde{G}_{kl}^{(ij)} \frac{f}{N}\right| + \left|\sum_{k,l}^{(ij)} \frac{f}{N} \widetilde{G}_{kl}^{(ij)} \frac{f}{N}\right| \end{aligned} \tag{7.17}$$

by Lemma 3.7 and (7.10).

The second term of (7.17) is bounded exactly as in (3.34) and (3.35); using (3.22) and (7.10), we estimate it by

$$(\log N)^\xi \frac{C}{q} + C(\log N)^{2\xi} \left(\sqrt{\frac{\operatorname{Im} \widetilde{m}}{N\eta}} + \frac{1}{N\eta}\right)$$

on $\widetilde{\Omega}(\eta)$ with (ξ, ν)-high probability.

The last term of (7.17) is bounded with (ξ,ν)-high probability by

$$\begin{aligned}\frac{f^2}{N^2}(N-2)\big|\langle \mathbf{e}, \widetilde{G}^{(ij)}\mathbf{e}\rangle\big| &\le C\frac{f^2}{N}\frac{|\langle \mathbf{e}, \mathbf{v}_{\max}^{(ij)}\rangle|^2}{|\mu_{\max}^{(ij)}-z|} + \frac{f^2}{N}\sum_{\alpha\neq\alpha_{\max}}\frac{|\langle \mathbf{e}, \mathbf{v}_{\alpha}^{(ij)}\rangle|^2}{|\mu_{\alpha}^{(ij)}-z|} \\ &\le \frac{Cf}{N}+\frac{C}{N\eta}+\frac{f^2}{N\eta}\sum_{\alpha\neq\alpha_{\max}}\big|\langle \mathbf{e}, \mathbf{v}_{\alpha}^{(ij)}\rangle\big|^2 \\ &\le \frac{Cf}{N}+\frac{C}{N\eta},\end{aligned} \tag{7.18}$$

where in the first step we used (6.8), and in the last step (6.32). Here we estimated the term arising from $\mu_{\max}^{(ij)}$ by $C(N\eta)^{-1}$ if $f\le 2\Sigma$, and by Cf/N if $f\ge 2\Sigma$.

Using Lemma 7.5, the third and fourth terms in (7.17) are bounded on $\widetilde{\Omega}(\eta)$ with (ξ,ν)-high probability by the right-hand side of (7.6). This concludes the proof of (7.14).

Next, we prove (7.15). By definition,

$$\widetilde{Z}_i = \sum_{k,l}^{(i)} h_{ik}\widetilde{G}_{kl}^{(i)}\frac{f}{N} + \sum_{k,l}^{(i)}\frac{f}{N}\widetilde{G}_{kl}^{(i)}h_{li} + \mathbb{IE}_i\sum_{k,l}^{(i)} h_{ik}\widetilde{G}_{kl}^{(i)}h_{li}. \tag{7.19}$$

The first two terms are bounded using Lemma 7.5, and the last one exactly as (3.33).

Finally, we prove (7.16). Using (6.8), (6.9), (6.32) and (7.10), we find on $\widetilde{\Omega}(\eta)$ with (ξ,ν)-high probability

$$\begin{aligned}\widetilde{\mathcal{A}}_i &= \frac{f}{N}-\frac{f^2}{N}\frac{N-1}{N}\frac{|\langle \mathbf{e}, \mathbf{v}_{\max}^{(i)}\rangle|^2}{\mu_{\max}^{(i)}-z} \\ &\quad -\frac{f^2}{N}\frac{N-1}{N}\sum_{\alpha\neq\alpha_{\max}}\frac{|\langle \mathbf{e}, \mathbf{v}_{\alpha}^{(i)}\rangle|^2}{\mu_{\alpha}^{(i)}-z}+\frac{1}{N}\sum_j\frac{\widetilde{G}_{ij}\widetilde{G}_{ji}}{\widetilde{G}_{ii}} \\ &= \frac{f}{N}-\frac{f^2}{N}\frac{N-1}{N}\frac{1}{f}\left[1+O\left(\frac{1}{f}\right)\right] \\ &\quad + O\left(\frac{1}{N\eta}\right)+O\left(\frac{f^2}{N\eta}\sum_{\alpha\neq\alpha_{\max}}\big|\langle \mathbf{e}, \mathbf{v}_{\alpha}^{(i)}\rangle\big|^2\right)+O\left(\frac{1}{N}\sum_j|\widetilde{G}_{ij}|^2\right) \\ &= O\left(\frac{1}{N\eta}\right),\end{aligned} \tag{7.20}$$

where in the second step we distinguished the two cases $f\le 2\Sigma$ and $f\ge 2\Sigma$, as in (7.18). □

We may now estimate $\widetilde{\Lambda}_d$ in terms of $\widetilde{\Lambda}$.

LEMMA 7.7. *Assume* (2.21). *For* $z = E + i\eta \in D_L$ *we have*

$$\max_i |\widetilde{G}_{ii}(z) - \widetilde{m}(z)| \le C\Phi(z) \tag{7.21}$$

on $\widetilde{\Omega}(\eta)$ *with* (ξ, ν)*-high probability. In particular, on* $\widetilde{\Omega}(\eta)$ *we have with* (ξ, ν)*-high probability*

$$\widetilde{\Lambda}_d(z) \le \widetilde{\Lambda}(z) + C\Phi(z). \tag{7.22}$$

PROOF. Using (7.15), (7.16) and Lemma 3.7, we find

$$\max_i |\widetilde{\Upsilon}_i| \le (\log N)^\xi \frac{C}{q} + C(\log N)^{2\xi}\left(\sqrt{\frac{\operatorname{Im}\widetilde{m}}{N\eta}} + \frac{1}{N\eta}\right) \tag{7.23}$$

on $\widetilde{\Omega}(\eta)$ with (ξ, ν)-high probability. From (7.2) we therefore get

$$\begin{aligned} |\widetilde{G}_{ii} - \widetilde{G}_{jj}| &= |\widetilde{G}_{ii}||\widetilde{G}_{jj}||\widetilde{\Upsilon}_i - \widetilde{\Upsilon}_j| \\ &\le (\log N)^\xi \frac{C}{q} + C(\log N)^{2\xi}\left(\sqrt{\frac{\operatorname{Im}\widetilde{m}}{N\eta}} + \frac{1}{N\eta}\right) \end{aligned} \tag{7.24}$$

on $\widetilde{\Omega}(\eta)$ with (ξ, ν)-high probability. Since $\widetilde{m} = \frac{1}{N}\sum_j \widetilde{G}_{jj}$, the proof is complete. □

7.2. *Establishing* $\widetilde{\Omega}(\eta)$ *with high probability.* What remains to complete the proof of Theorem 2.9 is to prove that the events $\widetilde{\Omega}(\eta)$ hold with (ξ, ν)-high probability. We do this using a simplified version of the continuity argument of Sections 3.4 and 3.6.

LEMMA 7.8. *If* $\eta \ge 2$, *then* $\widetilde{\Omega}(\eta)$ *holds with* (ξ, ν)*-high probability.*

PROOF. The proof is similar to that of Lemma 3.16; we merely sketch the modifications.

Let $z = E + \mathrm{i}\eta \in D_L$ for $\eta \ge 2$. We estimate $\widetilde{\Lambda}_o(z)$ following closely the proof of (7.14), using (7.6) and setting $R = 1$ in (7.12). Using the rough bound $|\widetilde{G}_{ij}| + |\widetilde{m}| \le 1$ as in (3.43), we find

$$\widetilde{\Lambda}_o \le (\log N)^\xi \frac{C}{q} + (\log N)^{2\xi}\frac{C}{\sqrt{N}} + \frac{Cf}{N} \le C(\log N)^{-2\xi}$$

with (ξ, ν)-high probability. Similarly, we find

$$|\widetilde{Z}_i| \le (\log N)^\xi \frac{C}{q} + (\log N)^{2\xi}\frac{C}{\sqrt{N}} \le (\log N)^{-2\xi}$$

with (ξ,ν)-high probability. In order to estimate $\widetilde{\mathcal{A}}_i$, we proceed similarly to (3.44) and find

$$|\widetilde{\mathcal{A}}_i| \le \frac{f}{N} + \frac{f^2}{N}|\langle \mathbf{e}, \widetilde{G}^{(i)}\mathbf{e}\rangle| + \frac{1}{N}\sum_j |\widetilde{G}^{(i)}_{jj}||\widetilde{G}_{ji}|(|a_{ij}| + |\widetilde{Z}_{ij}|).$$

The term $\widetilde{Z}_{ij}$ is estimated exactly as $\widetilde{\Lambda}_o$ above; using the calculation of (7.18), we therefore get

$$|\widetilde{\mathcal{A}}_i| \le (\log N)^{\xi}\frac{C}{q} + (\log N)^{2\xi}\frac{C}{\sqrt{N}} + \frac{Cf}{N} \le C(\log N)^{-2\xi}$$

with (ξ,ν)-high probability.

Now we may follow the proof of Lemma 3.16 to the letter, starting from (3.45) to get $\widetilde{\Lambda}_d \le C(\log N)^{-2\xi}$ with (ξ,ν)-high probability.

Thus, we have proved that $\widetilde{\Lambda}_d(z) + \widetilde{\Lambda}_o(z) \le C(\log N)^{-2\xi}$ with (ξ,ν)-high probability. A simple lattice argument along the lines of Corollary 3.19 then concludes the proof. $\square$

The following simple continuity argument establishes $\widetilde{\Omega}(\eta)$ with (ξ,ν)-high probability for smaller η. Let η_k be a sequence as in Section 3.6.

Note that, unlike in Section 3.6, each step $k \to k+1$ of the continuity argument has to establish a statement for all $z \in D(\eta_{k+1})$.

LEMMA 7.9. *We have*

$$\mathbb{P}(\widetilde{\Omega}(\eta_k)^c) \le k\mathrm{e}^{-\nu(\log N)^{\xi}}.$$

PROOF. We proceed by induction on k. The case $k=1$ was proved in Lemma 7.8. We write

$$\mathbb{P}(\widetilde{\Omega}(\eta_{k+1})^c) \le \mathbb{P}\bigl(\widetilde{\Omega}(\eta_k) \cap \widetilde{\Omega}(\eta_{k+1})^c\bigr) + \mathbb{P}(\widetilde{\Omega}(\eta_k)^c).$$

Now for any $w \in D(\eta_k)$ and on $\widetilde{\Omega}(\eta_k)$ we have, using (7.14), (7.22) and (2.20),

$$\widetilde{\Lambda}_d(w) + \widetilde{\Lambda}_o(w) \le C(\log N)^{-2\xi}$$

with (ξ,ν)-high probability. Using the estimate (3.56), we find, for any $z \in D(\eta_{k+1})$,

$$\widetilde{\Lambda}_d(z) + \widetilde{\Lambda}_o(z) \le C(\log N)^{-2\xi}$$

with (ξ,ν)-high probability. Using a lattice argument similar to Corollary 3.19, we therefore find

$$\mathbb{P}\bigl(\widetilde{\Omega}(\eta_k) \cap \widetilde{\Omega}(\eta_{k+1})^c\bigr) \le \mathrm{e}^{-\nu(\log N)^{\xi}}.$$

The claim follows. $\square$

This estimate (2.22) now follows from (7.14), (7.22), (2.20) and the lattice argument of Corollary 3.19. This concludes the proof of Theorem 2.9.

7.3. *Eigenvector delocalization*: *Proof of Theorem* 2.16. We may now prove Theorem 2.16. Delocalization for the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_{N-1}$ is an immediate consequence of the weak local semicircle law. From (6.1) and Lemma 4.3 we find that $\mu_1, \dots, \mu_{N-1} \in [-\Sigma, \Sigma]$ with (ξ, ν)-high probability. Let $L = 8\xi$ and set $\eta := (\log N)^L$. Using (3.5), Lemma 7.1 and (7.22), we therefore find with (ξ, ν)-high probability

$$C \geq \operatorname{Im} \widetilde{G}_{jj}(\mu_\alpha + \mathrm{i}\eta) = \sum_\beta \frac{\eta |v_\beta(j)|^2}{(\mu_\beta - \mu_\alpha)^2 + \eta^2} \geq \frac{|v_\alpha(j)|^2}{\eta} \qquad (\alpha < N). \tag{7.25}$$

This concludes the proof of (2.32). Moreover, the same argument [with $\alpha = N$ in (7.25)] proves (2.34) if $f \leq \Sigma - 3$, since in that case $\mu_{\max} \in D$ with (ξ, ν)-high probability by (6.5) and Lemma 4.3.

Next, we note that (2.33) is an immediate consequence of (6.9).

In order to prove (2.35), we use the following large deviation estimate which is proved in the Appendix.

LEMMA 7.10. *For $k \leq \log N$ and fixed i we have with (ξ, ν)-high probability*

$$|(H^k \mathbf{e})(i)| = \left| \sum_{i_1, \dots, i_k} h_{ii_1} h_{i_1 i_2} \cdots h_{i_{k-1} i_k} \right| \leq (\log N)^{k\xi}. \tag{7.26}$$

Now from the expansion (6.18) we get with (ξ, ν)-high probability

$$K^{-1} v_{\max}(i) = \frac{1}{\sqrt{N}} + O\left(\frac{(\log N)^\xi}{\sqrt{N} f} \right),$$

and (2.35) follows since $K = 1 + O(f^{-2})$ [see (6.27)]. In this argument we used that $f \sim \mu_{\max} \geq C_0 (\log N)^\xi$ for some large enough C_0 to overcome the logarithmic factors in (6.18) that arise from (7.26). This concludes the proof of Theorem 2.16.

Finally, we outline the proof of (2.36) for $1 \ll f \leq C(\log N)^\xi$. The idea is to use the same proof as for (2.34), relying on the estimate (7.25). In order to do this, we need the pointwise bound $C \geq \widetilde{G}_{ii}(\mu_N + \mathrm{i}\eta)$ which we get by extending the proof of Theorem 2.9 to a larger set D_L. Here D_L has to contain μ_N, so that we have to choose $\Sigma = C(\log N)^\xi$ in the definition (3.1) of D_L with some large constant C.

This extension requires some modifications in our proof of the local semicircle law. Now instead of the bounds (3.8), we have $c(\log N)^{-\xi} \leq |m_{sc}(z)| \leq 1$ for $z \in D_L$. We modify the definitions (3.25) of $\Omega(z)$ and (7.4) of $\widetilde{\Omega}(\eta)$ by replacing $(\log N)^{-\xi}$ with $(\log N)^{-2\xi}$. Then, on these events, we get the lower bound $|G_{ii}(z)| \geq c(\log N)^{-\xi}$ instead of (3.28). One can then check that all estimates of Sections 3–7 remain valid with some deterioration in the form of larger powers of $(\log N)^\xi$, provided that $L \geq C\xi$ for some large enough C; we omit the details.

7.4. *Control of the average of* $\sum_{k\neq i} \widetilde{G}_{1k}^{(i)} h_{ki}$. In this final section we estimate the averaged quantity

$$\Pi := \frac{1}{N} \sum_{i\neq 1} \sum_{k\neq i} \widetilde{G}_{1k}^{(i)} h_{ki}. \tag{7.27}$$

The estimate of Π is not needed for the local semicircle law, but we give its proof here, as it is a natural application of the abstract decoupling lemma, Theorem 5.6. The expression (7.27) arises as an error term when controlling resolvent expansions of the noncentered matrix A. Such expansions are used in the companion paper [13] to establish the universality of the extreme eigenvalues; see Section 6.3 in [13].

Note that a naive application of the large deviation bound (3.19) yields $|\Pi| \leq (\log N)^{C\xi} q^{-1}$ with (ξ, ν)-high probability. In order to establish universality of the extreme eigenvalues in [13], it is crucial that the factor q^{-1} be improved to q^{-2}. This is the content of the following proposition.

PROPOSITION 7.11. *Suppose that the assumptions of Theorem 2.9 hold. Then for any $z \in D_{120(\xi+2)}$ we have with (ξ, ν)-high probability*

$$|\Pi(z)| \leq (\log N)^{C\xi} \left(\frac{1}{q^2} + \frac{\operatorname{Im} m_{\mathrm{sc}}(z)}{N\eta} + \frac{1}{(N\eta)^2} \right). \tag{7.28}$$

PROOF. We shall apply Theorem 5.6 to the quantities

$$\begin{aligned} \mathcal{Z}_i &:= \mathbf{1}(i \neq 1) \sum_{k}^{(i)} \widetilde{G}_{1k}^{(i)} h_{ki}, \\ \mathcal{Z}_i^{[\mathbb{U}]} &:= \mathbf{1}(i \neq 1) \mathbf{1}(i, 1 \notin \mathbb{U}) \sum_{k}^{(\mathbb{U}i)} \widetilde{G}_{1k}^{(\mathbb{U}i)} h_{ki}. \end{aligned} \tag{7.29}$$

Thus, $\Pi = [\mathcal{Z}]$ and $\mathcal{Z}_i^{[\varnothing]} = \mathcal{Z}_i$.

We define the deterministic control parameters

$$X(z) := (\log N)^{40(\xi+2)} \left(\frac{1}{q} + \sqrt{\frac{\operatorname{Im} m_{\mathrm{sc}}(z)}{N\eta}} + \frac{1}{N\eta} \right), \qquad Y(z) := (\log N)^{\xi}$$

and the event

$$\Xi := \bigcap_{z \in D_{120(\xi+2)}} \left\{ \max_{1 \leq i,j \leq N} |\widetilde{G}_{ij}(z) - \delta_{ij} m_{\mathrm{sc}}(z)| \leq X(z) \right\}.$$

Recall that the collection of random variables $(\mathcal{Z}_i^{[\mathbb{U}]})_{\mathbb{U}}$ generates random variables $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ through (5.1). We choose $p := (\log N)^{\xi}$ in Theorem 5.6. It is immediate that the assumptions (i) and (iv) of Theorem 5.6 are satisfied. By Theorem 2.9, the assumption (v) of Theorem 5.6 holds as well.

We shall prove that, for any $\mathbb{U} \subset \mathbb{S}$ with $1, i \notin \mathbb{S}$, $|\mathbb{S}| \leq p$, and $r \leq p$, we have

$$\mathbb{E}(\mathbf{1}([\Xi]_i)|\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}|^r) \leq (Y(CuX)^u)^r, \tag{7.30}$$

where $u := |\mathbb{U}| + 1$. Supposing (7.30) is proved, both assumptions (ii) and (iii) of Theorem 5.6 are satisfied. Then the claim of Theorem 5.6, (5.12) and Markov's inequality yield (7.28).

It remains to prove (7.30). Throughout the following we abbreviate $u := |\mathbb{U}| + 1$. By the definition of $\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}$ in (5.1) and (7.29) we find, for $1, i \notin \mathbb{S}$,

$$\begin{aligned}
\mathcal{Z}_i^{\mathbb{S},\mathbb{U}} &= \mathbf{1}(i \neq 1)(-1)^{|\mathbb{S}\setminus\mathbb{U}|} \sum_{\mathbb{V}:\, \mathbb{S}\setminus\mathbb{U}\subset\mathbb{V}\subset\mathbb{S}} (-1)^{|\mathbb{V}|} \sum_k^{(\mathbb{V}i)} \widetilde{G}_{1k}^{(\mathbb{V}i)} h_{ki} \\
&= \mathbf{1}(i \neq 1) \sum_k^{((i\mathbb{S})\setminus\mathbb{U})} h_{ki}(-1)^{|\mathbb{S}\setminus\mathbb{U}|} \sum_{\mathbb{V}:\, \mathbb{S}\setminus\mathbb{U}\subset\mathbb{V}\subset\mathbb{S}\setminus\{k\}} (-1)^{|\mathbb{V}|} \widetilde{G}_{1k}^{(\mathbb{V}i)} \\
&= \mathbf{1}(i \neq 1) \sum_k^{((i\mathbb{S})\setminus\mathbb{U})} h_{ki} \widetilde{G}_{1k}^{(\mathbb{S}i)\setminus\{k\},\mathbb{U}\setminus\{k\}},
\end{aligned}$$

where in the last step we again used (5.1), as well as Definition 5.8 and the fact that $((\mathbb{S}i) \setminus \{k\}) \setminus (\mathbb{U} \setminus \{k\}) = (\mathbb{S}i) \setminus \mathbb{U}$. We split

$$\mathcal{Z}_i^{\mathbb{S},\mathbb{U}} = D_1 + D_2,$$

where

$$D_1 := \mathbf{1}(i \neq 1) \sum_{k\in\mathbb{U}} h_{ki} \widetilde{G}_{1k}^{(\mathbb{S}i)\setminus\{k\},\mathbb{U}\setminus\{k\}}, \qquad D_2 := \mathbf{1}(i \neq 1) \sum_k^{(\mathbb{S}i)} h_{ki} \widetilde{G}_{1k}^{(\mathbb{S}i),\mathbb{U}}.$$

Thus, we may estimate

$$\mathbb{E}(\mathbf{1}([\Xi]_i)|\mathcal{Z}_i^{\mathbb{S},\mathbb{U}}|^r) \leq 2^r \sum_{j=1}^{2} \mathbb{E}(\mathbf{1}([\Xi]_i)|D_j|^r).$$

To that end, we shall make use of (5.73). Note that Lemma 5.11 is entirely deterministic. In particular, it applies if all quantities are defined in terms of A rather than H (and hence bear a tilde in our convention). We shall apply it to the Green function $\widetilde{G}$.

We start by estimating D_1. Since $\widetilde{G}_{1k}^{(\mathbb{S}i)\setminus\{k\},\mathbb{U}\setminus\{k\}}$ in D_1 is by definition independent of the ith column of H, for $|\mathbb{U}| \leq |\mathbb{S}| \leq p = (\log N)^{\xi}$ we get from (5.73) that

$$\|\mathbf{1}([\Xi]_i)\widetilde{G}_{1k}^{(\mathbb{S}i)\setminus\{k\},\mathbb{U}\setminus\{k\}}\|_\infty = \|\mathbf{1}(\Xi)\widetilde{G}_{1k}^{(\mathbb{S}i)\setminus\{k\},\mathbb{U}\setminus\{k\}}\|_\infty \leq (C|\mathbb{U}|X)^{|\mathbb{U}|}.$$

Here we used that $\widetilde{\Lambda}_o \leq X$ on Ξ. Now we may estimate, using $|\mathbb{U}| \leq (\log N)^\xi$,

$$\begin{aligned}\mathbb{E}(\mathbf{1}([\Xi]_i)|D_1|^r) &\leq (\log N)^{r\xi} \max_{k,i} \mathbb{E}|h_{ki}|^r (C|\mathbb{U}|X)^{r|\mathbb{U}|} \\ &\leq \bigl(C(\log N)^\xi q^{-1}(C|\mathbb{U}|X)^{|\mathbb{U}|}\bigr)^r \\ &\leq (Y(CuX)^u)^r.\end{aligned}$$

Next, we estimate D_2. As above, since $\widetilde{G}_{1k}^{(\mathbb{S}i),\mathbb{U}}$ in D_2 is by definition independent of the ith column of H, for $|\mathbb{U}| \leq (\log N)^\xi$ we get from (5.73) that

$$\|\mathbf{1}([\Xi]_i)\widetilde{G}_{1k}^{(\mathbb{S}i),\mathbb{U}}\|_\infty = \|\mathbf{1}(\Xi)\widetilde{G}_{1k}^{(\mathbb{S}i),\mathbb{U}}\|_\infty \leq (C|\mathbb{U}|X)^{|\mathbb{U}|+1}.$$

Now we use the moment estimate (A.2) with $\alpha = 1$, $\beta = -2$, $\gamma = 1$ and

$$A_k := \mathbf{1}(k \notin \mathbb{S} \cup \{i\})\mathbf{1}([\Xi]_i)\widetilde{G}_{1k}^{(\mathbb{S}i),\mathbb{U}}.$$

This yields

$$\mathbb{E}(\mathbf{1}([\Xi]_i)|D_2|^r) \leq \bigl(r(C|\mathbb{U}|X)^{|\mathbb{U}|+1}\bigr)^r.$$

Here we used that A_k defined above is independent of the randomness in the ith column of H. Thus, we conclude that

$$\mathbb{E}(\mathbf{1}([\Xi]_i)|D_2|^r) \leq (Y(CuX)^u)^r.$$

This completes the proof of (7.30), and hence of (7.28). □

8. Density of states and eigenvalue locations.

8.1. *Local density of states.* The following estimate is the key tool for controlling the local density of states—and hence proving Theorems 2.10 and 2.12.

LEMMA 8.1. *Recall the definition* (2.14) *of the distance κ_E from E to the spectral edge. Suppose that the event*

$$\bigcap_{z \in D_L} \left\{ |\widetilde{m}(z) - m_{\mathrm{sc}}(z)| \leq (\log N)^{C\xi} \left(\min\left\{ \frac{1}{q^2\sqrt{\kappa_E + \eta}}, \frac{1}{q} \right\} + \frac{1}{N\eta} \right) \right\} \tag{8.1}$$

holds with (ξ, ν)-high probability for $L := C_0\xi$, where C_0 is a positive constant. For given $E_1 < E_2$ in $[-\Sigma, \Sigma]$ we abbreviate

$$\kappa := \min\{\kappa_{E_1}, \kappa_{E_2}\}, \qquad \mathcal{E} := \max\{E_2 - E_1, (\log N)^L N^{-1}\}. \tag{8.2}$$

Then, for any $-\Sigma \leq E_1 < E_2 \leq \Sigma$, we have

$$\bigl|\bigl(\widetilde{\mathfrak{n}}(E_2) - \widetilde{\mathfrak{n}}(E_1)\bigr) - \bigl(n_{\mathrm{sc}}(E_2) - n_{\mathrm{sc}}(E_1)\bigr)\bigr| \leq (\log N)^{C\xi} \left[\frac{1}{N} + \frac{\mathcal{E}}{q^2\sqrt{\kappa + \mathcal{E}}} \right] \tag{8.3}$$

with (ξ, ν)-high probability.

PROOF. Recall the definitions (2.11) and (2.12). Similarly, we have

$$\widetilde{\varrho}(x) = \frac{1}{N} \sum_{\alpha=1}^{N} \delta(x - \mu_\alpha), \qquad \widetilde{\mathfrak{n}}(E) = \int_{-\infty}^{E} \widetilde{\varrho}(x)\,\mathrm{d}x = \frac{1}{N} |\{\alpha : \mu_\alpha \le E\}|.$$

Thus, we may write

$$\widetilde{m}(z) = \frac{1}{N} \operatorname{Tr} G(z) = \int \frac{\widetilde{\varrho}(x)\,\mathrm{d}x}{x - z}.$$

We introduce the differences

$$\varrho^\Delta := \widetilde{\varrho} - \varrho_{\mathrm{sc}}, \qquad m^\Delta := \widetilde{m} - m_{\mathrm{sc}}.$$

Following [15], we use the Helffer–Sjöstrand functional calculus. Set $\widetilde{\eta} := (\log N)^L N^{-1}$. (Recall that $L = C_0 \xi$.) Let χ be a smooth cutoff function equal to 1 on $[-\mathcal{E}, \mathcal{E}]$ and vanishing on $[-2\mathcal{E}, 2\mathcal{E}]^c$, such that $|\chi'(y)| \le C\mathcal{E}^{-1}$. Set $\eta := N^{-1}$ and let f be a characteristic function of the interval $[E_1, E_2]$ smoothed on the scale η: $f(x) = 1$ on $[E_1, E_2]$, $f(x) = 0$ on $[E_1 - \eta, E_2 + \eta]^c$, $|f'(x)| \le C\eta^{-1}$, and $|f''(x)| \le C\eta^{-2}$. Note that the supports of f' and f'' have measure $O(\eta)$.

Then we have the estimate (see equation (B.13) in [15])

$$\begin{aligned}
\left| \int f(\lambda) \varrho^\Delta(\lambda)\,\mathrm{d}\lambda \right| &\le C \left| \int \mathrm{d}x \int_0^\infty \mathrm{d}y\, \big(f(x) + y f'(x)\big) \chi'(y) m^\Delta(x + \mathrm{i}y) \right| \\
&\quad + C \left| \int \mathrm{d}x \int_0^\eta \mathrm{d}y\, f''(x) \chi(y) y \operatorname{Im} m^\Delta(x + \mathrm{i}y) \right| \\
&\quad + C \left| \int \mathrm{d}x \int_\eta^\infty \mathrm{d}y\, f''(x) \chi(y) y \operatorname{Im} m^\Delta(x + \mathrm{i}y) \right|.
\end{aligned} \tag{8.4}$$

Since χ' vanishes away from $[\mathcal{E}, 2\mathcal{E}]$, the first term on the right-hand side of (8.4) is bounded with (ξ, ν)-high probability by

$$\begin{aligned}
&\frac{C}{\mathcal{E}} \int \mathrm{d}x \int_{\mathcal{E}}^{2\mathcal{E}} \mathrm{d}y\, |f(x) + y f'(x)| (\log N)^{C\xi} \left(\frac{1}{q^2 \sqrt{\kappa + \mathcal{E}}} + \frac{1}{N\mathcal{E}} \right) \\
&\qquad \le (\log N)^{C\xi} \left(\frac{\mathcal{E}}{q^2 \sqrt{\kappa + \mathcal{E}}} + \frac{1}{N} \right),
\end{aligned}$$

where we abbreviated $\kappa := \min\{\kappa_{E_1}, \kappa_{E_2}\}$. In order to estimate the two remaining terms of (8.4), we estimate $\operatorname{Im} m^\Delta(x + \mathrm{i}y)$. If $y \ge \widetilde{\eta}$, we may use (8.1). Consider therefore the case $0 < y \le \widetilde{\eta}$. From Lemma 3.2 we find

$$|\operatorname{Im} m_{\mathrm{sc}}(x + \mathrm{i}y)| \le C\sqrt{\kappa_x + y}. \tag{8.5}$$

By spectral decomposition of A, it is easy to see that the function $y \mapsto y \operatorname{Im} \widetilde{m}(x + \mathrm{i}y)$ is monotone increasing. Thus, we get, using (8.5), $x + \mathrm{i}\widetilde{\eta} \in D_L$ and (8.1), that

$$\begin{aligned}
y \operatorname{Im} \widetilde{m}(x + \mathrm{i}y) &\le \widetilde{\eta} \operatorname{Im} \widetilde{m}(x + \mathrm{i}\widetilde{\eta}) \le (\log N)^{C\xi} \widetilde{\eta} \left(\sqrt{\kappa_x + \widetilde{\eta}} + \frac{1}{q} + \frac{1}{N\widetilde{\eta}} \right) \\
&\le \frac{(\log N)^{C\xi}}{N} \qquad (y \le \widetilde{\eta})
\end{aligned} \tag{8.6}$$

with (ξ,ν)-high probability. From (8.5), $m^\Delta = \widetilde{m} - m_{sc}$, and the definition of $\widetilde{\eta}$, we find

$$|y \operatorname{Im} m^\Delta(x+\mathrm{i}y)| \le \frac{(\log N)^{C\xi}}{N} \qquad (y \le \widetilde{\eta}) \tag{8.7}$$

with (ξ,ν)-high probability. Since $\eta \le \widetilde{\eta}$, we therefore find that the second term of (8.4) is bounded with (ξ,ν)-high probability by

$$\frac{C(\log N)^{C\xi}}{N} \int \mathrm{d}x\, |f''(x)| \int_0^\eta \mathrm{d}y\, \chi(y) \le \frac{(\log N)^{C\xi}}{N}.$$

In order to estimate the third term on the right-hand side of (8.4), we integrate by parts, first in x and then in y, to obtain the bound

$$\begin{aligned} & C\left|\int \mathrm{d}x\, f'(x)\eta \operatorname{Re} m^\Delta(x+\mathrm{i}\eta)\right| \\ & \quad + C\left|\int \mathrm{d}x \int_\eta^\infty \mathrm{d}y\, f'(x)\chi'(y) y \operatorname{Re} m^\Delta(x+\mathrm{i}y)\right| \\ & \quad + C\left|\int \mathrm{d}x \int_\eta^\infty \mathrm{d}y\, f'(x)\chi(y) \operatorname{Re} m^\Delta(x+\mathrm{i}y)\right|. \end{aligned} \tag{8.8}$$

The second term of (8.8) is similar to the first term on the right-hand side of (8.4), and is easily seen to be bounded by

$$(\log N)^{C\xi}\left(\frac{\mathcal{E}}{q^2\sqrt{\kappa+\mathcal{E}}} + \frac{1}{N}\right).$$

In order to bound the first and third terms of (8.8), we estimate, for any $y \le \widetilde{\eta}$,

$$|m^\Delta(x+\mathrm{i}y)| \le |m^\Delta(x+\mathrm{i}\widetilde{\eta})| + \int_y^{\widetilde{\eta}} \mathrm{d}u\, \big(|\partial_u \widetilde{m}(x+\mathrm{i}u)| + |\partial_u m_{sc}(x+\mathrm{i}u)|\big).$$

Moreover, using (8.6) and (3.27), we find for any $u \le \widetilde{\eta}$ that

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

with (ξ,ν)-high probability. Similarly, we find from (2.12) that

$$|\partial_u m_{sc}(x+\mathrm{i}u)| \le \frac{1}{u^2}\widetilde{\eta} \operatorname{Im} m_{sc}(x+\mathrm{i}\widetilde{\eta}).$$

Thus, (8.1) yields

$$|m^\Delta(x+\mathrm{i}y)| \le (\log N)^{C\xi}\left(1 + \int_y^{\widetilde{\eta}} \mathrm{d}u\, \frac{\widetilde{\eta}}{u^2}\right) \le (\log N)^{C\xi}\frac{\widetilde{\eta}}{y} \qquad (y \le \widetilde{\eta}) \tag{8.9}$$

with (ξ, ν)-high probability.

Using (8.9), we may now bound the first term of (8.8) by $(\log N)^{C\xi} N^{-1}$.

What remains is the third term of (8.8). We first split the y-integration domain $[\eta, \infty)]$ into the pieces $[\eta, \widetilde{\eta}]$ and $[\widetilde{\eta}, \infty)$. Using (8.9), we estimate the integral over the first piece, with (ξ, ν)-high probability, by

$$\int \mathrm{d}x\, |f'(x)| \int_{\eta}^{\widetilde{\eta}} \mathrm{d}y\, |m^{\Delta}(x + \mathrm{i}y)| \leq \frac{(\log N)^{C\xi}}{N}.$$

Using (8.1), we may therefore estimate the third term of (8.8), with (ξ, ν)-high probability, by

$$\begin{aligned}
&\frac{(\log N)^{C\xi}}{N} + (\log N)^{C\xi} \int \mathrm{d}x \int_{\widetilde{\eta}}^{2\mathcal{E}} \mathrm{d}y\, |f'(x)| \biggl(\frac{1}{Ny} + \frac{1}{\sqrt{\kappa_x + y}} \frac{1}{q^2} \biggr) \\
&\quad \leq \frac{(\log N)^{C\xi}}{N} + (\log N)^{C\xi} \int \mathrm{d}x\, |f'(x)| \biggl[\int_{\widetilde{\eta}}^{2\mathcal{E}} \mathrm{d}y\, \frac{1}{Ny} + \frac{1}{q^2} \int_{\widetilde{\eta}}^{2\mathcal{E}} \mathrm{d}y\, \frac{1}{\sqrt{\kappa + y}} \biggr] \\
&\quad \leq (\log N)^{C\xi} \biggl(\frac{1}{N} + \frac{\mathcal{E}}{q^2 \sqrt{\kappa + \mathcal{E}}} \biggr).
\end{aligned}$$

Summarizing, we have proved that

$$\biggl| \int f(\lambda) \varrho^{\Delta}(\lambda)\, \mathrm{d}\lambda \biggr| \leq (\log N)^{C\xi} \biggl[\frac{1}{N} + \frac{\mathcal{E}}{q^2 \sqrt{\kappa + \mathcal{E}}} \biggr] \tag{8.10}$$

with (ξ, ν)-high probability.

In order to estimate $|\widetilde{\mathfrak{n}}(E) - n_{\mathrm{sc}}(E)|$, we observe that (8.6) implies

$$|\widetilde{\mathfrak{n}}(x + \eta) - \widetilde{\mathfrak{n}}(x - \eta)| \leq C\eta \operatorname{Im} \widetilde{m}(x + \mathrm{i}\eta) \leq \frac{(\log N)^{C\xi}}{N}$$

with (ξ, ν)-high probability. Thus, we get

$$\begin{aligned}
\biggl| \widetilde{\mathfrak{n}}(E_1) - \widetilde{\mathfrak{n}}(E_2) - \int f(\lambda) \varrho(\lambda)\, \mathrm{d}\lambda \biggr| &\leq C \sum_{i=1,2} \bigl(\widetilde{\mathfrak{n}}(E_i + \eta) - \widetilde{\mathfrak{n}}(E_i - \eta) \bigr) \\
&\leq \frac{C(\log N)^{C\xi}}{N}
\end{aligned}$$

with (ξ, ν)-high probability. Similarly, since ϱ_{sc} has a bounded density, we find

$$\biggl| n_{\mathrm{sc}}(E_1) - n_{\mathrm{sc}}(E_2) - \int f(\lambda) \varrho_{\mathrm{sc}}(\lambda)\, \mathrm{d}\lambda \biggr| \leq C\eta = \frac{C}{N}.$$

Together with (8.10), we therefore get (8.3). $\square$

We draw two simple consequences from Lemma 8.1.

PROPOSITION 8.2 (Uniform local density of states). *Suppose that A satisfies Definition 2.2 and that ξ and q satisfy (2.15). Then, for any E_1 and E_2 satisfying $E_2 \geq E_1 + (\log N)^{C\xi} N^{-1}$ we have*

$$\widetilde{\mathcal{N}}(E_1, E_2) = \mathcal{N}_{\mathrm{sc}}(E_1, E_2)\left[1 + O\left(\frac{(\log N)^{C\xi}}{\mathcal{N}_{\mathrm{sc}}(E_1, E_2)}\left(1 + \frac{N}{q^2}\frac{E_2 - E_1}{\sqrt{\kappa + E_2 - E_1}}\right)\right)\right] \tag{8.11}$$

with (ξ, ν)-high probability, where we abbreviated $\kappa := \min\{\kappa_{E_1}, \kappa_{E_2}\}$.

PROOF. By (2.20), the estimate (8.1) holds. Assuming $-\Sigma \leq E_1 \leq E_2 \leq \Sigma$, we get from (8.3), with (ξ, ν)-high probability,

$$|\widetilde{\mathcal{N}}(E_1, E_2) - \mathcal{N}_{\mathrm{sc}}(E_1, E_2)| \leq (\log N)^{C\xi}\left(1 + \frac{N}{q^2}\frac{E_2 - E_1}{\sqrt{\kappa + E_2 - E_1}}\right),$$

from which the claim follows. If $E_1 < -\Sigma$ and $\mathbb{E}_2 \leq \Sigma$, the claim follows by replacing E_1 with $-\Sigma$ and using Lemma 4.4. The other cases where $-\Sigma \leq E_1 \leq E_2 \leq \Sigma$ does not hold are treated similarly using Lemma 4.4. □

The proof of Theorem 2.10 is completed by observing that both its statements, (2.25) and (2.24), are special cases of (8.11). [Recall that in the bulk we have $\mathcal{N}_{\mathrm{sc}}(E_1, E_2) \sim N(E_2 - E_1)$; at the spectral edge we have $\mathcal{N}_{\mathrm{sc}}(E_1, E_2) \geq N(E_2 - E_1)^{3/2}$, which is sharp for $E_1 = -2$.]

PROOF OF THEOREM 2.12. Let us assume that $E \leq 0$; the case $E > 0$ is treated similarly. Setting

$$E_1 := -2 - (\log N)^{C_1\xi}(q^{-2} + N^{-2/3})$$

for some constant $C_1 > 0$, we find that $n_{\mathrm{sc}}(E_1) = 0$ and $\widetilde{\mathfrak{n}}(E_1) = 0$ with (ξ, ν)-high probability for C_1 large enough, by Lemma 4.4. We may assume that $E \geq E_1$.

By (2.20), the estimate (8.1) holds. Therefore, setting $E_2 = E$ in Lemma 8.1 yields

$$\begin{aligned} |\widetilde{\mathfrak{n}}(E) - n_{\mathrm{sc}}(E)| &\leq (\log N)^{C\xi}\left(\frac{1}{N} + \frac{1}{q^2}\sqrt{E - E_1 + (\log N)^{C\xi} N^{-1}}\right) \\ &\leq (\log N)^{C\xi}\left(\frac{1}{N} + \frac{1}{q^3} + \frac{\sqrt{\kappa_E}}{q^2}\right) \end{aligned}$$

with (ξ, ν)-high probability. This holds for any fixed E. The claim (2.27), which is uniform in E, now follows by a lattice argument similar to Corollary 3.19, whereby we choose a lattice of points $E_i \in [-\Sigma, \Sigma]$ with $|E_{i+1} - E_i| \leq N^{-1}$; we omit the details. □

8.2. *Eigenvalue locations*. The following result contains the main estimate on the locations of the eigenvalues μ_α of A. Recall the definition (2.28) of the classical location γ_α of the αth eigenvalue.

PROPOSITION 8.3. *Suppose that A satisfies Definition 2.2 and that ξ satisfies (2.15). Let ϕ be an exponent satisfying $0 < \phi \le 1/2$ and assume that $q = N^\phi$. Then the following statements hold with (ξ, ν)-high probability for all $\alpha = 1, \dots, N-1$, for some sufficiently large constant K*:

(i) *If* $\max\{\kappa_{\mu_\alpha}, \kappa_{\gamma_\alpha}\} \le (\log N)^{K\xi}(N^{-2/3} + N^{-2\phi})$, *then*

$$|\mu_\alpha - \gamma_\alpha| \le (\log N)^{C\xi}(N^{-2/3} + N^{-2\phi}). \tag{8.12}$$

(ii) *If* $\max\{\kappa_{\mu_\alpha}, \kappa_{\gamma_\alpha}\} \ge (\log N)^{K\xi}(N^{-2/3} + N^{-2\phi})$, *then*

$$|\mu_\alpha - \gamma_\alpha| \le (\log N)^{C\xi}(N^{-2/3}\widehat{\alpha}^{-1/3} + N^{2/3-4\phi}\widehat{\alpha}^{-2/3} + N^{-2\phi}), \tag{8.13}$$

where we abbreviated $\widehat{\alpha} := \min\{\alpha, N-\alpha\}$.

PROOF. To simplify the presentation, we concentrate only on the eigenvalues $\mu_1, \dots, \mu_{N/2}$. The remaining eigenvalues $\mu_{N/2+1}, \dots, \mu_{N-1}$ are dealt with similarly, using Lemma 4.4 and the estimate $\mu_N \ge 2 + c$ which holds with (ξ, ν)-high probability.

We define the event $\widetilde{\Omega}$ as the intersection of the events on which (4.30) holds and on which

$$|\widetilde{\mathfrak{n}}(E) - n_{\mathrm{sc}}(E)| \le (\log N)^{C_0\xi}\left(\frac{1}{N} + \frac{1}{q^3} + \frac{\sqrt{\kappa_E}}{q^2}\right) \tag{8.14}$$

holds for all $E \in [-\Sigma, \Sigma]$ and some positive constant C_0. Recalling (2.27) and (4.30), we find that $\widetilde{\Omega}$ holds with (ξ, ν)-high probability for large enough C_0. Note that on $\widetilde{\Omega}$ we have $\mu_{N/2} \le 1$. Indeed, the condition $\mu_{N/2} \le 1$ is equivalent to $\widetilde{\mathfrak{n}}(1) \ge 1/2$, which follows from (8.14) and the fact that $n_{\mathrm{sc}}(1) > 1/2$.

Abbreviate $\zeta := \min\{2\phi, 2/3\}$ and let $C_1 > C_0$. We use the dyadic decomposition

$$\{1, \dots, N/2\} = \bigcup_{k=0}^{2\log N} U_k,$$

where we defined

$$U_0 := \{\alpha \le N/2 : 2 + \max\{\gamma_\alpha, \mu_\alpha\} \le 2(\log N)^{C_1\xi}N^{-\zeta}\},$$

$$U_k := \{\alpha \le N/2 : 2^k(\log N)^{C_1\xi}N^{-\zeta} < 2 + \max\{\gamma_\alpha, \mu_\alpha\} \le 2^{k+1}(\log N)^{C_1\xi}N^{-\zeta}\} \quad \text{for } k \ge 1.$$

By definition of U_0 and Lemma 4.4, on $\widetilde{\Omega}$ we have

$$|\mu_\alpha - \gamma_\alpha| \le (\log N)^{C\xi}N^{-\zeta} \qquad (\alpha \in U_0).$$

This proves (8.12).

Next, let $k \geq 1$. From (8.14) we find that on $\widetilde{\Omega}$ we have

$$\frac{\alpha}{N} = n_{\mathrm{sc}}(\gamma_\alpha) = \widetilde{\mathfrak{n}}(\mu_\alpha) = n_{\mathrm{sc}}(\mu_\alpha) + (\log N)^{C_0 \xi} O\left(\frac{1}{N} + \frac{1}{q^3} + \frac{\sqrt{\kappa_{\mu_\alpha}}}{q^2}\right). \tag{8.15}$$

On $\widetilde{\Omega}$ and for $\alpha \in U_k$, the second term on the right-hand side of (8.15) may be estimated as

$$\begin{aligned}&(\log N)^{C_0\xi} O\left(\frac{1}{N} + \frac{1}{q^3} + \frac{\sqrt{\kappa_{\mu_\alpha}}}{q^2}\right)\\ &\qquad \leq (\log N)^{C_0\xi}(N^{-1} + N^{-3\phi}) + C2^{(k+1)/2}(\log N)^{(C_0 + C_1/2)\xi} N^{-\zeta/2 - 2\phi},\end{aligned}$$

since $\kappa_{\mu_\alpha} \leq 2 + \mu_\alpha$. Moreover, on $\widetilde{\Omega}$ and for $\alpha \in U_k$ we have

$$n_{\mathrm{sc}}(\gamma_\alpha) + n_{\mathrm{sc}}(\mu_\alpha) \geq c 2^{3k/2} (\log N)^{(3/2)C_1\xi} N^{-3\zeta/2},$$

where we used the simple estimate $n_{\mathrm{sc}}(-2 + x) \sim x^{3/2}$ for $0 \leq x \leq 3$. Thus, we have, on $\widetilde{\Omega}$ and for $\alpha \in U_k$,

$$(\log N)^{C_0\xi} O\left(\frac{1}{N} + \frac{1}{q^3} + \frac{\sqrt{\kappa_{\mu_\alpha}}}{q^2}\right) \ll n_{\mathrm{sc}}(\gamma_\alpha) + n_{\mathrm{sc}}(\mu_\alpha),$$

from which we deduce using (8.15) that

$$n_{\mathrm{sc}}(\mu_\alpha) = n_{\mathrm{sc}}(\gamma_\alpha)\bigl(1 + O\bigl[(\log N)^{-(C_1 - C_0)\xi}\bigr]\bigr).$$

Thus, we find, on $\widetilde{\Omega}$ and for $\alpha \in U_k$, that $2 + \gamma_\alpha \sim 2 + \mu_\alpha$ and, hence, $n'_{\mathrm{sc}}(x) \sim n'_{\mathrm{sc}}(\gamma_\alpha)$ for any x between γ_α and μ_α. Here we used that $n'_{\mathrm{sc}}(x) \sim (n_{\mathrm{sc}}(x))^{1/3} \sim \sqrt{2 + x}$ for $-2 \leq x \leq 1$. Thus, the mean value theorem and (8.15) imply, on $\widetilde{\Omega}$ and for $\alpha \in U_k$,

$$\begin{aligned}&|\mu_\alpha - \gamma_\alpha|\\ &\qquad \leq \frac{C|n_{\mathrm{sc}}(\mu_\alpha) - n_{\mathrm{sc}}(\gamma_\alpha)|}{n'_{\mathrm{sc}}(\gamma_\alpha)}\\ &\qquad \leq \frac{C(\log N)^{C_0\xi}}{(\alpha/N)^{1/3}}\bigl(N^{-1} + N^{-3\phi} + N^{-2\phi}\sqrt{\kappa_{\mu_\alpha}}\bigr)\\ &\qquad \leq \frac{C(\log N)^{C_0\xi}}{\alpha^{1/3}}\bigl(N^{-2/3} + N^{1/3 - 3\phi} + N^{-2\phi}\alpha^{1/3} + N^{1/3 - 2\phi}\sqrt{|\mu_\alpha - \gamma_\alpha|}\bigr),\end{aligned}$$

where we used that $\kappa_{\mu_\alpha} \leq \kappa_{\gamma_\alpha} + |\mu_\alpha - \gamma_\alpha|$ and $\kappa_{\gamma_\alpha} \sim (\alpha/N)^{2/3}$. Thus, we find, on $\widetilde{\Omega}$ and for $\alpha \in U_k$,

$$|\mu_\alpha - \gamma_\alpha| \leq (\log N)^{C\xi}(N^{-2/3}\alpha^{-1/3} + N^{2/3 - 4\phi}\alpha^{-2/3} + N^{-2\phi}).$$

This proves (8.13). □

PROOF OF THEOREM 2.13. We apply Proposition 8.3. As before, we only deal with the eigenvalues $\alpha \le N/2$; the proof for the eigenvalues $N/2 < \alpha \le N-1$ is the same. Suppose that α satisfies Case (i) of Proposition 8.3. Using $\alpha/N = n_{\mathrm{sc}}(\gamma_\alpha) \sim (2+\gamma_\alpha)^{3/2}$, we find that

$$\alpha \le (\log N)^{K\xi}(1+N^{1-3\phi}). \tag{8.16}$$

Therefore, we get, squaring (8.12) and (8.13) and summing over α,

$$\sum_{\alpha=1}^{N-1} |\mu_\alpha - \gamma_\alpha|^2 \le (\log N)^{C\xi}(1+N^{1-3\phi})(N^{-4/3}+N^{-4\phi}) + (\log N)^{C\xi}(N^{-1}+N^{4/3-8\phi}+N^{1-4\phi})$$

with (ξ,ν)-high probability. This concludes the proof of (2.29).

Finally, we note that (2.30) follows from (8.12) and (8.13) as well as the above observation that Case (i) in Proposition 8.3 implies (8.16). □

APPENDIX: MOMENT ESTIMATES: PROOFS OF LEMMAS 3.8, 4.3, 6.5 AND 7.10

In order to prove Lemma 3.8, we prove the following high moment bounds, which are also independently useful.

LEMMA A.1. (i) *Let (a_i) be a family of centered and independent random variables satisfying*

$$\mathbb{E}|a_i|^p \le \frac{C^p}{N^\gamma q^{\alpha p+\beta}} \tag{A.1}$$

for all $2 \le p \le (\log N)^{A_0 \log\log N}$, where $\alpha \ge 0$ and $\beta, \gamma \in \mathbb{R}$. Then for all even p satisfying $2 \le p \le (\log N)^{A_0\log\log N}$ we have

$$\mathbb{E}\Big|\sum_i A_i a_i\Big|^p \le (Cp)^p\bigg[\frac{\sup_i |A_i|}{q^\alpha} + \bigg(\frac{1}{N^\gamma q^{\beta+2\alpha}}\sum_i |A_i|^2\bigg)^{1/2}\bigg]^p \tag{A.2}$$

for some constant $C>0$ depending only on the constant in (A.1).

(ii) *Let $a_1,\dots,a_N$ be centered and independent random variables satisfying*

$$\mathbb{E}|a_i|^p \le \frac{C^p}{Nq^{p-2}} \tag{A.3}$$

for all $2 \le p \le (\log N)^{A_0\log\log N}$. Then for all even p satisfying $2 \le p \le (\log N)^{A_0\log\log N}$ and all $B_{ij} \in \mathbb{C}$ we have

$$\mathbb{E}\Big|\sum_{i\ne j} \bar a_i B_{ij} a_j\Big|^p \le (Cp)^{2p}\bigg[\frac{\max_{i\ne j}|B_{ij}|}{q} + \bigg(\frac{1}{N^2}\sum_{i\ne j}|B_{ij}|^2\bigg)^{1/2}\bigg]^p \tag{A.4}$$

for some C depending only on the constant in (A.3).

PROOF. We begin with (i). To prove (A.2), we set $p = 2r$ and compute

$$\mathbb{E}\biggl|\sum_i A_i a_i\biggr|^{2r} = \sum_{i_1,\dots,i_{2r}} \bar{A}_{i_1} \cdots \bar{A}_{i_r} A_{i_{r+1}} \cdots A_{i_{2r}} \mathbb{E} \bar{a}_{i_1} \cdots \bar{a}_{i_r} a_{i_{r+1}} \cdots a_{i_{2r}}. \tag{A.5}$$

Each configuration of labels $(i_1, \dots, i_{2r})$ defines an equivalence relation (or partition) Γ on the index set $\{1, \dots, 2r\}$ by requiring that the indices j and k are in the same equivalence class if and only if their labels satisfy $i_j = i_k$. We organize the summation over the labels $i_1, \dots, i_{2r}$ by (i) prescribing a partition Γ of the set of indices, (ii) summing over all label configurations yielding the partition Γ, and (iii) summing over all partitions Γ. Thus, let a partition Γ be given. Let l denote the number of equivalence classes of Γ, and order the equivalence classes in some arbitrary fashion. Let r_s be the size of equivalence class s; clearly, we have $r_1 + \cdots + r_l = 2r$. Moreover, since the random variables a_i are centered, we find that each equivalence class has size at least 2; in particular, $r_s \geq 2$ for each s and, hence, $l \leq r$. Using the independence of the a_i's, we thus find that the contribution of the partition Γ to (A.5) is bounded in absolute value by

$$\sum_{i_1,\dots,i_l} \prod_{s=1}^{l} |A_{i_s}|^{r_s} \mathbb{E}|a_{i_s}|^{r_s} \leq \prod_{s=1}^{l} \biggl(\sum_i |A_i|^{r_s} \frac{C^{r_s}}{N^\gamma q^{\alpha r_s + \beta}} \biggr). \tag{A.6}$$

Abbreviating $A := \max_i |A_i|$, we find that (A.6) is bounded by

$$\begin{aligned}
&\prod_{s=1}^{l} \biggl((CAq^{-\alpha})^{r_s} A^{-2} N^{-\gamma} q^{-\beta} \sum_i |A_i|^2 \biggr) \\
&\quad = (CAq^{-\alpha})^{2r} \biggl(\frac{1}{A^2 N^\gamma q^\beta} \sum_i |A_i|^2 \biggr)^l \\
&\quad \leq (CAq^{-\alpha})^{2r} \max\biggl\{ 1, \biggl(\frac{1}{A^2 N^\gamma q^\beta} \sum_i |A_i|^2 \biggr)^r \biggr\} \\
&\quad \leq C^r \biggl[\frac{A}{q^\alpha} + \biggl(\frac{1}{N^\gamma q^{\beta + 2\alpha}} \sum_i |A_i|^2 \biggr)^{1/2} \biggr]^{2r}.
\end{aligned}$$

Next, it is easy to see that the total number of partitions of $2r$ elements is bounded by $(Cr)^{2r}$, so that we get

$$\mathbb{E}\biggl|\sum_i A_i a_i\biggr|^{2r} \leq (Cr)^{2r} \biggl[\frac{A}{q^\alpha} + \biggl(\frac{1}{N^\gamma q^{\beta + 2\alpha}} \sum_i |A_i|^2 \biggr)^{1/2} \biggr]^{2r}.$$

This concludes the proof of (A.2).

The proof of (A.4) needs more effort. Without loss of generality, we set $B_{ii} = 0$ for all i. As above, we set $p = 2r$. We find

$$\mathbb{E}\bigg|\sum_{i\neq j} \bar{a}_i B_{ij} a_j\bigg|^{2r} = \sum_{i_1,\ldots,i_{4r}} \bar{B}_{i_1 i_2} \cdots \bar{B}_{i_{2r-1} i_{2r}} B_{i_{2r+1} i_{2r+2}} \cdots B_{i_{4r-1} i_{4r}} \times \mathbb{E} a_{i_1} \bar{a}_{i_2} \cdots a_{i_{2r-1}} \bar{a}_{i_{2r}} \bar{a}_{i_{2r+1}} a_{i_{2r+2}} \cdots \bar{a}_{i_{4r-1}} a_{i_{4r}}. \tag{A.7}$$

As above, we associate a partition $\Gamma(\mathbf{i}) \equiv \Gamma = \{\gamma\}$ of the index set $\{1, \ldots, 4r\}$ with every label configuration $\mathbf{i} = (i_1, \ldots, i_{4r})$ by requiring that k and l are in the same equivalence class of Γ if and only if $i_k = i_l$. We rewrite (A.7) by first specifying a partition Γ and summing over all label configurations $\mathbf{i}$ satisfying $\Gamma(\mathbf{i}) = \Gamma$, and subsequently summing over all partitions Γ. Note that a partition Γ yields a nonzero contribution to the right-hand side of (A.7) only if (i) each equivalence class contains at least two indices, and (ii) $[2k-1] \neq [2k]$ for all $k = 1, \ldots, 2r$; here $[n]$ denotes the equivalence class $\gamma \ni n$ of n in Γ.

Next, we encode Γ using a multigraph (i.e., a graph which may have multiple edges) $G \equiv G(\Gamma)$ defined as follows. The vertex set of G is the set of equivalence classes $\{\gamma\}$ of Γ. Each factor $\bar{B}_{i_{2k-1} i_{2k}}$ or $B_{i_{2k-1} i_{2k}}$ gives rise to an edge of G connecting the vertices $[2k-1]$ and $[2k]$. Note that, by property (ii) of Γ, no edge of G connects a vertex to itself. Moreover, G has $2r$ edges.

Let G be a multigraph with v vertices. We define the *value* of G through

$$\mathscr{V}(G) := \sum_{i_1,\ldots,i_v} \bigg(\prod_{\{\gamma,\gamma'\}\in E(G)} |B_{i_\gamma i_{\gamma'}}| \bigg) \prod_{\gamma=1}^{v} \frac{1}{N q^{[\delta_\gamma - 2]_+}}, \tag{A.8}$$

where δ_γ is the degree of γ in G.

Fix a partition Γ. We claim that the contribution to the right-hand side of (A.7) of all label configurations $\mathbf{i}$ satisfying $\Gamma(\mathbf{i}) = \Gamma$ is bounded in absolute value by $C^r \mathscr{V}(G(\Gamma))$. This is an easy consequence of the definition of $G(\Gamma)$: each vertex γ carries a label i_γ, and the contribution of vertex γ is bounded by $\mathbb{E}|a_{i_\gamma}|^{\delta_\gamma} \leq C^{\delta_\gamma}(N q^{\delta_\gamma - 2})^{-1}$. [Note that, by the property (i) of Γ, we have $\delta_\gamma \geq 2$. Here we also used that $\sum_\gamma \delta_\gamma = 4r$.]

Next, we estimate $\mathscr{V}(G(\Gamma))$. Let $G_0 := G(\Gamma)$. The idea is to construct a sequence of multigraphs $G_0, G_1, \ldots, G_s$ by successively removing edges incident to vertices of degree greater than two, until all vertices have degree at most two.

If all vertices of G_0 have degree at most two, set $s = 0$. Otherwise, pick a vertex $\widetilde{\gamma}$ of G_0 with degree greater than two, and let $\widetilde{\gamma}'$ be adjacent to $\widetilde{\gamma}$. Define $R(G_0)$ as the multigraph obtained from G_0 by removing an edge connecting $\widetilde{\gamma}$ and $\widetilde{\gamma}'$. We claim that

$$\mathscr{V}(G_0) \leq \frac{B_o}{d} \mathscr{V}(R(G_0)) \tag{A.9}$$

(regardless of the choice of the removed edge). Here we abbreviated $B_o := \max_{i\neq j} |B_{ij}|$. The estimate (A.9) is obtained by estimating $|B_{i_{\widetilde{\gamma}} i_{\widetilde{\gamma}'}}| \leq B_o$ in (A.8), and by noting that $[\delta_\gamma - 2]_+$ in G_0 is strictly greater than in $R(G_0)$. Now set $G_1 := R(G_0)$.

We continue inductively in this manner, generating a sequence $G_0, \dots, G_s$ of multigraphs with the properties that $G_{k+1} = R(G_k)$ (for an immaterial choice of R), G_s has $2r - s$ edges, and all vertices of G_s have degree at most two. By (A.9), we have

$$\mathscr{V}(G_0) \le \left(\frac{B_o}{q}\right)^s \mathscr{V}(G_s). \tag{A.10}$$

Next, it is immediate from its definition that G_s is a disjoint union of simple closed and open paths. Here a simple open path of length $l \ge 0$ is the graph with vertices $1, \dots, l+1$ and edges $\{1,2\}, \dots, \{l, l+1\}$; similarly, a simple closed path of length $l \ge 2$ is the graph with vertices $1, \dots, l$ and edges $\{1,2\}, \dots, \{l-1, l\}, \{l, 1\}$.

From the definition (A.8) we immediately find

$$\mathscr{V}(G \cup G') = \mathscr{V}(G)\mathscr{V}(G'), \tag{A.11}$$

where $\cup$ denotes disjoint union. We shall now prove that, if G is a simple (open or closed) path of length l, we have

$$\mathscr{V}(G) \le \left(\frac{1}{N^2}\sum_{i,j} |B_{ij}|^2\right)^{l/2}. \tag{A.12}$$

Using (A.10), (A.11) and (A.12), we find that

$$\mathscr{V}(G(\Gamma)) \le \left(\frac{B_o}{q}\right)^s \left(\frac{1}{N^2}\sum_{i,j} |B_{ij}|^2\right)^{(2r-s)/2}. \tag{A.13}$$

Let us now prove (A.12). We start with a simple closed path of length l, whose value (A.8) is given by

$$\mathcal{C}_l := \frac{1}{N^l} \sum_{i_1, \dots, i_l} |B_{i_1 i_2}| \cdots |B_{i_{l-1} i_l}||B_{i_l i_1}|.$$

Assume first that $l = 2k$ is even. Then

$$\begin{aligned}
\mathcal{C}_{2k} &\le \frac{1}{N^{2k}}\left(\sum_{i_1, \dots, i_{2k}} |B_{i_1 i_2}|^2 |B_{i_3 i_4}|^2 \cdots |B_{i_{2k-1} i_{2k}}|^2\right)^{1/2} \\
&\quad \times \left(\sum_{i_1, \dots, i_{2k}} |B_{i_2 i_3}|^2 |B_{i_4 i_5}|^2 \cdots |B_{i_{2k} i_1}|^2\right)^{1/2} \\
&\le \left(\frac{1}{N^2}\sum_{i,j} |B_{ij}|^2\right)^{l/2}.
\end{aligned}$$

If $l = 2k+1$ is odd, we find

$$\begin{aligned}
\mathcal{C}_{2k+1} &= \frac{1}{N^{2k+1}} \sum_{i_1,i_2} |B_{i_1 i_2}| \biggl(\sum_{i_3,\dots,i_{2k+1}} |B_{i_2 i_3}| \cdots |B_{i_{2k+1} i_1}| \biggr) \\
&\le \frac{1}{N^{2k+1}} \biggl(\sum_{i_1,i_2} |B_{i_1 i_2}|^2 \biggr)^{1/2} \\
&\quad \times \biggl(\sum_{i_1,\dots,i_{2k+1}} \sum_{i'_3,\dots,i'_{2k+1}} |B_{i_2 i_3}| \cdots |B_{i_{2k+1} i_1}| |B_{i_2 i'_3}| \cdots |B_{i'_{2k+1} i_1}| \biggr)^{1/2} \\
&\le \biggl(\frac{1}{N^2} \sum_{i,j} |B_{ij}|^2 \biggr)^{1/2} \mathcal{C}_{4k}^{1/2} \le \biggl(\frac{1}{N^2} \sum_{i,j} |B_{ij}|^2 \biggr)^{l/2}.
\end{aligned}$$

This proves (A.12) for closed simple paths. Consider now an open simple path of length l, whose value (A.8) is

$$\mathcal{O}_l := \frac{1}{N^{l+1}} \sum_{i_1,\dots,i_{l+1}} |B_{i_1 i_2}| \cdots |B_{i_l i_{l+1}}|.$$

If $l = 2k$ is even, we get

$$\begin{aligned}
\mathcal{O}_l &\le \frac{1}{N^{2k+1}} \biggl(\sum_{i_1,\dots,i_{2k+1}} |B_{i_1 i_2}| |B_{i_3 i_4}| \cdots |B_{i_{2k-1} i_{2k}}| \biggr)^{1/2} \\
&\quad \times \biggl(\sum_{i_1,\dots,i_{2k+1}} |B_{i_2 i_3}| |B_{i_4 i_5}| \cdots |B_{i_{2k} i_{2k+1}}| \biggr)^{1/2} \\
&\le \biggl(\frac{1}{N^2} \sum_{i,j} |B_{ij}|^2 \biggr)^{l/2}.
\end{aligned}$$

Finally, if $l = 2k+1$ is odd, we find

$$\begin{aligned}
\mathcal{O}_l &\le \frac{1}{N^{2k+2}} \sum_{i_1,i_2} |B_{i_1 i_2}| \biggl(\sum_{i_3,\dots,i_{2k+2}} |B_{i_2 i_3}| \cdots |B_{i_{2k+1} i_{2k+2}}| \biggr) \\
&\le \frac{1}{N^{2k+2}} \biggl(\sum_{i_1,i_2} |B_{i_1 i_2}|^2 \biggr)^{1/2} \\
&\quad \times \biggl(\sum_{i_1,\dots,i_{2k+2}} \sum_{i'_3,\dots,i'_{2k+2}} |B_{i_2 i_3}|^2 \cdots |B_{i_{2k+1} i_{2k+2}}|^2 |B_{i_2 i'_3}|^2 \cdots |B_{i'_{2k+1} i'_{2k+2}}|^2 \biggr)^{1/2} \\
&\le \biggl(\frac{1}{N^2} \sum_{i,j} |B_{ij}|^2 \biggr)^{1/2} \mathcal{O}_{4k}^{1/2} \le \biggl(\frac{1}{N^2} \sum_{i,j} |B_{ij}|^2 \biggr)^{l/2}.
\end{aligned}$$

This concludes the proof of (A.12).

Thus, we get from (A.13) that the contribution to the right-hand side of (A.7) of all label configurations $\mathbf{i}$ satisfying $\Gamma(\mathbf{i}) = \Gamma$ is bounded in absolute value by

$$C^r \biggl(\frac{B_o}{q} + \biggl(\frac{1}{N^2} \sum_{i \neq j} |B_{ij}|^2 \biggr)^{1/2} \biggr)^{2r}.$$

In order to conclude the proof of (A.4), we need a combinatorial bound on the number of multigraphs of the above type containing $2r$ edges, as well as on the number of partitions Γ associated with any given multigraph G. Their product is easily seen to be bounded by $(Cr)^{4r}$. This completes the proof of (A.4). $\square$

PROOF OF LEMMA 3.8. The proof is a simple application of Lemma A.1 and Markov's inequality.

In order to prove (i), we choose $p = \nu(\log N)^\xi$ in (A.2) and apply a high moment Markov inequality.

Next, we prove (ii). The bound (3.19) follows immediately from (i) by setting $\alpha = 1$, $\beta = -2$ and $\gamma = 1$. Similarly, the bound (3.20) follows easily by applying (i) to the random variables $|a_i|^2 - \sigma_i^2$ and setting $A_i = B_{ii}$; here $\alpha = 2$, $\beta = -2$ and $\gamma = 1$, as can be easily seen using (3.18). Moreover, the claim (3.21) follows by setting $p = \nu(\log N)^\xi$ in (A.4) and applying a high moment Markov inequality.

Finally, we prove (iii). Write

$$\biggl| \sum_{i,j} a_i B_{ij} b_j \biggr| \leq \biggl| \sum_i a_i B_{ii} b_i \biggr| + \biggl| \sum_{i \neq j} a_i B_{ij} b_j \biggr|.$$

The first term is dealt with by noting that the random variables $a_1 b_1, \ldots, a_N b_N$ are independent and satisfy (3.16) for $\alpha = 2$, $\beta = -4$ and $\gamma = 2$. Therefore, (3.17) yields with (ξ, ν)-high probability

$$\begin{aligned} \biggl| \sum_i a_i B_{ii} b_i \biggr| &\leq (\log N)^\xi \biggl[\frac{B_d}{q^2} + \biggl(\frac{1}{N^2} \sum_i |B_{ii}|^2 \biggr)^{1/2} \biggr] \\ &\leq 2(\log N)^\xi \frac{B_d}{q^2}. \end{aligned}$$

In order to bound the off-diagonal terms, we set $A_i := \sum_{j \neq i} B_{ij} b_j$. Then we may again apply (3.17) to get with (ξ, ν)-high probability

$$|A_i| \leq (\log N)^\xi \biggl[\frac{B_o}{q} + \biggl(\frac{1}{N} \sum_{j \neq i} |B_{ij}|^2 \biggr)^{1/2} \biggr].$$

Since A_i is independent of a_j, we therefore get from (3.17)

$$\begin{aligned}\left|\sum_{i\neq j} a_i B_{ij} b_j\right| &= \left|\sum_i A_i a_i\right| \\ &\leq (\log N)^{\xi}\left[\frac{\max_i |A_i|}{q} + \left(\frac{1}{N}\sum_i |A_i|^2\right)^{1/2}\right] \\ &\leq C(\log N)^{2\xi}\left[\frac{B_o}{q} + \left(\frac{1}{N^2}\sum_{i\neq j}|B_{ij}|^2\right)^{1/2}\right]\end{aligned}$$

with (ξ, ν)-high probability. We remark finally that the constant C may be absorbed into the small constant ν when applying the high moment Markov inequality used to prove (3.17). □

PROOF OF LEMMA 6.5. To prove (6.19) for $k = 1$, we estimate with (ξ, ν)-high probability

$$|\langle \mathbf{e}, H\mathbf{e}\rangle| = \left|\frac{1}{N}\sum_{i,j} h_{ij}\right| = O((\log N)^{\xi} N^{-1/2}),$$

where we invoked (2.5) and applied (3.17) to the $O(N^2)$ variables $\{h_{ij} : i < j\}$ (and similarly for $i \geq j$) with $\alpha = 1$, $\beta = -2$ and $\gamma = 1$.

If $k \geq 2$, we use a high moment expansion. The following notation will prove helpful. We abbreviate $\alpha = (i, j)$ and write $h_\alpha := h_{ij}$. Defining

$$B_{(i,j)(k,l)} := \delta_{jk},$$

we may thus write

$$\mathbb{IE}\langle \mathbf{e}, H^k\mathbf{e}\rangle = \frac{1}{N}\sum_{\alpha_1,\dots,\alpha_k} B_{\alpha_1\alpha_2}B_{\alpha_2\alpha_3}\cdots B_{\alpha_{k-1}\alpha_k}\mathbb{IE}(h_{\alpha_1}\cdots h_{\alpha_k}), \tag{A.14}$$

where $\mathbb{IE}(\cdot) := (\cdot) - \mathbb{E}(\cdot)$. In order to make all matrix entries independent of each other, we split $H = H' + H''$ into two triangular matrices, where

$$h'_{ij} := h_{ij}\mathbf{1}(i \leq j), \qquad h''_{ij} := h_{ij}\mathbf{1}(i > j).$$

This results in a splitting of (A.14) into 2^k terms, of which we only consider

$$X_k := \frac{1}{N}\sum_{\alpha_1,\dots,\alpha_k} B_{\alpha_1\alpha_2}B_{\alpha_2\alpha_3}\cdots B_{\alpha_{k-1}\alpha_k}\mathbb{IE}(h'_{\alpha_1}\cdots h'_{\alpha_k})$$

(the other terms are dealt with in exactly the same manner and the resulting factor 2^k is immaterial).

We abbreviate $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)$ and write

$$X_k = \sum_{\boldsymbol{\alpha}}(\zeta_{\boldsymbol{\alpha}} - \mathbb{E}\zeta_{\boldsymbol{\alpha}}),$$

where we defined

$$\zeta_{\boldsymbol{\alpha}} := \frac{1}{N} B_{\alpha_1\alpha_2} B_{\alpha_2\alpha_3} \cdots B_{\alpha_{k-1}\alpha_k} h'_{\alpha_1} \cdots h'_{\alpha_k}.$$

For even $p \in \mathbb{N}$ we get therefore

$$\mathbb{E}X_k^p = \sum_{\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^p} \mathbb{E}[(\zeta_{\boldsymbol{\alpha}^1} - \mathbb{E}\zeta_{\boldsymbol{\alpha}^1}) \cdots (\zeta_{\boldsymbol{\alpha}^p} - \mathbb{E}\zeta_{\boldsymbol{\alpha}^p})]. \tag{A.15}$$

By independence of the family $\{h'_\alpha\}$, we find that a summand in (A.15) indexed by $\boldsymbol{\alpha}$ vanishes if there is an r such that $[\boldsymbol{\alpha}^r] \cap [\boldsymbol{\alpha}^{r'}] = \varnothing$ for all $r' \neq r$. Here $[(\alpha_1, \dots, \alpha_k)] := \{\alpha_1, \dots, \alpha_k\}$. Thus, we find

$$\mathbb{E}X_k^p = \sum_{\boldsymbol{\alpha}^1,\dots,\boldsymbol{\alpha}^p} \mathbb{E}[(\zeta_{\boldsymbol{\alpha}^1} - \mathbb{E}\zeta_{\boldsymbol{\alpha}^1}) \cdots (\zeta_{\boldsymbol{\alpha}^p} - \mathbb{E}\zeta_{\boldsymbol{\alpha}^p})] \chi(\boldsymbol{\alpha}^1, \dots, \boldsymbol{\alpha}^p), \tag{A.16}$$

where

$$\chi(\boldsymbol{\alpha}^1, \dots, \boldsymbol{\alpha}^p) := \prod_{r=1}^p \mathbf{1}(\exists r' : [\boldsymbol{\alpha}^r] \cap [\boldsymbol{\alpha}^{r'}] \neq \varnothing).$$

For each given label configuration $\boldsymbol{\alpha} = (\boldsymbol{\alpha}^r) = (\alpha_l^r)$, we define a partition $\Gamma(\boldsymbol{\alpha})$ of the index set $\{(r,l) : r = 1, \dots, p, l = 1, \dots, k\}$ by imposing that (r,l) and (r',l') are in the same equivalence class of $\Gamma(\boldsymbol{\alpha})$ if and only if $\alpha_{rl} = \alpha_{r'l'}$. We now perform the sum over $\boldsymbol{\alpha}$ in (A.16) by first specifying a partition Γ and summing over all $\boldsymbol{\alpha}$ satisfying $\Gamma = \Gamma(\boldsymbol{\alpha})$, and then summing over all partitions Γ. Note that any partition Γ yielding a nonzero contribution to (A.16) satisfies the two following conditions:

(i) Each equivalence class of Γ contains at least two elements.

(ii) For each $r = 1, \dots, p$ there are $r' = 1, \dots, p$ and $l, l' = 1, \dots, k$ such that (r,l) and (r',l') are in the same equivalence class of Γ.

Condition (i) follows from the fact that h'_α is centered, and condition (ii) from the definition of χ.

Let us fix a partition Γ satisfying (i) and (ii). Its contribution to (A.16) is

$$\begin{aligned} &\left| \sum_{\boldsymbol{\alpha} : \Gamma(\boldsymbol{\alpha}) = \Gamma} \mathbb{E}[(\zeta_{\boldsymbol{\alpha}^1} - \mathbb{E}\zeta_{\boldsymbol{\alpha}^1}) \cdots (\zeta_{\boldsymbol{\alpha}_p} - \mathbb{E}\zeta_{\boldsymbol{\alpha}^p})] \chi(\boldsymbol{\alpha}) \right| \\ &\qquad \le \sum_{\boldsymbol{\alpha} : \Gamma(\boldsymbol{\alpha}) = \Gamma} \mathbb{E}[(|\zeta_{\boldsymbol{\alpha}^1}| + \mathbb{E}|\zeta_{\boldsymbol{\alpha}^1}|) \cdots (|\zeta_{\boldsymbol{\alpha}^p}| + \mathbb{E}|\zeta_{\boldsymbol{\alpha}^p}|)] \chi(\boldsymbol{\alpha}). \end{aligned} \tag{A.17}$$

Next, we note that Γ gives rise to a multigraph $G \equiv G(\Gamma)$ defined as follows. The vertex set $V(G)$ is given by the equivalence classes of Γ. Each pair $\{(r,l), (r,l+1)\}$, $l = 1, \dots, k-1$, gives rise to an edge that connects the vertices $\gamma \ni (r,l)$ and $\gamma' \ni (r,l+1)$. Thus, the set of edges $E(G)$ of G contains $p(k-1)$

edges. The interpretation of the edges is that each factor $B_{\alpha_{r,l}\alpha_{r,l+1}}$ on the right-hand side of (A.17) is represented with an edge.

The expectation on the right-hand side of the identity

$$\mathbb{E}[(|\zeta_{\boldsymbol{\alpha}^1}|+\mathbb{E}|\zeta_{\boldsymbol{\alpha}^1}|)\cdots(|\zeta_{\boldsymbol{\alpha}^p}|+\mathbb{E}|\zeta_{\boldsymbol{\alpha}^p}|)]$$
$$=\frac{1}{N^p}\Biggl[\prod_{l=1}^{k-1}B_{\alpha_l^r\alpha_{l+1}^r}\Biggr]\mathbb{E}\Biggl[\prod_{r=1}^{p}\Biggl(\prod_{l=1}^{p}|h'_{\alpha_l^r}|+\mathbb{E}\prod_{l=1}^{p}|h'_{\alpha_l^r}|\Biggr)\Biggr]$$

is bounded by

$$2^p\prod_{\gamma\in V(G)}\frac{C^{|\gamma|}}{Nq^{|\gamma|-2}},$$

where $|\gamma|$ denotes the size of the equivalence class γ; this is a simple consequence of (2.5) and the constraint $q\le CN^{1/2}$. By construction of G, each vertex γ of G carries a label α_γ. Thus, we may bound (A.17) by

$$\frac{2^p}{N^p}\sum_{\alpha_1,\dots,\alpha_v}\Biggl[\prod_{\{\gamma,\gamma'\}\in E(G)}B_{\alpha_\gamma\alpha_{\gamma'}}\Biggr]\prod_{\gamma\in V(G)}\frac{C^{|\gamma|}}{Nq^{|\gamma|-2}},$$

where $v=|V(G)|$ denotes the number of vertices of G. Here we dropped the factor χ, and the restriction that $\alpha_1,\dots,\alpha_v$ be distinct, to obtain an upper bound. By property (i) above, we have $|\gamma|-2\ge 0$ and we get the bound

$$\frac{2^pC^{pk}}{N^{p+v}}\sum_{\alpha_1,\dots,\alpha_v}\prod_{\{\gamma,\gamma'\}\in E(G)}B_{\alpha_\gamma\alpha_{\gamma'}}. \tag{A.18}$$

Next, we split $G=G_1\cup\cdots\cup G_l$ into its connected components; here l denotes the number of connected components. An immediate consequence of the property (ii) of Γ is the bound

$$l\le p/2. \tag{A.19}$$

Thus, (A.18) becomes

$$\frac{C^{pk}}{N^{p+v}}\prod_{j=1}^{l}\Biggl[\sum_{\alpha_1,\dots,\alpha_{v_j}}\prod_{\{\gamma,\gamma'\}\in E(G_j)}B_{\alpha_\gamma\alpha_{\gamma'}}\Biggr], \tag{A.20}$$

where $v_j=|V(G_j)|$ denotes the number of vertices in G_j.

In order to estimate the contribution of the jth connected component, we pick a root $r_j\in V(G_j)$ and a spanning tree T_j of G_j. First, we use the trivial bound $B_{\alpha_\gamma\alpha_{\gamma'}}\le 1$ for edges that do not belong to T_j. Second, we sum over all of the v_j-1 nonroot labels α_γ, starting from the leaves of T_j, and using the identity

$$\sum_{\alpha_{\gamma'}}B_{\alpha_\gamma\alpha_{\gamma'}}=N$$

at each step. Third, we sum over the root label γ_{r_j}, which yields a factor bounded by N^2. Putting everything together yields

$$\sum_{\alpha_1,\dots,\alpha_{v_j}} \prod_{\{\gamma,\gamma'\}\in E(G_j)} B_{\alpha_\gamma \alpha_{\gamma'}} \le N^{v_j+1}.$$

Returning to (A.20), we thus find that the right-hand side of (A.17) is bounded by

$$\frac{C^{pk}}{N^{p+v}} N^{v+l} \le \frac{C^{pk}}{N^{p/2}},$$

where we used (A.19).

Since the number of partitions Γ is bounded by $(kp)^{kp}$, we get the bound

$$\mathbb{E} X_k^p \le \left(\frac{(Ckp)^k}{N^{1/2}}\right)^p.$$

Choosing $p = \frac{1}{2Ck}(\log N)^\xi$ and applying a high moment Markov inequality completes the proof. □

PROOF OF LEMMA 7.10. The proof is similar to (in fact, considerably simpler than) the proof of Lemma 6.5. We only sketch the argument, using the notation of the proof of Lemma 6.5 without further comment. Write

$$X_k := \sum_{i_1,\dots,i_k} h_{ii_1} h_{i_1 i_2} \cdots h_{i_{k-1} i_k} = \sum_{\alpha_1,\dots,\alpha_k} B_{\alpha_0 \alpha_1} B_{\alpha_1 \alpha_2} \cdots B_{\alpha_{k-1}\alpha_k} h_{\alpha_1} \cdots h_{\alpha_k},$$

where $\alpha_0 := (1, i)$. Then, as in the proof of Lemma 6.5, we write $\mathbb{E} X_k^p$ as a sum over partitions Γ which give rise to multigraphs $G \equiv G(\Gamma)$ whose edges are given by the factors B and whose vertices are given by equivalence classes γ of the set $\{1,\dots,k\} \times \{1,\dots,p\}$, to which has been adjoined a distinguished vertex γ_0. The vertex γ_0 corresponds to the fixed label α_0, and it has degree p. Each multigraph G has pk edges, and is connected. In this fashion we find that the contribution of the multigraph G to $\mathbb{E} X_k^p$ is bounded by

$$\sum_{(\alpha_\gamma)_{\gamma \ne \gamma_0}} \left[\prod_{\{\gamma,\gamma'\}\in E(G)} B_{\alpha_\gamma \alpha_{\gamma'}}\right] \prod_{\gamma \in V(G)\setminus\{\gamma_0\}} \mathbb{E}|h_{\alpha_\gamma}|^{|\gamma|}, \tag{A.21}$$

where the first sum ranges over families $(\alpha_\gamma)_{\gamma \in V(G)\setminus\{\gamma_0\}}$ of labels; every vertex $\gamma \ne \gamma_0$ carries a label α_γ which is summed over. The vertex γ_0 carries the label α_0 which is fixed.

Since $\mathbb{E} h_\alpha = 0$, it is easy to see that $|\gamma| \ge 2$ for all γ. Choosing a spanning tree of the connected graph G, one therefore finds that (A.21) is bounded by

$$N^{|V(G)|-1} \prod_{\gamma \in V(G)\setminus\{\gamma_0\}} \left(\max_\alpha \mathbb{E}|h_\alpha|^2\right) = 1.$$

Since the number of partitions Γ is bounded by $(kp)^{kp}$, we find $\mathbb{E}X_k^p \le (Ckp)^{kp}$ for $p \le (\log N)^\xi$. Choosing $p = \frac{1}{2Ck}(\log N)^\xi$ and applying Markov's inequality completes the proof. $\square$

PROOF OF LEMMA 4.3. Our proof is a standard application of the moment method, along the lines of [22], Lemma 7.2.

In a first step, we truncate the entries h_{ij}. Let $C_1 = C$ be a constant for which Lemma 3.7 holds. Define

$$\mu_{ij} := \mathbb{E}h_{ij}\mathbf{1}(|h_{ij}| \le C_1 q^{-1}).$$

Choose an independent family (X_{ij}) of random variables, independent of H, such that

$$\mathbb{P}(X_{ij} = q^{-1}) = \mu_{ij}q, \qquad \mathbb{P}(X_{ij} = 0) = 1 - \mu_{ij}q.$$

Now set

$$\widehat{h}_{ij} := h_{ij}\mathbf{1}(|h_{ij}| \le C_1 q^{-1}) - X_{ij}.$$

It is easy to see that $|\mu_{ij}| \le \mathrm{e}^{-\nu(\log N)^\xi}$ and, therefore,

$$\mathbb{P}(h_{ij} \ne \widehat{h}_{ij}) \le \mathrm{e}^{-\nu(\log N)^\xi}. \tag{A.22}$$

Moreover, we have

$$\mathbb{E}\widehat{h}_{ij} = 0, \qquad |\widehat{h}_{ij}| \le \frac{C_1+1}{q}, \qquad \mathbb{E}|\widehat{h}_{ij}|^2 \le \frac{1}{N}\big(1 + \mathrm{e}^{-\nu(\log N)^\xi}\big). \tag{A.23}$$

By (A.22), it suffices to prove that $\|\widehat{H}\| \le 2 + (\log N)^\xi q^{-1/2}$ with (ξ, ν)-high probability. We shall prove that, for even $k \le c\sqrt{q}$, we have

$$|\mathbb{E}\operatorname{Tr}\widehat{H}^k| \le 3Nk2^k. \tag{A.24}$$

In order to prove (A.24), we write

$$\mathbb{E}\operatorname{Tr}\widehat{H}^k = \sum_{i_1,\dots,i_k} \mathbb{E}\widehat{h}_{i_1i_2}\cdots\widehat{h}_{i_{k-1}i_k}\widehat{h}_{i_ki_1} \tag{A.25}$$

and apply a graphical expansion to the right-hand side. Before giving its precise definition, we outline how it arises from (A.25). Let the label configuration $i_1, \dots, i_k$ be fixed. We represent each index $j = 1, \dots, k$ by a vertex $[j]$, whereby two indices j and j' correspond to the same vertex if their labels agree, $i_j = i_{j'}$. Let p be the number of vertices. We then construct a closed walk through the sequence of edges $([1], [2]), ([2], [3]), \dots, ([k], [1])$. The walk has k steps. We name the p vertices $1, \dots, p$, whereby vertex v is reached after all vertices $1, \dots, v-1$. Since $\mathbb{E}\widehat{h}_{ij} = 0$, it is easy to see from (A.25) that each edge of the walk must appear at least twice.

We may now give a precise definition of such walks. Let $\mathbf{w} = (w_1, \ldots, w_k)$ be a sequence with $w_v \in \{1, \ldots, p\}$. With $\mathbf{w}$ we associate a multigraph $G(\mathbf{w})$ as follows. The vertex set of $G(\mathbf{w})$ is $\{1, \ldots, p\}$; the edge set of $G(\mathbf{w})$ is given by the undirected edges $\{w_1, w_2\}, \ldots, \{w_{k-1}, w_k\}, \{w_k, w_1\}$. [Note that $G(\mathbf{w})$ may contain multiple edges as well as loops.] We say that $\mathbf{w}$ is an *ordered closed walk* of length k on p vertices if:

(i) A vertex that is visited for the first time at time j is greater than all vertices visited before time $j : \max_{j' \le j} w_{j'} \le \max_{j' < j} w_{j'} + 1$.
(ii) All vertices are visited: $\{w_1, \ldots, w_k\} = \{1, \ldots, p\}$.
(iii) Every edge of $G(\mathbf{w})$ appears at least twice.

Let $\mathcal{W}(k, p)$ denote the set of ordered closed walks of length k on p vertices. The key combinatorial estimate of our proof is the bound

$$|\mathcal{W}(k, p)| \le \binom{k}{2p-2} p^{2(k-2p+2)} 2^{2p-2},$$

proved in [39]. Using the notion of ordered closed walks, it is not hard to see that (A.25) may be rewritten as

$$\mathbb{E} \operatorname{Tr} \widehat{H}^k = \sum_{p=1}^{k/2+1} \sum_{\mathbf{w} \in \mathcal{W}(k,p)} \sum_{\boldsymbol{\ell} \in \mathcal{L}(p)} \mathbb{E} \widehat{h}_{\ell(w_1)\ell(w_2)} \cdots \widehat{h}_{\ell(w_{k-1})\ell(w_k)} \widehat{h}_{\ell(w_k)\ell(w_1)}, \tag{A.26}$$

where $\mathcal{L}(p)$ is the set of all p-tuples $\boldsymbol{\ell} = (\ell(1), \ldots, \ell(p)) \in \{1, \ldots, N\}^p$ whose components are disjoint. See [22], Section 7.1, for a detailed proof.

Next, associate with the multigraph $G(\mathbf{w})$ its *skeleton* $S(\mathbf{w})$, obtained from $G(\mathbf{w})$ by discarding the multiplicity of every edge (i.e., by successively removing edges until it has no multiple edges). For $e \in E(S(\mathbf{w}))$, we denote by $\nu(e)$ the multiplicity of the edge e in $G(\mathbf{w})$. We have the obvious relation $\sum_{e \in E(S(\mathbf{w}))} \nu(e) = k$. If $e = \{v, v'\}$, we write $\ell(e) := (\ell(v), \ell(v'))$ [the chosen order of the pair $\ell(e)$ is immaterial]. Then it is easy to see that

$$\begin{aligned} &\bigl|\mathbb{E} \widehat{h}_{\ell(w_1)\ell(w_2)} \cdots \widehat{h}_{\ell(w_{k-1})\ell(w_k)} \widehat{h}_{\ell(w_k)\ell(w_1)}\bigr| \\ &\quad \le \prod_{e \in E(S(\mathbf{w}))} \mathbb{E} \bigl|\widehat{h}_{\ell(e)}\bigr|^{\nu(e)} \\ &\quad \le \prod_{e \in E(S(\mathbf{w}))} \frac{1}{N} \bigl(1 + \mathrm{e}^{-\nu (\log N)^{\xi}}\bigr) \frac{C^{\nu(e)-2}}{q^{\nu(e)-2}} \\ &\quad \le \biggl[\frac{q^2}{C^2 N} \bigl(1 + \mathrm{e}^{-\nu (\log N)^{\xi}}\bigr)\biggr]^{E_S} \biggl(\frac{C}{q}\biggr)^k, \end{aligned}$$

where we used (A.23) and introduced the shorthand $E_S := |E(S(\mathbf{w}))|$. Therefore, summing over $\ell \in \mathcal{L}(p)$ in (A.26) yields

$$|\mathbb{E} \operatorname{Tr} \widehat{H}^k| \le \sum_{p=1}^{k/2+1} \sum_{\mathbf{w} \in \mathcal{W}(k,p)} N^p \biggl[\frac{q^2}{C^2 N} \bigl(1 + \mathrm{e}^{-c (\log N)^{\xi}}\bigr)\biggr]^{E_S} \biggl(\frac{C}{q}\biggr)^k.$$

Next, it is immediate that we have the relations $p - 1 \le E_S \le k/2$; these inequalities follow from the above properties (ii) and (iii), respectively. Since $d^2/N \ll 1$, we therefore get

$$|\mathbb{E}\operatorname{Tr}\widehat{H}^k| \le N \sum_{p=1}^{k/2+1} |\mathcal{W}(k,p)| \big(1 + \mathrm{e}^{-c(\log N)^\xi}\big)^p \left(\frac{C}{q}\right)^{k-2p+2}.$$

For $k \le N$ this yields

$$|\mathbb{E}\operatorname{Tr}\widehat{H}^k| \le 3N \sum_{p=1}^{k/2+1} S(k,p),$$

where

$$S(k,p) := \binom{k}{2p-2} p^{2(k-2p+2)} 2^{2p-2} \left(\frac{C}{q}\right)^{k-2p+2}.$$

It is elementary to check that $S(k, k/2+1) = 2^k$ and

$$S(k,p) \le \frac{k^4}{8}\left(\frac{C}{q}\right)^2 S(k,p+1).$$

Therefore, choosing $k \le c\sqrt{q}$ implies $S(k,p) \le 2^k$. This concludes the proof of (A.24).

The claim now follows by setting $k = c\sqrt{q}$ with a sufficiently small constant c, applying a high moment Markov inequality and recalling that $\sqrt{q} \gg (\log N)^\xi$ by (2.6). □

REFERENCES

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

[2] BAI, Z. D., MIAO, B. and TSAY, J. (2002). Convergence rates of the spectral distributions of large Wigner matrices. *Int. Math. J.* **1** 65–90. MR1825933

[3] BENAYCH-GEORGES, F. and NADAKUDITI, R. R. (2011). The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* **227** 494–521. MR2782201

[4] BLEHER, P. and ITS, A. (1999). Semiclassical asymptotics of orthogonal polynomials, Riemann–Hilbert problem, and universality in the matrix model. *Ann. of Math.* (2) **150** 185–266. MR1715324

[5] CAPITAINE, M., DONATI-MARTIN, C. and FÉRAL, D. (2009). The largest eigenvalues of finite rank deformation of large Wigner matrices: Convergence and nonuniversality of the fluctuations. *Ann. Probab.* **37** 1–47. MR2489158

[6] DEIFT, P. and GIOEV, D. (2009). *Random Matrix Theory*: *Invariant Ensembles and Universality*. *Courant Lecture Notes in Mathematics* **18**. Amer. Math. Soc., Providence, RI. MR2514781

[7] DEIFT, P., KRIECHERBAUER, T., MCLAUGHLIN, K. T. R., VENAKIDES, S. and ZHOU, X. (1999). Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.* **52** 1335–1425. MR1702716

[8] DEIFT, P., KRIECHERBAUER, T., MCLAUGHLIN, K. T.-R., VENAKIDES, S. and ZHOU, X. (1999). Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.* **52** 1491–1552. MR1711036

[9] DEIFT, P. A. (1999). *Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach. Courant Lecture Notes in Mathematics* **3**. Amer. Math Soc., Providence, RI. MR1677884

[10] DEKEL, Y., LEE, R. L. and LINIAL, N. (2011). Eigenvectors of random graphs: Nodal domains. *Random Structures Algorithms* **39** 39–58. MR2839984

[11] DUMITRIU, I. and PAL, S. (2011). Sparse regular random graphs: Spectral density and eigenvectors. Preprint. Available at arXiv:0910.5306.

[12] DYSON, F. J. (1962). A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3** 1191–1198. MR0148397

[13] ERDŐS, L., KNOWLES, A., YAU, H. T. and YIN, J. (2012). Spectral statistics of Erdős–Rényi graphs II: Eigenvalue spacing and the extreme eigenvalues. *Comm. Math. Phys.* **314** 587–640. MR2964770

[14] ERDŐS, L., PÉCHÉ, S., RAMÍREZ, J. A., SCHLEIN, B. and YAU, H.-T. (2010). Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63** 895–925. MR2662426

[15] ERDŐS, L., RAMÍREZ, J. A., SCHLEIN, B. and YAU, H.-T. (2010). Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation. *Electron. J. Probab.* **15** 526–603. MR2639734

[16] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2009). Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.* **37** 815–852. MR2537522

[17] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2009). Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* **287** 641–655. MR2481753

[18] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2010). Wegner estimate and level repulsion for Wigner random matrices. *Int. Math. Res. Not. IMRN* **3** 436–479. MR2587574

[19] ERDŐS, L., SCHLEIN, B. and YAU, H.-T. (2011). Universality of random matrices and local relaxation flow. *Invent. Math.* **185** 75–119. MR2810797

[20] ERDŐS, L., SCHLEIN, B., YAU, H. T. and YIN, J. (2012). The local relaxation flow approach to universality of the local statistics for random matrices. *Ann. Inst. H. Poincaré Probab. Statist.* **48** 1–46.

[21] ERDŐS, L., YAU, H.-T. and YIN, J. (2011). Universality for generalized Wigner matrices with Bernoulli distribution. *J. Comb.* **2** 15–81. MR2847916

[22] ERDŐS, L., YAU, H. T. and YIN, J. (2012). Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields* **154** 341–407. MR2981427

[23] ERDŐS, L., YAU, H.-T. and YIN, J. (2012). Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math.* **229** 1435–1515. MR2871147

[24] ERDŐS, P. and RÉNYI, A. (1959). On random graphs. I. *Publ. Math. Debrecen* **6** 290–297. MR0120167

[25] ERDŐS, P. and RÉNYI, A. (1960). On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** 17–61. MR0125031

[26] FÉRAL, D. and PÉCHÉ, S. (2007). The largest eigenvalue of rank one deformation of large Wigner matrices. *Comm. Math. Phys.* **272** 185–228. MR2291807

[27] GAUDIN, M. (1961). Sur la loi limite de l'espacement des valeurs propres d'une matrice aléatoire. *Nuclear Phys.* **25** 447–458.

[28] GUIONNET, A. (2009). *Large Random Matrices*: *Lectures on Macroscopic Asymptotics*. *Lecture Notes in Math*. **1957**. Springer, Berlin. MR2498298

[29] GUIONNET, A. and ZEITOUNI, O. (2000). Concentration of the spectral measure for large matrices. *Electron*. *Commun*. *Probab*. **5** 119–136 (electronic). MR1781846

[30] MEHTA, M. L. (1991). *Random Matrices*, 2nd ed. Academic Press, Boston, MA. MR1083764

[31] PASTUR, L. and SHCHERBINA, M. (2008). Bulk universality and related properties of Hermitian matrix models. *J*. *Stat*. *Phys*. **130** 205–250. MR2375744

[32] PIZZO, A., RENFREW, D. and SOSHNIKOV, A. (2013). On finite rank deformations of Wigner matrices. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **49** 64–94.

[33] SINAĬ, Y. G. and SOSHNIKOV, A. B. (1998). A refinement of Wigner's semicircle law in a neighborhood of the spectrum edge for random symmetric matrices. *Funct*. *Anal*. *Appl*. **32** 114–131.

[34] SODIN, S. (2010). The spectral edge of some random band matrices. *Ann*. *of Math*. (2) **172** 2223–2251. MR2726110

[35] SOSHNIKOV, A. (1999). Universality at the edge of the spectrum in Wigner random matrices. *Comm*. *Math*. *Phys*. **207** 697–733. MR1727234

[36] TAO, T. and VU, V. (2010). Random matrices: Universality of local eigenvalue statistics up to the edge. *Comm*. *Math*. *Phys*. **298** 549–572. MR2669449

[37] TAO, T. and VU, V. (2011). Random matrices: Universality of local eigenvalue statistics. *Acta Math*. **206** 127–204. MR2784665

[38] TRAN, L., VU, V. and WANG, K. (2010). Sparse random graphs: Eigenvalues and eigenvectors. Preprint. Available at arXiv:1011.6646.

[39] VU, V. H. (2007). Spectral norm of random matrices. *Combinatorica* **27** 721–736. MR2384414

[40] WIGNER, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Ann*. *of Math*. (2) **62** 548–564. MR0077805

L. ERDŐS
INSTITUTE OF MATHEMATICS
UNIVERSITY OF MUNICH
THERESIENSTRASSE 39
D-80333 MUNICH
GERMANY
E-MAIL: lerdos@math.lmu.de

A. KNOWLES
H.-T. YAU
DEPARTMENT OF MATHEMATICS
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS 02138
USA
E-MAIL: knowles@math.harvard.edu
htyau@math.harvard.edu

J. YIN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706
USA
E-MAIL: jyin@math.wisc.edu