

Spectral statistics of sparse Erdős–Rényi graph Laplacians

Jiaoyang Huang^a and Benjamin Landon^b

^a*Department of Mathematics, Harvard University, Cambridge, MA, USA. E-mail: jiaoyang@math.harvard.edu*

^b*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA. E-mail: blandon@mit.edu*

Received 8 March 2016; revised 28 November 2018; accepted 13 January 2019

Abstract. We consider the bulk eigenvalue statistics of Laplacian matrices of large Erdős–Rényi random graphs in the regime $p \geq N^\delta/N$ for any fixed $\delta > 0$. We prove a local law down to the optimal scale $\eta \gtrsim N^{-1}$ which implies that the eigenvectors are delocalized. We consider the local eigenvalue statistics and prove that both the gap statistics and averaged correlation functions coincide with the GOE in the bulk.

Résumé. Nous nous intéressons aux statistiques, dans l'intérieur du spectre, des valeurs propres de matrices laplaciennes de grands graphes d'Erdős–Rényi aléatoires dans le régime où $p \geq N^\delta/N$ pour un $\delta > 0$ fixé arbitraire. Nous montrons une loi locale jusqu'à l'échelle optimale $\eta \gtrsim N^{-1}$ qui implique que les vecteurs propres sont délocalisés. Nous considérons les statistiques locales des valeurs propres et montrons que les statistiques des intervalles et les fonctions de corrélation moyennées coïncident avec le GOE dans l'intérieur du spectre.

MSC: 15B52

Keywords: Random matrix theory; Universality

1. Introduction

In the late 1950s Erdős and Rényi introduced what is perhaps the simplest random graph model, the now well-known Erdős–Rényi random graph $\mathcal{G}(N, p)$. This model is a graph on N vertices formed by choosing each (undirected) edge (i, j) independently with probability p . We will say that the model is sparse when $p \rightarrow 0$ as $N \rightarrow \infty$ as each vertex has on average $pN \ll N$ edges.

Random graphs have numerous applications in the modeling of complex physical, biological and social systems. A common method to understand the structure of the underlying graph is to associate to the graph a matrix, the most intensively studied being the adjacency and Laplacian matrices. The spectral properties of these matrices carry a wealth of information about the underlying graph. For example, the eigenvalues of the adjacency matrix reflect the topological features of the corresponding graphs, such as connectivity, while the eigenvalues and eigenvectors of the Laplacian matrix can be used to study random walks on the graph. For a comprehensive introduction to the spectral properties of random graphs we refer the interested reader to [16].

Our main interest in this paper is the spectral properties of the Laplacian of the Erdős–Rényi random graph. There is an enormous literature on the study of the Laplacian of general random graphs and we do not attempt to give an exhaustive list of existing work, only mentioning some results on spectral properties related to ours. The limiting spectral measure of the non-sparse Erdős–Rényi Laplacian was first derived in [12], and the sparse case was obtained in [33]. Laws of large numbers of extreme eigenvalues of the Erdős–Rényi Laplacian were proved in [18,34] in the non-sparse case. For the spectral properties of the Laplacian matrices of more general random graphs, we refer the reader to, e.g., [13–15,17,38].

Random matrix theory provides a natural framework for studying the spectral properties of random graphs. One of the most well-studied models is the Wigner ensemble consisting of matrices with iid entries (up to the symmetry constraint $H = H^*$), introduced by Wigner [41] to model the spectra of heavy ions. The Wigner–Dyson–Gaudin–Mehta conjecture asserts that the local spectral statistics of Wigner matrices exhibit universality and are independent of the underlying

distribution of the matrix entries, and instead depend only on the symmetry class (real symmetric or complex Hermitian) of the ensemble. In particular, the local statistics coincide with the case in which the entries are Gaussians for which there are explicit formulas. This conjecture has recently been solved for all symmetry classes in [10,23,24,27,29,30] (parallel results in certain cases were obtained in [39,40]).

The adjacency matrix of a non-sparse Erdős–Rényi graph (i.e., p fixed independent of N) essentially falls into the Wigner class and so the results [10,23,24,27,29,30] provide detailed information of the local spectral statistics of this ensemble. The sparse case can be viewed as a singular Wigner ensemble, as the law of the entries is highly concentrated around 0, and its analysis is much harder than the non-sparse case. Universality of the local statistics of the sparse Erdős–Rényi adjacency matrix was achieved in the regime $p \geq N^{2/3+\delta}/N$ in the works [20,22]; universality (of the bulk statistics) in the regime $p \geq N^\delta/N$ was only achieved quite recently in [32].

A key property of the Wigner ensemble is that all of the entries are independent (up to the symmetry constraint $H = H^*$). In this regard, the new major challenge in the analysis of the Erdős–Rényi Laplacian is the correlation between the entries, an obstacle which is present even in the non-sparse case. To our knowledge, detailed information about local spectral statistics of (non-invariant) matrix ensembles with correlated entries has been limited to only a few cases. A local law was proved for sparse d -regular graphs in [19] down to the scale $\eta \gtrsim \log(N)^{-1}$. A local law down to the optimal scale was achieved for the d -regular random graph in the regime $d \gtrsim (\log N)^4$ in the work [5] and universality of the local spectral statistics was subsequently obtained in the regime $d \gtrsim N^\varepsilon$ in [4]. A local law was obtained for matrices having with correlated entries having a certain four-fold symmetry in [2] and universality for Gaussian matrices with a translation invariant correlation structure was obtained in [1].

Our first contribution is to prove a local law for the Erdős–Rényi Laplacian down to the optimal scale. More precisely we prove that the Stieltjes transform of the empirical eigenvalue distribution is well-approximated by the free convolution of the semicircle law and a standard Gaussian down to the scale $\eta \gtrsim N^{-1}$. This implies a rigidity estimate for any fixed fraction of the non-extremal eigenvalues and delocalization of the corresponding eigenvectors. Our proof of the local law is inspired by methods developed for symmetric matrices with uncorrelated entries [21,22,25,26,31,36]. However, these proofs rely on the fact that the i th row and column of the matrix is independent of i th minor. This difficulty was overcome in the work [5] on the d -regular graph using methods exploiting the specific structure of the underlying random graph, and it is not clear that these methods would be applicable to the general class of models we will consider in this paper. We instead use the fact that the correlation between the matrix entries is quite weak and implement a high-order resolvent expansion in order to estimate high moments of certain error terms arising in the proof. We introduce a graphical notation to keep track of terms appearing in our estimates inspired by the methods of [9]. Our expansion shows that the correlation between the i th minor and the i th row can be ignored.

The solution of the Wigner–Dyson–Gaudin–Mehta conjecture saw the development of a general three-step strategy for proving bulk universality for random matrix ensembles. The first step is to prove a local law controlling the eigenvalue density down to the optimal scale. The second step concerns the proof of the optimal time to local equilibrium of Dyson Brownian motion (DBM), and the third step is either a comparison or perturbative argument comparing the random matrix ensemble to DBM. Our result discussed above completes the first step.

In the second step we introduce a modification of the DBM, defining a matrix process that preserves the structure of the Laplacian matrix. Note that the Laplacian matrix has a trivial eigenvalue corresponding to the constant eigenvector. Once we restrict the matrix process to the space orthogonal to this eigenvector, we get an ensemble that is well approximated by the sum of a diagonal matrix and a GOE matrix. Universality for this process then follows from the recent work [35], where the optimal time to local equilibrium was established for a wide class of initial data; we remark that the result [35] is required to deal with the sparse ensembles considered here, and previous approaches along the lines of [27,37] are insufficient (see [28] for related results on DBM with general initial data). The work [35] introduces a local approach to DBM, made possible by a coupling idea of [10]. Previous approaches [27,37] were based on hydrodynamics and moreover require estimates on the local eigenvalue density which are stronger than those found in the case of sparse random matrices.

In order to complete the third step we rely on the observation in [11] that the change of the eigenvalues is negligible for a short time if one views DBM as a matrix process. As a result we obtain bulk universality for the Erdős–Rényi Laplacian in the sense of gap universality and averaged correlation functions in the regime $p \geq N^\delta/N$.

Our analysis applies to a wider class of random matrices that includes the Erdős–Rényi Laplacian as a specific example. In Section 2 we introduce this model which we call the random Laplacian-type matrix and state our main results. In Section 3 we prove the local law for random Laplacian-type matrices. In Section 4 we prove universality of the local statistics.

2. Definition of model and main results

2.1. Definition of model

In this section we introduce the class of random matrices under consideration which we call random Laplacian-type matrices. The motivating example is the graph Laplacian of the Erdős–Rényi random graph. The Erdős–Rényi graph on $N + 1$ vertices is constructed by choosing each undirected edge (i, j) , $i \neq j$ independently with probability p (for reasons that will become clear later it is more natural to consider a graph on $N + 1$ vertices instead of N vertices. As a result our matrices will be of size $(N + 1) \times (N + 1)$). It will be convenient to introduce the *sparsity parameter* q by

$$p := \frac{q^2}{N} \quad (2.1)$$

We denote the Laplacian of the Erdős–Rényi random graph by $M = (M_{ij})_{ij}$. The off-diagonal entries are given by

$$M_{ij} = m_{ij}, \quad 1 \leq i \neq j \leq N + 1 \quad (2.2)$$

where the m_{ij} are independent Bernoulli random variables (up to the symmetry constraint $m_{ij} = m_{ji}$) satisfying

$$\mathbb{P}(m_{ij} = 1) = \frac{q^2}{N}, \quad \mathbb{P}(m_{ij} = 0) = 1 - \frac{q^2}{N}. \quad (2.3)$$

The diagonal entries of M are given by

$$M_{ii} = - \sum_{1 \leq k \leq N+1, k \neq i} m_{ki}. \quad (2.4)$$

Note that on average each row has $pN = q^2$ nonzero elements and so the matrix is sparse in the regime $q = o(N^{1/2})$. In order to study the spectral statistics of M we will extract the mean and introduce a rescaling so that the typical eigenvalue spacing in the bulk is of the order N^{-1} . We write

$$M = q(1 - q^2/N)^{1/2} H + q^2 \mathbf{e} \mathbf{e}^* - q^2(N + 1)I_{N+1}. \quad (2.5)$$

Here, we have introduced the $(N + 1) \times 1$ column vector

$$\mathbf{e} = ((N + 1)^{-1/2}, (N + 1)^{-1/2}, \dots, (N + 1)^{-1/2})^*. \quad (2.6)$$

The matrix $H = (H_{ij})$ is of the same form as M ,

$$H_{ij} = h_{ij}, \quad 1 \leq i \neq j \leq N + 1, \quad H_{ii} = - \sum_{1 \leq k \leq N+1, k \neq i} h_{ki}, \quad (2.7)$$

where the h_{ij} are independent random variables (up to the symmetry constraint $h_{ij} = h_{ji}$) defined by

$$h_{ij} := \frac{m_{ij}}{q(1 - q^2/N)^{1/2}} - \frac{q}{N(1 - q^2/N)^{1/2}}, \quad 1 \leq i \neq j \leq N + 1. \quad (2.8)$$

Note that the entries of H are centered and for any indices $1 \leq i \neq j \leq N + 1$, the off-diagonal element h_{ij} satisfies

$$\mathbb{E}[|h_{ij}|^2] = \frac{1}{N}, \quad \mathbb{E}[|h_{ij}|^k] \leq \frac{1}{Nq^{k-2}}, \quad k \geq 3. \quad (2.9)$$

Remark 2.1. Note that since \mathbf{e} is an eigenvector of H , all spectral properties of M can be deduced from H as defined in (2.5). We will therefore state and prove results only for H . Analogous statements for M follow trivially.

The above discussion leads us to the following definition of random Laplacian-type matrices.

Definition 2.2. A real symmetric matrix $H = (H_{ij})_{ij}$ is called a *random Laplacian-type matrix* if it is of the form

$$H = \begin{pmatrix} -\sum_{j \neq 1} h_{1j} & h_{12} & \cdots & h_{1N+1} \\ h_{21} & -\sum_{j \neq 2} h_{2j} & \cdots & h_{2N+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N+11} & h_{N+12} & \cdots & -\sum_{j \neq N+1} h_{N+1j} \end{pmatrix} \quad (2.10)$$

where h_{ij} are independent random variables (up to the symmetry constraint $h_{ij} = h_{ji}$) with mean zero and variance $1/N$. We assume that they satisfy the moment conditions: for any $p \geq 3$, there exists some constant c_p such that

$$\mathbb{E}[|h_{ij}|^p] \leq \frac{c_p}{q^{p-2}N}, \quad (2.11)$$

where the *sparsity parameter* q satisfies

$$N^\beta \leq q \leq N^{1/2} \quad (2.12)$$

for some $\beta > 0$.

Remark 2.3. We state and prove our results only in the real symmetric case, however our results and proofs carry over to the complex Hermitian case with only minor notational changes.

2.2. Free convolution

The results of [12,33] state that the empirical eigenvalue distribution of the random Laplacian-type matrix converges weakly to the free convolution of the semicircle law and standard Gaussian, which we define in this section. We denote this free convolution by ρ_{fc} ,

$$\rho_{fc} = \rho_{sc} \boxplus \rho_G. \quad (2.13)$$

Above, ρ_{sc} and ρ_G are the semicircle and standard Gaussian distributions, respectively, and are given by

$$\rho_{sc}(E) = \frac{1}{2\pi} \sqrt{(4 - E^2)_+}, \quad \rho_G(E) = \frac{1}{\sqrt{2\pi}} e^{-E^2/2}. \quad (2.14)$$

In Figure 1, we plot the numerically calculated free convolution $\rho_{sc} \boxplus \rho_G$ of the semicircle and standard Gaussian distributions. The precise definition of (2.13) is through its Stieltjes transform which we denote by m_{fc} . The Stieltjes transform m_{fc} is defined as the solution to the following functional equation on the upper half complex plane $\text{Im}[z] > 0$,

$$m_{fc}(z) = \int \frac{\rho_G(x)}{x - z - m_{fc}(z)} dx. \quad (2.15)$$

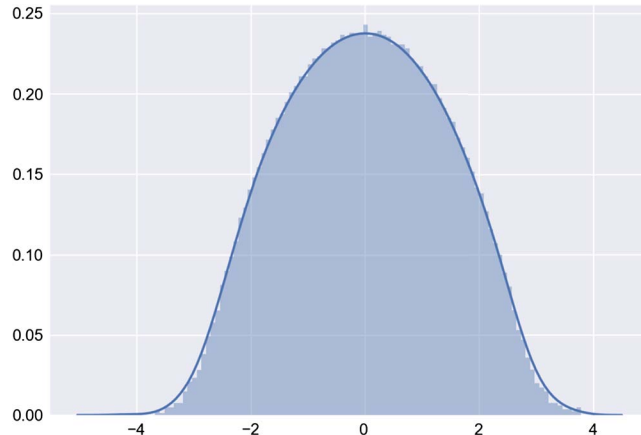


Fig. 1. Free convolution of the semicircle and standard Gaussian distributions.

It is well known [8] that there is a unique solution satisfying $\text{Im}[m_{\text{fc}}(z)] \geq 0$ and $m_{\text{fc}}(z) \sim z^{-1}$ as $z \rightarrow \infty$, and its properties are well-studied. In particular, m_{fc} is the Stieltjes transform of a measure with density ρ_{fc} that is analytic and can be recovered by the Stieltjes inversion formula

$$\rho_{\text{fc}}(E) = \lim_{\eta \downarrow 0} \frac{1}{\pi} \text{Im}[m_{\text{fc}}(E + i\eta)]. \quad (2.16)$$

We denote the classical eigenvalues of ρ_{fc} and ρ_{sc} by γ_i and $\gamma_{i,\text{sc}}$ respectively. They are defined by the equations

$$\int_{-\infty}^{\gamma_i} \rho_{\text{fc}}(x) dx = \frac{i - \frac{1}{2}}{N}, \quad \int_{-\infty}^{\gamma_{i,\text{sc}}} \rho_{\text{sc}}(x) dx = \frac{i - \frac{1}{2}}{N}. \quad (2.17)$$

2.3. Main results

A random Laplacian-type matrix H has $N + 1$ eigenvalues, one of which is always 0 with eigenvector \mathbf{e} . We call this the trivial eigenvalue. We denote the remaining N nontrivial eigenvalues of H as $\lambda_1(H) \leq \lambda_2(H) \leq \dots \leq \lambda_N(H)$ and corresponding eigenvectors as $u_1(H), u_2(H), \dots, u_N(H)$. We will sometimes write them as λ_i and u_i for simplicity of notation.

In this paper we are interested in the properties of the ‘‘bulk’’ eigenvalues of H . We fix a large constant $L > 0$ and consider eigenvalues in the interval $[-L, L]$. We fix the parameter

$$\xi := N^\nu \quad (2.18)$$

where $\nu > 0$ is a small parameter. The parameter $\nu > 0$ can be arbitrarily small, but in proofs its value may change from line to line; this will only occur at finitely many steps and we will not track its value explicitly.

We define the spectral domain $\mathcal{D}_L = \mathcal{D}_L^{(0)} \cup \mathcal{D}_L^{(1)} \cup \mathcal{D}_L^{(2)}$, where

$$\begin{aligned} \mathcal{D}_L^{(0)} &= \{z = E + i\eta : |E| \leq 2\xi + L, 2 \leq \eta \leq L\}, \\ \mathcal{D}_L^{(1)} &= \{z = E + i\eta : |E| \leq L, \xi^3 N^{-1} \leq \eta \leq 2\}, \\ \mathcal{D}_L^{(2)} &= \{z = E + i\eta : 2\xi \leq |E| \leq 2\xi + L, \xi^3 N^{-1} \leq \eta \leq 2\}. \end{aligned} \quad (2.19)$$

On the domain $\mathcal{D}_L^{(1)}$ we will prove a local law down to the optimal scale. In order to obtain information on the eigenvalue locations in the interval $[-L, L]$ we will also require estimates on the domains $\mathcal{D}_L^{(0)}$ and $\mathcal{D}_L^{(2)}$.

The spectral parameter z will play a fundamental role in this paper. We will write it as the sum of its real and imaginary parts as

$$z = E + i\eta, \quad E = \text{Re}[z], \eta = \text{Im}[z] \quad (2.20)$$

We define the control parameter

$$\Psi(z) = \frac{1}{q} + \frac{1}{(N\eta)^{1/2}}. \quad (2.21)$$

On the domain \mathcal{D}_L we always have

$$\xi \Psi \leq \log(N)^{-1} \quad (2.22)$$

for N large enough.

Throughout the paper we use the following notion of *overwhelming probability*.

Definition 2.4. We say that a family of events $\mathcal{F}(u)$ indexed by some parameter(s) u holds with overwhelming probability if for any large $D > 0$ we have for all N large enough,

$$\mathbb{P}[\mathcal{F}(u)] \geq 1 - N^{-D}, \quad (2.23)$$

uniformly in u .

Given two families of events $\mathcal{F}_1(u)$ and $\mathcal{F}_2(u)$ we say that \mathcal{F}_2 holds with overwhelming probability on \mathcal{F}_1 if for any large $D > 0$, we have for all N large enough, $\mathbb{P}[\mathcal{F}_2(u)^c \cap \mathcal{F}_1(u)] \leq N^{-D}$, uniformly in u .

Often we will take u to lie in the spectral domain \mathcal{D}_L or the indices $\llbracket 1, N+1 \rrbracket$ or some combination thereof. Unless it is mentioned otherwise, the notion of overwhelming probability will always be uniform in the parameters under consideration.

We denote the resolvent of the matrix H by

$$G(z) := (H - z)^{-1}, \quad (2.24)$$

and the Stieltjes transform of the empirical eigenvalue distribution of H by

$$m_N(z) := \frac{1}{N} \text{Tr} G(z). \quad (2.25)$$

Sometimes we will omit the parameter z and simply denote them by G and m_N , if the context is clear.

Now we can state the main results of this paper.

Theorem 2.5 (Local Law). *Let H be a random Laplacian-type matrix as in Definition 2.2 and let $G(z)$ be its resolvent. There is a constant $C > 0$ such that with overwhelming probability,*

$$|m_N(z) - m_{\text{fc}}(z)| \leq C\xi\Psi \quad (2.26)$$

and

$$\left| G_{ij}(z) - \delta_{ij} \frac{1}{\sum_{k \neq i} h_{ik} - z - m_{\text{fc}}(z)} \right| \leq C\xi\Psi \quad (2.27)$$

uniformly for $z \in \mathcal{D}_L$.

Theorem 2.5 implies the following rigidity estimates on the eigenvalues and complete delocalization of eigenvectors.

Corollary 2.6 (Rigidity of Eigenvalues). *Let H be a random Laplacian-type matrix as in Definition 2.2. Fix a small $\kappa > 0$. With overwhelming probability,*

$$|\lambda_i(H) - \gamma_i| \leq C\xi^2 q^{-1}, \quad (2.28)$$

uniformly for indices $i \in \llbracket \kappa N, (1 - \kappa)N \rrbracket$.

Corollary 2.7 (Delocalization of Eigenvectors). *Let H be a random Laplacian-type matrix as in Definition 2.2. There is a constant $C > 0$ so that the following holds. Fix a small $\kappa > 0$. With overwhelming probability we have*

$$\|u_i(H)\|_\infty^2 \leq C\xi^3 N^{-1}, \quad (2.29)$$

uniformly for indices $i \in \llbracket \kappa N, (1 - \kappa)N \rrbracket$.

Remark 2.8. In this note, we focus on the bulk eigenvalues of the random Laplacian-type matrix. As for the behavior of extreme eigenvalues, we refer the interested reader to the papers [18,34], where it is proved that for the non-sparse random Laplacian-type matrix (i.e. $q = N^{1/2}$) that for any fixed index i ,

$$\frac{\lambda_i(H)}{\sqrt{2N \log N}} \longrightarrow 1, \quad (2.30)$$

almost surely. The eigenvectors corresponding to those extreme eigenvalues can be studied in the Gaussian case (2.35) by finite rank perturbation theory [7, Theorem 2.2]. We cannot expect delocalization results such as Corollary 2.7.

Finally, we have the following result on the bulk universality of the eigenvalue statistics. To state it we define the n -point correlation functions. Denote by

$$\rho_H(\lambda_1, \dots, \lambda_N) \quad (2.31)$$

the unordered eigenvalue density of the matrix H . The n -point correlation function is defined by

$$p_H^{(n)}(\lambda_1, \dots, \lambda_n) = \int p_H(\lambda_1, \dots, \lambda_N) d\lambda_{n+1} \dots d\lambda_N. \quad (2.32)$$

We denote the corresponding quantities for the GOE by $p_{GOE}^{(n)}$ and p_{GOE} .

Theorem 2.9. *Let H be a random Laplacian-type matrix as in Definition 2.2. Then the nontrivial eigenvalues of H exhibit universality in the following two forms. Firstly, the nontrivial eigenvalues of H have gap universality. For any $\kappa > 0$ and index $i \in \llbracket \kappa N, (1 - \kappa)N \rrbracket$*

$$\begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{E}^{(H)} \left[O(N\rho_{fc}(\gamma_i)(\lambda_i - \lambda_{i+1}), \dots, N\rho_{fc}(\gamma_i)(\lambda_i - \lambda_{i+n})) \right] \\ & - \mathbb{E}^{(GOE)} \left[O(N\rho_{sc}(\gamma_{i,sc})(\lambda_i - \lambda_{i+1}), \dots, N\rho_{sc}(\gamma_{i,sc})(\lambda_i - \lambda_{i+n})) \right] = 0, \end{aligned} \quad (2.33)$$

for test functions $O \in C_0^\infty(\mathbb{R}^n)$, where $\gamma_{i,sc}$'s are the classical eigenvalue locations with respect to semi-circle law ρ_{sc} .

Secondly, the averaged n -point correlation functions of the nontrivial eigenvalues of H are universal. For any $\delta > 0$, $E \in [-L, L]$, and $b \geq N^{-1+\delta}$ we have

$$\begin{aligned} & \lim_{N \rightarrow \infty} \int_{E-b}^{E+b} \int_{\mathbb{R}^n} O(\alpha_1, \dots, \alpha_n) \left\{ \frac{1}{\rho_{fc}(E)^n} \rho_H^{(n)} \left(E' + \frac{\alpha_1}{N\rho_{fc}(E)}, \dots, E' + \frac{\alpha_n}{N\rho_{fc}(E)} \right) \right. \\ & \left. - \frac{1}{\rho_{sc}(E)^n} \rho_{GOE}^{(n)} \left(E' + \frac{\alpha_1}{N\rho_{sc}(E)}, \dots, E' + \frac{\alpha_n}{N\rho_{sc}(E)} \right) \right\} d\alpha_1 \dots d\alpha_n \frac{dE'}{2b} = 0, \end{aligned} \quad (2.34)$$

where the test function $O \in C_c^\infty(\mathbb{R}^n)$.

2.4. Gaussian toy model

In this section we introduce a Gaussian toy model of the random Laplacian-type matrix. We let W be the following $(N+1) \times (N+1)$ real symmetric matrix

$$W = \begin{pmatrix} -\sum_{j \neq 1} w_{1j} & w_{12} & \dots & w_{1N+1} \\ w_{21} & -\sum_{j \neq 2} w_{2j} & \dots & w_{2N+1} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N+11} & w_{N+12} & \dots & -\sum_{j \neq N+1} w_{N+1j} \end{pmatrix} \quad (2.35)$$

where $\{w_{ij}, 1 \leq i < j \leq N+1\}$ are independent standard Gaussian random variables with mean 0 and variance $1/N$.

We sketch a proof of Theorem 2.5 and 2.9 for the above toy model. This argument heavily depends on the Gaussian structure and cannot be generalized to the random Laplacian-type matrix of Definition 2.2. However, the argument provides some intuition as to why one expects that Theorem 2.9 holds. More importantly, the structure of W exhibited in the decomposition (2.38) will play a role in the proof for the general case.

The matrix W always has a trivial eigenvalue 0 with eigenvector \mathbf{e} and so we need to remove this eigenvalue in order to obtain universality. We project W down to an $N \times N$ matrix by conjugation with an orthogonal matrix R . Let \tilde{R} be any fixed $(N+1) \times (N+1)$ deterministic orthogonal matrix with last column \mathbf{e} . That is, $\tilde{R} = (R|\mathbf{e})$, where R is the $(N+1) \times N$ matrix consisting of the first N columns of \tilde{R} . We then have

$$\tilde{R}^* W \tilde{R} = \hat{W} \oplus 0, \quad (2.36)$$

where \hat{W} is the $N \times N$ matrix

$$\hat{W} := R^* W R. \quad (2.37)$$

The entries of the matrix \hat{W} are joint Gaussian random variables. We compute the covariance structure of the matrix \hat{W} , which will prove the following.

Proposition 2.10. *We have the following equality in law,*

$$\hat{W} \stackrel{d}{=} GOE + R^* D R + g I_N, \quad (2.38)$$

where GOE is an $N \times N$ GOE matrix, D is an $(N + 1) \times (N + 1)$ diagonal matrix with diagonal entries independent Gaussian random variables with variance $(N + 1)/N$, and g is a Gaussian random variable with variance $1/N$. All three are independent of each other.

Proof. The matrix elements of both sides of (2.38) are jointly distributed centered Gaussians. Hence we need only to check that the covariances of the matrix elements are the same. For the left hand side of (2.38) we first compute the covariance structure of W . We introduce auxiliary independent Gaussian variables w_{ii} 's, for $1 \leq i \leq N + 1$, with variance $2/N$. We have,

$$\begin{aligned} \mathbb{E}[W_{ij}W_{kl}] &= \mathbb{E}\left[\left(w_{ij} - \delta_{ij} \sum_{m=1}^{N+1} w_{mi}\right)\left(w_{kl} - \delta_{kl} \sum_{m=1}^{N+1} w_{mk}\right)\right] \\ &= \frac{1}{N}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \delta_{ijk} - \delta_{ijl} - \delta_{ikl} - \delta_{jkl} + \delta_{ij}\delta_{kl} + (N + 1)\delta_{ijkl}) \end{aligned} \quad (2.39)$$

where $\delta_{abc} = 1$ (resp $\delta_{abcd} = 1$) if and only if $a = b = c$ (resp $a = b = c = d$), otherwise it is zero. For indices $a, b, c, d \in \llbracket 1, N \rrbracket$ we have

$$\begin{aligned} \mathbb{E}[\hat{W}_{ab}\hat{W}_{cd}] &= \mathbb{E}\left[\sum_{i,j,k,l} R_{ia}W_{ij}R_{jb}R_{kc}W_{kl}R_{ld}\right] \\ &= \frac{1}{N} \sum_{i,j,k,l} R_{ia}R_{jb}R_{kc}R_{ld}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \delta_{ijk} - \delta_{ijl} - \delta_{ikl} - \delta_{jkl} + \delta_{ij}\delta_{kl} + (N + 1)\delta_{ijkl}) \\ &= \frac{1}{N} \left(\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc} + \delta_{ab}\delta_{cd} + (N + 1) \sum_{i=1}^{N+1} R_{ia}R_{ib}R_{ic}R_{id} \right), \end{aligned} \quad (2.40)$$

where we used the fact that all columns of R are orthogonal to \mathbf{e} . For the righthand side of (2.38) we compute

$$\begin{aligned} &\mathbb{E}[(GOE + R^*DR + gI_N)_{ab} \times (GOE + R^*DR + gI_N)_{cd}] \\ &= \mathbb{E}[GOE_{ab}GOE_{cd}] + \mathbb{E}[(R^*DR)_{ab}(R^*DR)_{cd}] + \mathbb{E}[(gI_N)_{ab}(gI_N)_{cd}] \\ &= \frac{1}{N}(\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}) + \frac{N + 1}{N} \sum_{i=1}^{N+1} R_{ia}R_{ib}R_{ic}R_{id} + \frac{1}{N}\delta_{ab}\delta_{cd} \end{aligned} \quad (2.41)$$

This finishes the proof. \square

The advantage of the decomposition (2.38) is that the two summands $R^*DR + gI_N$ and GOE are independent. In this case, both the global empirical eigenvalue distribution and the local eigenvalue statistics are well-understood. For the global eigenvalue distribution, one notices that the empirical distribution of $R^*DR + gI_N$ converges to the law of a standard Gaussian. From the theory of free probability, the global law of eigenvalue distribution of \hat{W} converges to free convolution of semi-circle and standard gaussian distributions. We refer to [3, Chapter 5] for an introduction on free probability. For the local eigenvalue statistics, one can show that the Stieltjes transform of the empirical eigenvalue distribution of $R^*DR + gI_N$ is close to that of a Gaussian down to the optimal scale, i.e. for $\eta \gtrsim N^{-1}$, with overwhelming probability. The RHS of (2.38) is a deformed Wigner ensemble, and universality of the bulk eigenvalue statistics can be concluded from the main results of [37] or [35].

3. Local laws

The goal of this section is to prove Theorem 2.5, the local law for random Laplacian-type matrices. The proof is based on the analysis of the self-consistent equation for the empirical Stieltjes transform of H . The main difficulty here is the correlation between diagonal entries and off-diagonal entries of H . As a result, the analysis is more involved than in the case of Wigner or deformed Wigner ensembles.

3.1. Notation and preliminaries

We introduce some notation. We denote the k -th minor of H by $H^{(k)}$, which is the submatrix of H with the k -th row and column removed. We denote the random Laplacian-type matrix of the subgraph obtained by removing the k -th vertex by $\tilde{H}^{(k)}$. More precisely, we define $\tilde{H}^{(k)}$ by

$$H_{ij}^{(k)} =: \tilde{H}_{ij}^{(k)} - \delta_{ij} h_{ik}, \quad i, j \neq k. \quad (3.1)$$

We denote the Green's functions of H , $H^{(k)}$ and $\tilde{H}^{(k)}$ by G , $G^{(k)}$ and $\tilde{G}^{(k)}$ respectively. We furthermore define the normalized Stieltjes transforms by

$$m_N(z) = \frac{1}{N} \operatorname{tr} G(z), \quad m_N^{(k)}(z) = \frac{1}{N} \operatorname{tr} G^{(k)}(z), \quad \tilde{m}_N^{(k)}(z) = \frac{1}{N} \operatorname{tr} \tilde{G}^{(k)}(z). \quad (3.2)$$

Note that we normalize the Stieltjes transform of H by $1/N$ even though it is an $(N+1) \times (N+1)$ matrix. Finally, we denote by

$$\sum_i^{(k)} = \sum_{\substack{1 \leq i \leq N+1 \\ i \neq k}}^{(k)}, \quad \sum_{ij}^{(k)} = \sum_{\substack{1 \leq i, j \leq N+1 \\ i, j \neq k}}^{(k)} \quad (3.3)$$

summation in the index i or i, j omitting the index k .

In the following we collect some resolvent identities which will be used repeatedly in the paper. A proof can be found in [30].

Proposition 3.1. *The Schur complement formula*

$$G_{kk}(z) = \frac{1}{-\sum_i^{(k)} h_{ik} - z - \sum_{ij}^{(k)} h_{ki} G_{ij}^{(k)} h_{kj}}. \quad (3.4)$$

The identities for $i, j \neq k$

$$G_{ij} = G_{ij}^{(k)} + \frac{G_{ik} G_{kj}}{G_{kk}}. \quad (3.5)$$

The identities for $i \neq k$

$$G_{ik} = -G_{kk} \sum_j^{(k)} h_{kj} G_{ji}^{(k)}. \quad (3.6)$$

We have also the Ward identity

$$\sum_j |G_{ij}|^2 = \frac{1}{\eta} \operatorname{Im}[G_{ii}]. \quad (3.7)$$

All of the above identities hold with G replaced by $G^{(k)}$ or $\tilde{G}^{(k)}$. Lastly we have the resolvent identities, which connect $G^{(k)}$ and $\tilde{G}^{(k)}$, for $i, j \neq k$,

$$G_{ij}^{(k)} = \tilde{G}_{ij}^{(k)} + \sum_l^{(k)} G_{il}^{(k)} h_{kl} \tilde{G}_{lj}^{(k)} = \tilde{G}_{ij}^{(k)} + \sum_l^{(k)} \tilde{G}_{il}^{(k)} h_{kl} G_{lj}^{(k)}. \quad (3.8)$$

By the Schur complement formula (3.4), we have

$$G_{kk}(z) = \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_N(z) - \mathcal{E}_k}, \quad (3.9)$$

where the error term is given by

$$\mathcal{E}_k(z) := \sum_{i \neq j}^{(k)} h_{ki} G_{ij}^{(k)}(z) h_{jk} - \sum_i^{(k)} \left(h_{ki}^2 - \frac{1}{N} \right) G_{ii}^{(k)}(z) - (m_N(z) - m_N^{(k)}(z)). \quad (3.10)$$

Definition 3.2. Let c_* be the constant introduced in (3.13) below. Note that c_* depends only on the choice of L in \mathcal{D}_L . For any $z \in \mathcal{D}_L$ we define the events,

$$\Xi^{(k)}(z) = \left\{ \max_{i, j \neq k} |\tilde{G}_{ij}^{(k)}(z)| \leq 2/c_* \right\}, \quad k = 1, 2, \dots, N+1, \quad (3.11)$$

and the event

$$\Xi(z) := \bigcap_k \Xi^{(k)}(z) \cap \left\{ |m_N(z) - m_{\text{fc}}(z)| \leq (\log(N))^{-1} \right\}. \quad (3.12)$$

Remark 3.3. We will often omit the parameter z and write $\Xi^{(k)}$ and Ξ if the context is clear.

We will need the following regularity estimates for m_{fc} . Parts of the proof use ideas of [8].

Lemma 3.4. For any fixed $L > 0$ we have that the imaginary part of m_{fc} is bounded above and away from zero,

$$c_* \leq \text{Im}[m_{\text{fc}}(z)] \leq |m_{\text{fc}}(z)| \leq 1, \quad (3.13)$$

uniformly for $z \in \mathcal{D}_L^{(1)}$. The constant c_* depends only on L .

Moreover, m_{fc} is Lipschitz

$$|m'_{\text{fc}}(z)| \leq C \quad (3.14)$$

uniformly for all $\text{Im}[z] \geq 0$.

Proof. We have by the Cauchy–Schwarz inequality,

$$|m_{\text{fc}}(z)|^2 \leq \int \frac{\rho_G(x) dx}{|x - z - m_{\text{fc}}(z)|^2} = \frac{\text{Im}[m_{\text{fc}}(z)]}{\text{Im}[z + m_{\text{fc}}(z)]} \leq 1, \quad (3.15)$$

which gives us the upper bound in (3.13). The lower bound follows from the fact that ρ_{fc} is supported on the whole real axis.

To prove that m_{fc} is Lipschitz, we differentiate both sides of the defining equation (2.15) to obtain

$$\left(1 - \int \frac{\rho_G(x) d(x)}{(x - z - m_{\text{fc}}(z))^2} \right) m'_{\text{fc}}(z) = \int \frac{\rho_G(x) dx}{(x - z - m_{\text{fc}}(z))^2}. \quad (3.16)$$

From (3.15), the right-hand side is bounded by 1, and so we need only prove that the factor in front of $m'_{\text{fc}}(z)$ is bounded away from 0. In the following we prove that

$$\text{Re} \left[\int \frac{d\mu_G(x)}{(x - z - m_{\text{fc}}(z))^2} \right] \leq 1 - c \quad (3.17)$$

for a $c > 0$ and (3.14) follows.

If the imaginary part of z , $\text{Im}[z] \geq 2$, then $|x - z - m_{\text{fc}}(z)| \geq 2$, and (3.17) is trivial. Therefore in the following we assume that the imaginary part of z is bounded by 2.

We first consider the regime $|\text{Re}[z]| \geq L'$ for a fixed large $L' > 0$ chosen later. Denote $z + m_{\text{fc}}(z) = a + bi$. Without loss of generality take $\text{Re}[z] \geq 0$ and so $a \geq L' - 1$ because $|m_{\text{fc}}| \leq 1$. Integration by parts yields,

$$\int \frac{d\mu_G(x)}{(x - z - m_{\text{fc}}(z))^2} = \int \frac{-x\rho_G(x) dx}{x - z - m_{\text{fc}}(z)} = \int_{a-4}^{a+4} \frac{-x\rho_G(x) dx}{x - z - m_{\text{fc}}(z)} + \int_{[a-4, a+4]^c} \frac{-x\rho_G(x) dx}{x - z - m_{\text{fc}}(z)}. \quad (3.18)$$

Choosing $L' > 0$ large enough we see that the second term is bounded by

$$\left| \int_{[a-4, a+4]^c} \frac{-x\rho_G(x) dx}{x-z-m_{fc}(z)} \right| \leq \frac{1}{4} \int |x|\rho_G(x) dx \leq \frac{1}{4}. \quad (3.19)$$

We rearrange the first term as

$$\int_{a-4}^{a+4} \frac{-x\rho_G(x) dx}{x-z-m_{fc}(z)} = - \int_{a-4}^{a+4} \rho_G(x) dx - (z+m_{fc}(z)) \int_{a-4}^{a+4} \frac{\rho_G(x) dx}{x-z-m_{fc}(z)}. \quad (3.20)$$

For the first integral in (3.20) we have the bound,

$$\left| - \int_{a-4}^{a+4} \rho_G(x) dx \right| \leq \frac{1}{\sqrt{2\pi}(a-4)} e^{-(a-4)^2/2}. \quad (3.21)$$

The second term in (3.20) is somewhat harder due to the ‘‘singularity’’ at $x = a$. We write,

$$\begin{aligned} & (z+m_{fc}(z)) \int_{a-4}^{a+4} \frac{\rho_G(x) dx}{x-z-m_{fc}(z)} \\ &= (a+ib) \left(\int_0^4 dy \frac{\rho_G(a+y)}{y-ib} - \frac{\rho_G(a-y)}{y+ib} \right) \\ &= (a+ib) \left(\int_0^4 dy \frac{\rho_G(a+y) - \rho_G(a-y)}{y-ib} \right) + (a+ib) \int_0^4 dy \frac{2ib}{y^2+b^2} \rho_G(a-y). \end{aligned} \quad (3.22)$$

The real part of above expression is bounded above by

$$(a+3) \int_0^4 dy \frac{\rho_G(a+y) - \rho_G(a-y)}{y} + \int_0^4 dy \frac{2b^2}{y^2+b^2} \rho_G(a-y) \leq \frac{2(a^2+3a+1)}{\sqrt{2\pi}(a-4)} e^{-(a-4)^2/2}. \quad (3.23)$$

Therefore (3.17) follows after taking $L' > 0$ sufficiently large.

We now consider the case $|\operatorname{Re}[z]| \leq L'$. We have that the imaginary part of m_{fc} is bounded below by c' . In this case we have

$$\begin{aligned} \operatorname{Re} \left[\int \frac{\rho_G(x) dx}{(x-z-m_{fc}(z))^2} \right] &= \int \rho_G(x) dx \frac{(\operatorname{Re}[x-z-m_{fc}])^2 - (\operatorname{Im}[z+m_{fc}])^2}{|x-z-m_{fc}|^4} \\ &\leq \int \frac{\rho_G(x) dx}{|x-z-m_{fc}|^2} - (\operatorname{Im}[z+m_{fc}])^2 \int \frac{\rho_G(x) dx}{|x-z-m_{fc}|^4} \\ &\leq 1 - (\operatorname{Im}[m_{fc}(z)])^2 \leq 1 - (c')^2, \end{aligned} \quad (3.24)$$

where in the second inequality we have used the Cauchy–Schwarz inequality

$$(\operatorname{Im}[z+m_{fc}])^2 \int \frac{\rho_G(x) dx}{|x-z-m_{fc}|^4} \geq (\operatorname{Im}[z+m_{fc}])^2 \left(\int \frac{\rho_G(x) dx}{|x-z-m_{fc}|^2} \right)^2 = (\operatorname{Im}[m_{fc}])^2. \quad (3.25)$$

This finishes the proof of (3.17). \square

The following proposition provides a discrete version of the estimate (3.17).

Proposition 3.5. *There is a constant c_1 depending only on the constant c_* from (3.13) so that the following holds uniformly for $z \in \mathcal{D}_L^{(1)}$. Suppose that w_1, w_2, \dots, w_{N+1} is a sequence of real numbers satisfying*

$$\frac{1}{N} \sum_k \frac{1}{w_k - z - m_{fc}(z)} = m_{fc}(z) + \epsilon(z) \quad (3.26)$$

where the error term satisfies

$$|\epsilon(z)| \leq c_1 \quad (3.27)$$

Then we have

$$\frac{c_*^2}{16} \leq \left| 1 - \frac{1}{N} \sum_k \frac{1}{(w_k - z - m_{\text{fc}}(z))^2} \right| \leq 1 + \frac{1}{c_*^2} \quad (3.28)$$

In fact we can take $c_1 = \min\{c_*^3/16, c_*/2\}$.

Proof. The upper bound of (3.28) is an immediate consequence of the lower bound of (3.13). For the lower bound of (3.28) we compute

$$\begin{aligned} \operatorname{Re} \left[\frac{1}{N} \sum_k \frac{1}{(w_k - z - m_{\text{fc}}(z))^2} \right] &= \frac{1}{N} \sum_k \frac{(\operatorname{Re}[w_k - z - m_{\text{fc}}])^2 - (\operatorname{Im}[z + m_{\text{fc}}])^2}{|w_k - z - m_{\text{fc}}|^4} \\ &\leq \frac{1}{N} \sum_k \frac{1}{|w_k - z - m_{\text{fc}}|^2} - \frac{1}{N} \sum_k \frac{(\operatorname{Im}[z + m_{\text{fc}}])^2}{|w_k - z - m_{\text{fc}}|^4} \\ &\leq \frac{\operatorname{Im}[m_{\text{fc}} + \epsilon]}{\operatorname{Im}[m_{\text{fc}} + \eta]} - \frac{1}{2} (\operatorname{Im}[m_{\text{fc}} + \epsilon])^2 \\ &\leq 1 + \frac{|\epsilon|}{\operatorname{Im}[m_{\text{fc}}]} - \frac{(\operatorname{Im}[m_{\text{fc}}])^2}{8} \leq 1 - \frac{c_*^2}{16}. \end{aligned} \quad (3.29)$$

The last line follows from our choice of c_1 . In the third inequality we used the Cauchy–Schwarz inequality

$$\frac{1}{N+1} \sum_k \frac{(\operatorname{Im}[z + m_{\text{fc}}])^2}{|w_k - z - m_{\text{fc}}|^4} \geq \left(\frac{1}{N+1} \sum_k \frac{\operatorname{Im}[z + m_{\text{fc}}]}{|w_k - z - m_{\text{fc}}|^2} \right)^2 = (\operatorname{Im}[m_{\text{fc}} + \epsilon])^2. \quad (3.30)$$

□

We also need the following concentration of measure result.

Proposition 3.6. *For any $z \in \mathcal{D}_L$, with overwhelming probability, we have*

$$\left| \frac{1}{N} \sum_k \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_{\text{fc}}(z)} - m_{\text{fc}}(z) \right| = O\left(\frac{\xi}{q}\right) \quad (3.31)$$

Proof. From the assumption (2.11), we know that with overwhelming probability, we have

$$\max_{i < j} |h_{ij}| \leq \frac{\xi}{q \log N}, \quad \max_k \left| \sum_i^{(k)} h_{ki} \right| \leq \xi. \quad (3.32)$$

Thus we can do a truncation. For this, we define two cutoff functions. First, take $\rho(x) = x$ on $[-\xi/(q \log N), \xi/(q \log N)]$, and $\rho(x) = \operatorname{sign}(x)\xi/(q \log N)$ outside the interval. Secondly, let $\chi(x) = x$ on $[-\xi, \xi]$, and $\chi(x) = \operatorname{sign}(x)(\xi + 1)$ for $|x| > \xi + 1$, and with χ having its first three derivatives bounded. We also define the functions

$$F(\{h_{ij}\}_{i < j}) = \frac{1}{N} \sum_k \frac{1}{g_k}, \quad g_k := g_k(\{h_{ik}\}_{i \neq k}) = -\chi\left(\sum_i^{(k)} \rho(h_{ik})\right) - z - m_{\text{fc}}(z). \quad (3.33)$$

We claim that g_k is bounded away from 0 on \mathcal{D}_L . We have $|g_k| \geq |\operatorname{Im}[z]| \geq 2$ on $\mathcal{D}_L^{(0)}$. On $\mathcal{D}_L^{(1)}$ we have $|g_k| \geq \operatorname{Im}[m_{\text{fc}}] \geq c_*$ and on $\mathcal{D}_L^{(2)}$ we have $|g_k| \geq \xi - 2$. Therefore, for any pair $i < j$, we have

$$\begin{aligned} &|F(h_{12}, \dots, h'_{ij}, \dots, h_{NN+1}) - F(h_{12}, \dots, h_{ij}, \dots, h_{NN+1})| \\ &= \frac{1}{N} \left| \frac{g_i(h_{ij}) - g_i(h'_{ij})}{g_i(h_{ij})g_i(h'_{ij})} + \frac{g_j(h_{ij}) - g_j(h'_{ij})}{g_j(h_{ij})g_j(h'_{ij})} \right| \leq \frac{C}{N} |\rho(h_{ij}) - \rho(h'_{ij})| \leq \frac{2C\xi}{Nq \log N} \end{aligned} \quad (3.34)$$

Therefore by McDiarmid's inequality, we get

$$\mathbb{P}\left(\left|F(\{h_{ij}\}_{i<j}) - \mathbb{E}[F(\{h_{ij}\}_{i<j})]\right| \geq \frac{t\xi}{q \log N}\right) \leq e^{-Ct^2}. \quad (3.35)$$

Hence, by taking $t = \log N$, we have that the following holds with overwhelming probability

$$\left|F(\{h_{ij}\}_{i<j}) - \mathbb{E}[F(\{h_{ij}\}_{i<j})]\right| \leq \xi/q. \quad (3.36)$$

Moreover, since $\max_{i<j} |h_{ij}| \leq \xi/(q \log N)$ with overwhelming probability we have

$$F(\{h_{ij}\}_{i<j}) = \frac{1}{N} \sum_k \frac{1}{-\chi(\sum_i^{(k)} h_{ik}) - z - m_{\text{fc}}(z)}. \quad (3.37)$$

By a similar argument as above we see that $|-\chi(\sum_i^{(k)} h_{ik}) - z - m_{\text{fc}}(z)|$ is bounded away from zero uniformly for $z \in \mathcal{D}_L$. Therefore both sides of (3.37) are uniformly bounded. By taking expectation on both sides of (3.37), for any fixed $D > 0$, we have

$$\mathbb{E}[F(\{h_{ij}\}_{i<j})] = \mathbb{E}\left[\frac{1}{N} \sum_k \frac{1}{-\chi(\sum_i^{(k)} h_{ik}) - z - m_{\text{fc}}(z)}\right] + O(N^{-D}), \quad (3.38)$$

given that N is large enough. Moreover, from our choice of χ , we see that the first three derivatives of $(-\chi(\sum_i^{(k)} h_{ik}) - z - m_{\text{fc}}(z))^{-1}$ with respect to h_{ik} 's are bounded. By the standard Lindeberg replacement trick one can show that

$$\begin{aligned} \mathbb{E}\left[\frac{1}{-\chi(\sum_i^{(k)} h_{ik}) - z - m_{\text{fc}}(z)}\right] &= \mathbb{E}\left[\frac{1}{-\chi(\sum_{i=1}^N x_i) - z - m_{\text{fc}}(z)}\right] + O\left(\frac{1}{q}\right) \\ &= \mathbb{E}\left[\frac{1}{-\sum_{i=1}^N x_i - z - m_{\text{fc}}(z)}\right] + O\left(\frac{1}{q}\right) \end{aligned} \quad (3.39)$$

where x_j 's are Gaussian random variables with mean zero and variance $1/N$. The implicit constant in the error term depends only on the bound for third moment of the h_{ij} . Since $\sum_{i=1}^N x_i \stackrel{d}{=} \mathcal{N}(0, 1)$ we have by definition,

$$\mathbb{E}\left[\frac{1}{-\sum_{i=1}^N x_i - z - m_{\text{fc}}(z)}\right] = m_{\text{fc}}(z). \quad (3.40)$$

By combining (3.38), (3.39) and (3.40), we get

$$\left|\mathbb{E}[F(\{h_{ij}\}_{i<j})] - m_{\text{fc}}\right| = O(\xi/q). \quad (3.41)$$

The claim follows from this and (3.36). \square

3.2. Proof of local law

In this subsection we prove Theorem 2.5. Recall the definition of the events $\Xi^{(k)}$ and Ξ in Definition 3.2, the Green's functions G , $G^{(k)}$, $\tilde{G}^{(k)}$ and error term \mathcal{E}_k in (3.10).

We first prove that with overwhelming probability $\Xi(z)$ holds uniformly for $\eta \geq 2$. Then we will bootstrap the estimate by taking advantage of the self-consistent equation of the Stieltjes transform and the following error control, the proof of which is postponed to Section 3.3.

Proposition 3.7. *For any $z \in \mathcal{D}_L$ all of the following holds with overwhelming probability. We have,*

$$\mathbb{1}_{\Xi^{(k)}(z)} \left|G_{ij}^{(k)}(z) - \tilde{G}_{ij}^{(k)}(z)\right| \leq \xi \Psi, \quad (3.42)$$

and

$$\mathbb{1}_{\Xi^{(k)}(z)} \left|\sum_j^{(k)} h_{kj} G_{ji}^{(k)}\right| \leq \xi \Psi \quad (3.43)$$

and

$$\mathbb{1}_{\Xi^{(k)}(z)} \mathcal{E}_k(z) \leq \xi \Psi, \quad (3.44)$$

where the control parameter Ψ is introduced in (2.21).

The proof of the local law is decomposed into two propositions.

Proposition 3.8. *With overwhelming probability, $\Xi(z)$ holds uniformly for $\eta \geq 2$.*

Proof. Since $|G_{ij}|$, $|G_{ij}^{(k)}|$ and $|\tilde{G}_{ij}^{(k)}|$ are all bounded by $1/\text{Im}[z]$, $\Xi^{(k)}(z)$ holds deterministically for any z with $\eta \geq 2$. As for the error estimate of $m_N(z)$, we prove the following stronger result: for $\eta \geq 2$ we have,

$$|m_N(z) - m_{\text{fc}}(z)| \leq \xi \Psi \quad (3.45)$$

with overwhelming probability.

For any z with $\text{Im}[z] \geq 2$, we denote the event $\mathcal{A}(z)$ such that (3.31) and $\bigcap_{k=1}^{N+1} \{|\mathcal{E}_k| \leq \xi \Psi\}$ hold. Then from Propositions 3.6 and 3.7, $\mathcal{A}(z)$ holds with overwhelming probability and so we can assume for the remainder of the proof that $\mathcal{A}(z)$ holds. Starting from the Schur complement formula (3.4) we have,

$$G_{kk}(z) = \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_N(z) - \mathcal{E}_k} = \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_N(z)} + O(\xi \Psi), \quad (3.46)$$

where we used the fact that $\text{Im}[z] \geq 2$ to expand the denominator which is bounded away from 0. After averaging over k we get,

$$m_N(z) = \frac{1}{N} \sum_k \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_N} + O(\xi \Psi) \quad (3.47)$$

We obtain the following by taking the difference of (3.47) and (3.31)

$$m_N - m_{\text{fc}} = \frac{m_N - m_{\text{fc}}}{N} \sum_k \frac{1}{(-\sum_i^{(k)} h_{ik} - z - m_N)(-\sum_i^{(k)} h_{ik} - z - m_{\text{fc}})} + O(\xi \Psi). \quad (3.48)$$

The imaginary part of the denominators on the RHS of (3.48) is bounded below by

$$\text{Im} \left[\sum_{j \neq k} h_{kj} + z + m_N \right] \geq 2, \quad \text{Im} \left[\sum_{j \neq k} h_{kj} + z + m_{\text{fc}} \right] \geq 2. \quad (3.49)$$

Therefore,

$$\left| \frac{1}{N} \sum_k \frac{1}{(\sum_i^{(k)} h_{ik} + z + m_N)(\sum_i^{(k)} h_{ik} + z + m_{\text{fc}})} \right| \leq \frac{1}{2}. \quad (3.50)$$

Combining (3.48) and (3.50) we obtain

$$|m_N(z) - m_{\text{fc}}(z)| = O(\xi \Psi). \quad (3.51)$$

This finishes the proof. \square

The following proposition is the key bootstrap principle. The weak estimates of the event $\Xi(z)$ are used to prove stronger estimates on the Green's function and Stieltjes transform.

Proposition 3.9. *For $z \in \mathcal{D}_L$ the following holds with overwhelming probability on the event $\Xi(z)$. For the Stieltjes transform we have*

$$|m_N(z) - m_{\text{fc}}(z)| \leq \xi \Psi. \quad (3.52)$$

Let X_{ij} denote any of the resolvent entries $G_{ij}(z)$, $G_{ij}^{(k)}(z)$ or $\tilde{G}_{ij}^{(k)}(z)$. We have,

$$\left| X_{ij}(z) - \frac{\delta_{ij}}{-\sum_l^{(j)} h_{lj} - z - m_{fc}(z)} \right| \leq \xi \Psi. \quad (3.53)$$

Recall Ψ is the control parameter introduced in (2.21).

Proof. For any $z \in \mathcal{D}_L$, we define the event $\mathcal{A}(z)$ such that (3.31), (3.32) and $\Xi(z)$ hold, and the following holds:

$$\bigcap_{k=1}^{N+1} \left(\{ \mathcal{E}_k(z) \leq \xi \Psi \} \cap \left\{ \max_{i,j \neq k} |G_{ij}^{(k)}(z) - \tilde{G}_{ij}^{(k)}(z)| \leq \xi \Psi \right\} \cap \left\{ \max_{j \neq k} \left| \sum_j^{(k)} h_{kj} G_{jj}^{(k)} \right| \leq \xi \Psi \right\} \right). \quad (3.54)$$

By Propositions 3.6 and 3.7, and our assumption, $\mathcal{A}(z)$ holds with overwhelming probability on $\Xi(z)$, and so we need only to prove that on $\mathcal{A}(z)$, (3.52) and (3.53) hold.

We claim that on the event $\mathcal{A}(z)$, we have the uniform bound for sufficiently large N ,

$$\left| -\sum_i^{(k)} h_{ik} - z - m_N(z) \right| \geq \min \left\{ 2, \frac{c_*}{2} \right\} > 0, \quad (3.55)$$

where c_* is the constant from (3.13). If $z \in \mathcal{D}_L^{(0)}$ then $\text{Im}[z] \geq 2$. If $z \in \mathcal{D}_L^{(1)}$ then $\text{Im}[m_{fc}(z)] \geq c_*$ and so for sufficiently large N , $\text{Im}[m_N(z)] \geq c_*/2$ by the definition of $\mathcal{A}(z)$. Lastly, if $z \in \mathcal{D}_L^{(2)}$ then by the definition of $\mathcal{A}(z)$

$$\left| -\sum_i^{(k)} h_{ik} - z - m_N \right| \geq |z| - \left| \sum_i^{(k)} h_{ik} \right| - |m_N| \geq \xi - 2 \geq 2. \quad (3.56)$$

Therefore we may expand the Schur complement formula (3.4) and obtain,

$$m_N(z) = \frac{1}{N} \sum_k \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_N} + O(\xi \Psi). \quad (3.57)$$

Again, by the definition of $\mathcal{A}(z)$ we have

$$m_{fc}(z) = \frac{1}{N} \sum_k \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_{fc}} + O(\xi \Psi). \quad (3.58)$$

Taking the difference of (3.57) and (3.58) we get

$$\left(1 - \frac{1}{N} \sum_k \frac{1}{(-\sum_i^{(k)} h_{ik} - z - m_{fc})^2} - \mathcal{E} \right) (m_N - m_{fc}) = O(\xi \Psi) \quad (3.59)$$

where the error term is given by

$$|\mathcal{E}| = \left| \frac{1}{N} \sum_k \frac{m_N - m_{fc}}{(-\sum_i^{(k)} h_{ik} - z - m_N)(-\sum_i^{(k)} h_{ik} - z - m_{fc})^2} \right| \leq \frac{C}{\log N}. \quad (3.60)$$

We used again that the denominator is bounded away from zero on \mathcal{D}_L . By Proposition 3.5 and the assumption that (3.31) holds on \mathcal{A} , we see that the prefactor in front of $(m_N - m_{fc})$ in (3.59) is bounded away from zero by a constant on $\mathcal{D}_L^{(1)}$. This is also true on $\mathcal{D}_L^{(0)}$ due to the fact that $\text{Im}[z] \geq 2$, and is true on $\mathcal{D}_L^{(2)}$ by the fact that $|\sum_i^{(k)} h_{ik} - z - m_{fc}| \geq \xi - 2$ on $\mathcal{A}(z)$. We conclude (3.52).

On the event \mathcal{A} we have for the diagonal resolvent terms,

$$\begin{aligned} G_{kk} &= \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_{fc}(z) + (m_N(z) - m_{fc}(z)) + \mathcal{E}_k} \\ &= \frac{1}{-\sum_i^{(k)} h_{ik} - z - m_{fc}(z)} + O(\xi \Psi), \end{aligned} \quad (3.61)$$

where we used (3.52). For the off-diagonal resolvent entries we apply the identity (3.6) and obtain

$$|G_{ik}| = |G_{kk}| \left| \sum_j^{(k)} h_{kj} G_{ji}^{(k)} \right| = O(\xi \Psi), \quad (3.62)$$

using the fact that $|G_{kk}| \leq 2/c_*$ and $|\sum_j^{(k)} h_{kj} G_{ji}^{(k)}| \leq \xi \Psi$ on \mathcal{A} (recall (3.54)). For the entries of the k th minor we have by (3.5) and (3.6)

$$|G_{ij} - G_{ij}^{(k)}| = \left| \frac{G_{ik} G_{kj}}{G_{kk}} \right| = \left| \sum_j^{(k)} h_{kj} G_{ji}^{(k)} \right| |G_{kj}| = O(\xi \Psi). \quad (3.63)$$

The estimates for $\tilde{G}_{ij}^{(k)}$ then follow from the definition (3.54) of \mathcal{A} on which we have $|G_{ij}^{(k)}(z) - \tilde{G}_{ij}^{(k)}(z)| \leq \xi \Psi$. \square

Proof of Theorem 2.5. Once Propositions 3.8 and 3.9 have been established, the local law follows from a standard induction scheme as given in [21]. We provide here a sketch for the sake of completeness. Note that on \mathcal{D}_L all the resolvent entries are Lipschitz in the parameter z with constant less than CN^2 . It therefore suffices to work on the lattice

$$\Lambda = \mathcal{D}_L \cap N^{-4}(\mathbb{Z} + i\mathbb{Z}) \quad (3.64)$$

which has cardinality $|\Lambda| \leq 8LN^8\xi$. Fix now a sequence of sets of points $L_k = \{E + (2 - jN^{-4})i \in \Lambda : j \leq k\}$, and let \mathcal{A}_k be the event that the conclusion of Theorem 2.5 holds uniformly for points in L_k . We know by Proposition 3.8 that \mathcal{A}_0 holds with overwhelming probability. By the Lipschitz continuity of the resolvent entries, for any $z \in L_{k+1}$, the event $\Xi(z)$ holds on \mathcal{A}_k . Hence, by Proposition 3.9, the event \mathcal{A}_{k+1} holds with overwhelming probability on \mathcal{A}_k . Hence fixing a large number $D > 0$ we see by induction that

$$\mathbb{P}[\mathcal{A}_k^c] \leq \mathbb{P}[\mathcal{A}_{k-1}^c] + \mathbb{P}[\mathcal{A}_{k-1} \cap \mathcal{A}_k^c] \leq kN^{-D} + N^{-D} = (k+1)N^{-D}. \quad (3.65)$$

Hence,

$$\mathbb{P} \left[\bigcap_{k=0}^M \mathcal{A}_k \right] \geq 1 - \sum_{k=0}^{N^4-1} (k+1)N^{-D} \geq 1 - N^{-D+8}. \quad (3.66)$$

This yields the claim. \square

Proof of Corollary 2.6. We want to apply the arguments of Section 7 of [35] to conclude rigidity. However, since the norm of H is not bounded we need a slight modification of the argument. Using the local law (2.26), we apply [35, Lemma 7.17] with $E_1 = -2\xi$ and $E_2 \in (-L/2, L/2)$ and obtain

$$\left| \int_{E_1}^{E_2} \rho_{\text{fc}}(E) \, dE - \int_{E_1}^{E_2} \rho_N(E) \, dE \right| \leq C\xi^2 q^{-1}, \quad (3.67)$$

with overwhelming probability, where ρ_N the empirical eigenvalue distribution of H . We also have that ρ_N is supported in $[-2\xi, 2\xi]$ with overwhelming probability. By letting $\text{Im}[z]$ go to zero on both sides of (2.15), we get

$$\begin{aligned} \sqrt{2\pi} &= \int_{\mathbb{R}} \frac{e^{-x^2/2} \, dx}{(x - E - \text{Re}[m_{\text{fc}}(E)])^2 + \pi^2 \rho_{\text{fc}}(E)^2} \\ &\leq \int_{|x-E| \leq |E|/2} \frac{e^{-x^2/2} \, dx}{\pi^2 \rho_{\text{fc}}(E)^2} + \int_{|x-E| > |E|/2} \frac{e^{-x^2/2} \, dx}{(x - E - \text{Re}[m_{\text{fc}}(E)])^2} \\ &\leq \frac{2e^{-E^2/8}}{\pi^2 |E| \rho_{\text{fc}}(E)^2} + \frac{4}{(|E| - 2)^2} \end{aligned} \quad (3.68)$$

where we used that $|m_{\text{fc}}| \leq 1$ from (3.13). We therefore have that $\rho_{\text{fc}}(E) \leq e^{-E^2/8}$ for sufficiently large $|E|$, and therefore for any $E \leq E_1$. Therefore,

$$\int_{-\infty}^{E_1} \rho_{\text{fc}}(E) \, dE \leq N^{-D} \quad (3.69)$$

for any $D > 0$. Combining (3.67) and (3.69) we get

$$\left| \int_{-\infty}^{E_2} \rho_{\text{fc}}(E) \, dE - \int_{-\infty}^{E_2} \rho_N(E) \, dE \right| \leq C \xi^2 q^{-1} \quad (3.70)$$

By a standard argument (see, e.g., [31, Theorem 2.2]), we obtain the following rigidity estimate for the bulk eigenvalues. For any index i such that $\gamma_i \in [-L/2, L/2]$, we have

$$|\lambda_i(H) - \gamma_i| \leq C \xi^2 q^{-1}. \quad (3.71)$$

□

The proofs of Corollary 2.7 from Theorem 2.5 is standard and we omit the details. This argument can be found, for example, in [20, Theorem 2.16].

3.3. Proof of Proposition 3.7

We show that Proposition 3.7 follows from the following large deviation estimates.

Proposition 3.10. *For any $z \in \mathcal{D}_L$, the following holds with overwhelming probability: for $l \geq 1$ and $j \neq k$ we have,*

$$\mathbb{1}_{\Xi^{(k)}(z)} \sum_{i_1, \dots, i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots h_{ki_l} \tilde{G}_{i_l j}^{(k)} \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/4}} \right)^l, \quad (3.72)$$

and for $l \geq 1$ and $i, j \neq k$

$$\mathbb{1}_{\Xi^{(k)}(z)} \sum_{i_1, \dots, i_l}^{(k)} \tilde{G}_{i_1 i_1}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots h_{ki_l} \tilde{G}_{i_l j}^{(k)} \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/4}} \right)^l. \quad (3.73)$$

For $l \geq 3$ we have

$$\mathbb{1}_{\Xi^{(k)}(z)} \sum_{i_1, \dots, i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots \tilde{G}_{i_{l-1} i_l}^{(k)} h_{ki_l} \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/4}} \right)^{l-2}. \quad (3.74)$$

We have better estimates for smaller l .

Proposition 3.11. *For any $z \in \mathcal{D}_L$, the following holds with overwhelming probability: for $j \neq k$ we have,*

$$\mathbb{1}_{\Xi^{(k)}(z)} \sum_i^{(k)} h_{ki} \tilde{G}_{i,j}^{(k)} \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/2}} \right), \quad (3.75)$$

and for $i, j \neq k$

$$\mathbb{1}_{\Xi^{(k)}(z)} \sum_{i_1}^{(k)} \tilde{G}_{i_1 i_1}^{(k)} h_{ki_1} \tilde{G}_{i_1 j}^{(k)} \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/2}} \right). \quad (3.76)$$

We have also

$$\mathbb{1}_{\Xi^{(k)}(z)} \sum_{i,j}^{(k)} \tilde{G}_{ij}^{(k)} (h_{ki} h_{kj} - \delta_{ij} N^{-1}) \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/2}} \right), \quad (3.77)$$

and

$$\mathbb{1}_{\Xi^{(k)}(z)} \sum_{i_1, i_2, i_3}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \tilde{G}_{i_2 i_3}^{(k)} h_{ki_3} \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/2}} \right). \quad (3.78)$$

Proof of Proposition 3.7. In this proof we assume the event $\Xi^{(k)}(z)$ holds and omit it from the notation. We define the following three quantities corresponding to (3.72), (3.73) and (3.74),

$$\begin{aligned} A_l(i, j) &:= \sum_{i_1, \dots, i_l}^{(k)} \tilde{G}_{ii_1}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} \dots h_{ki_l} \tilde{G}_{i_l j}^{(k)}, \\ A_l(j) &:= \sum_{i_1, \dots, i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots h_{ki_l} \tilde{G}_{i_l j}^{(k)}, \\ A_l &:= \sum_{i_1, \dots, i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots \tilde{G}_{i_{l-1} i_l}^{(k)} h_{ki_l}, \end{aligned} \quad (3.79)$$

for any $l \geq 1$. In order to prove (3.42) we iterate (3.8) and obtain

$$G_{ij}^{(k)} - \tilde{G}_{ij}^{(k)} = \sum_{l=1}^m A_l(i, j) + \sum_n^{(k)} A_m(i, n) h_{kn} G_{nj}^{(k)}. \quad (3.80)$$

By (3.76) and (3.73) we have with overwhelming probability,

$$A_1(i, j) \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/2}} \right), \quad (3.81)$$

and for $l \geq 2$,

$$A_l(i, j) \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/4}} \right)^l. \quad (3.82)$$

In order to control the last term on the right-hand side of (3.80) we use the trivial bound $|G_{nj}^{(k)}| \leq 1/\text{Im}[z] \leq N$ on \mathcal{D}_L and take m large (but independent of N , say $m \geq 4(1/\nu + 1)$, where ν is as in (2.18)) and obtain

$$|G_{ij}^{(k)} - \tilde{G}_{ij}^{(k)}| \leq \sum_{l=1}^m |A_l(i, j)| + \frac{1}{N} \leq C\xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/2}} \right) \quad (3.83)$$

with overwhelming probability.

In order to prove (3.43) we iterate (3.8) and obtain

$$\sum_i^{(k)} h_{ki} G_{ij}^{(k)} = A_1(j) + \sum_{l=2}^m A_l(j) + \sum_n^{(k)} A_m(n) h_{kn} G_{nj}^{(k)}. \quad (3.84)$$

The first term $A_1(j)$ is controlled by (3.75) and the terms in the summation are controlled by (3.73). For the last term we again use the trivial bound $|G_{nj}^{(k)}| \leq N$ and take m large. This yields (3.43).

For (3.44) we write

$$\begin{aligned} \mathcal{E}_k &= \sum_{ij}^{(k)} G_{ij}^{(k)} (h_{ki} h_{kj} - \delta_{ij} N^{-1}) - (m_N - m_N^{(k)}) \\ &= \sum_{ij}^{(k)} \tilde{G}_{ij}^{(k)} (h_{ki} h_{kj} - \delta_{ij} N^{-1}) \\ &\quad + \sum_{l=3}^m A_l + \sum_{i_m, j}^{(k)} A_{m-1}(i_m) h_{ki_m} G_{i_m j}^{(k)} h_{kj} - (m_N^{(k)} - \tilde{m}_N^{(k)}) - (m_N - m_N^{(k)}). \end{aligned} \quad (3.85)$$

The claim follows from (3.74), (3.77), (3.78), (3.42), and the deterministic bound [6, Lemma 6.5]

$$|m_N(z) - m_N^{(k)}(z)| \leq (N\eta)^{-1}. \quad (3.86)$$



Fig. 2. The graph on the left corresponds to the sum $N^{-5} \sum_{i_1, \dots, i_5}^{(k)} \tilde{G}_{i_1 i_2}^{(k)} \tilde{G}_{i_2 i_3}^{(k)} \tilde{G}_{i_2 i_5}^{(k)} \tilde{G}_{i_4 i_5}^{(k)}$. If the white vertex appearing in the graph on the right is associated with the index j , then the graph corresponds to the sum $N^{-4} \sum_{i_1, \dots, i_4}^{(k)} \tilde{G}_{i_1 i_2}^{(k)} \tilde{G}_{i_2 i_3}^{(k)} \tilde{G}_{i_2 j}^{(k)} \tilde{G}_{i_4 j}^{(k)}$.

□

For the proofs of Propositions 3.10 and 3.11 we introduce a graphical notation to encode terms appearing in our high moment estimates. This is inspired by the methodology of [9]. We require the following combinatorial definition and lemma.

Definition 3.12. We define a labeled colored undirected graph \mathcal{G} to be a graph on m black vertices $\{1, 2, \dots, m\}$ and on either 0, 1 or 2 white vertices, labeled as $m + 1$ and $m + 2$ if they exist. Furthermore we demand that each white vertex is associated to a specific index which we will denote by $i, j \in \{1, 2, \dots, N + 1\}$. \mathcal{G} may have multiple edges and self loops. In this section we will refer to such labeled colored undirected graphs simply as graphs.

We will use the graphs defined above to encode various summations of monomials of Green’s function elements. Black vertices will encode an index to be summed over, while white vertices will encode fixed indices that are not summed over.

Let \mathcal{G} be a graph as defined in Definition 3.12. We associate to \mathcal{G} and an index $k \in \llbracket 1, N + 1 \rrbracket$ the following sum,

$$\mathcal{V}(\mathcal{G}, k) := \frac{1}{N^m} \sum_{i_1, \dots, i_m}^{(k)} \prod_{(a,b) \in \mathcal{G}} |\tilde{G}_{i_a i_b}^{(k)}(z)|. \tag{3.87}$$

When a or b equal the white vertex then i_a and i_b denote the associated index, as defined in Definition 3.12. Note that since $|\tilde{G}_{ij}^{(k)}| = |\tilde{G}_{ji}^{(k)}|$ the above sum is well-defined even though \mathcal{G} is undirected. An example of a graph and its associated sum is given in Figure 2.

Lemma 3.13. Let \mathcal{G} be a connected graph with m black vertices and E edges. If \mathcal{G} has no white vertices then

$$\mathbb{1}_{\Xi^{(k)}} \mathcal{V}(\mathcal{G}, k) \leq C^E (N\eta)^{-(m-1)/2}. \tag{3.88}$$

If \mathcal{G} has at least one white vertex, then

$$\mathbb{1}_{\Xi^{(k)}} \mathcal{V}(\mathcal{G}, k) \leq C^E (N\eta)^{-m/2}. \tag{3.89}$$

Above $C = \max\{2/c_*, 1\}$, where c_* is the constant in (3.11).

Remark 3.14. We pause to point out the usage of the Ward identity (3.7) in the following proof, as it is the only non-graphical tool used. Roughly speaking, the point of the graphical notation is to organize the sum into terms to which we can apply first the Cauchy–Schwarz inequality and then the Ward identity. The Ward identity is what results in the factors $(N\eta)^{-1}$ which is important for closing the bootstrap argument used to prove the local law.

Proof. In the following calculations we assume that $\Xi^{(k)}$ holds and we omit it from the notation. The reader should therefore keep in mind that $|\tilde{G}_{ij}^{(k)}(z)| \leq C$.

First suppose that \mathcal{G} has no white vertices. We can assume that m is at least 2 or the statement is trivial. Let \mathcal{T} be a spanning tree of \mathcal{G} . By a relabeling of the vertices, we can assume that the vertex m is a leaf of \mathcal{T} and that the vertex $m - 1$ is adjacent to vertex m . We will use the Cauchy–Schwarz inequality and the Ward identity (3.7) to sum over the summation index associated with the black vertex m . Let \mathcal{T}_1 be the tree obtained by removing the vertex m and the edge $(m - 1, m)$ from \mathcal{T} . We have,

$$\mathcal{V}(\mathcal{G}, k) \leq C^{E-m+1} \mathcal{V}(\mathcal{T}, k) = C^{E-m+1} \frac{1}{N^m} \sum_{i_1, \dots, i_m}^{(k)} \prod_{(a,b) \in \mathcal{T}} |\tilde{G}_{i_a i_b}^{(k)}|$$

$$\begin{aligned}
&= C^{E-m+1} \frac{1}{N^{m-1}} \sum_{i_1, \dots, i_{m-1}}^{(k)} \prod_{(a,b) \in \mathcal{T}_1} |\tilde{G}_{i_a i_b}^{(k)}| \frac{1}{N} \sum_{i_m}^{(k)} |\tilde{G}_{i_{m-1} i_m}^{(k)}| \\
&\leq C^{E-m+1} \frac{1}{N^{m-1}} \sum_{i_1, \dots, i_{m-1}}^{(k)} \prod_{(a,b) \in \mathcal{T}_1} |\tilde{G}_{i_a i_b}^{(k)}| \left(\frac{1}{N} \sum_{i_m}^{(k)} |\tilde{G}_{i_{m-1} i_m}^{(k)}|^2 \right)^{1/2} \\
&\leq C^{E-m+2} \frac{1}{(N\eta)^{1/2}} \mathcal{V}(\mathcal{T}_1, k). \tag{3.90}
\end{aligned}$$

The second to last line follows from Cauchy–Schwarz and in the last line we have used the Ward identity,

$$\frac{1}{N} \sum_{i_m}^{(k)} |\tilde{G}_{i_{m-1} i_m}^{(k)}|^2 = \frac{\text{Im}[\tilde{G}_{i_{m-1} i_{m-1}}^{(k)}]}{N\eta} \leq \frac{C}{N\eta}. \tag{3.91}$$

We can now continue this procedure with tree \mathcal{T}_1 which consists of $m-1$ black vertices $\{1, 2, \dots, m-1\}$. That is, we choose a leaf of the tree \mathcal{T}_1 and repeat the above procedure of summing over the summation index associated with the leaf, applying Cauchy–Schwarz and the Ward identity. To be more precise, we can assume that the vertex $m-1$ is a leaf attached to the vertex $m-2$. If we denote by \mathcal{T}_2 the subtree obtained by removing the vertex $m-1$ and edge $(m-2, m-1)$ we obtain,

$$\begin{aligned}
\mathcal{V}(\mathcal{T}_1, k) &= \frac{1}{N^{m-1}} \sum_{i_1, \dots, i_{m-1}}^{(k)} \prod_{(a,b) \in \mathcal{T}_1} |\tilde{G}_{i_a i_b}^{(k)}| \\
&= \frac{1}{N^{m-2}} \sum_{i_1, \dots, i_{m-2}}^{(k)} \prod_{(a,b) \in \mathcal{T}_2} |\tilde{G}_{i_a i_b}^{(k)}| \frac{1}{N} \sum_{i_{m-1}}^{(k)} |\tilde{G}_{i_{m-2} i_{m-1}}^{(k)}| \\
&\leq \frac{1}{N^{m-2}} \sum_{i_1, \dots, i_{m-2}}^{(k)} \prod_{(a,b) \in \mathcal{T}_2} |\tilde{G}_{i_a i_b}^{(k)}| \left(\frac{1}{N} \sum_{i_{m-1}}^{(k)} |\tilde{G}_{i_{m-2} i_{m-1}}^{(k)}|^2 \right)^{1/2} \\
&\leq C \frac{1}{(N\eta)^{1/2}} \mathcal{V}(\mathcal{T}_2, k). \tag{3.92}
\end{aligned}$$

This yields another factor of $C(N\eta)^{-1/2}$. We then inductively repeat this procedure on the subtree \mathcal{T}_2 consisting of $m-2$ back vertices, and so on. Each summation removes a leaf vertex and the edge adjacent to it, and results in a factor $C(N\eta)^{-1/2}$.

We get a sequence of subtrees $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_{m-1}$, where the n th tree \mathcal{T}_n is obtained after the n th summation. The last tree \mathcal{T}_{m-1} consists only of the root vertex and clearly $\mathcal{V}(\mathcal{T}_{m-1}, k) = 1$ by definition. We see that

$$\begin{aligned}
\mathcal{V}(\mathcal{G}, k) &\leq C^{E-m+2} \frac{1}{(N\eta)^{1/2}} \mathcal{V}(\mathcal{T}_1, k) \leq C^{E-m+3} \frac{1}{(N\eta)} \mathcal{V}(\mathcal{T}_2, k) \leq \dots \\
&\leq (C(N\eta)^{-1/2})^{m-1} \mathcal{V}(\mathcal{T}_{m-1}, k) \leq C^E (N\eta)^{-(m-1)/2}. \tag{3.93}
\end{aligned}$$

This finishes the proof of (3.88). Suppose that \mathcal{G} has a single white vertex. We repeat the same procedure as above but choose the spanning tree to be rooted at the white vertex. After summing over m black vertices we will have

$$\mathcal{V}(\mathcal{G}, k) \leq C^E (N\eta)^{-m/2} \mathcal{V}(\mathcal{T}_m, k) \tag{3.94}$$

where \mathcal{T}_m consists of only the white vertex. Clearly $\mathcal{V}(\mathcal{T}_m, k) = 1$ by definition and we get the second inequality (3.89) in this case.

Now we can assume that \mathcal{G} has two white vertices. Instead of taking out a spanning tree we choose a spanning forest consisting of two disjoint trees $\mathcal{T} \cup \mathcal{T}'$ each of which is rooted at a white vertex. We then have,

$$\mathcal{V}(\mathcal{G}, k) \leq C^{E-m} \mathcal{V}(\mathcal{T}, k) \mathcal{V}(\mathcal{T}', k) \leq C^E (N\eta)^{-m/2}, \tag{3.95}$$

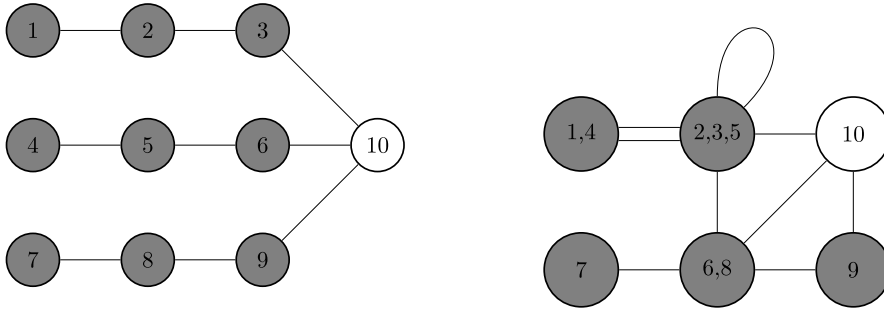


Fig. 3. The graph on the left is the original graph \mathcal{G} with 9 black vertices and 1 white vertices. The graph on the right is the new graph $\mathcal{G}(\mathcal{P})$ after grouping together black vertices with respect to the partition $\mathcal{P} = \{\{1, 4\}, \{2, 3, 5\}, \{6, 8\}, \{7, 9\}\}$.

where the second inequality follows from the argument in the case that \mathcal{G} has a single white vertex. This finishes the proof. \square

We will lastly need the following construction for forming new graphs by grouping together black vertices.

Definition 3.15. Let \mathcal{G} be a colored graph with m black vertices labeled $\{1, 2, \dots, m\}$ and 0, 1 or 2 white vertices. Let \mathcal{P} be a partition on $\{1, 2, \dots, m\}$. Note that \mathcal{P} induces a partition on the black vertices of \mathcal{G} . The graph $\mathcal{G}(\mathcal{P})$ is formed as follows.

Firstly, $\mathcal{G}(\mathcal{P})$ has the same number of white vertices as \mathcal{G} and the same associated index. To each block of \mathcal{P} we create one black vertex for $\mathcal{G}(\mathcal{P})$. We now define the edges of $\mathcal{G}(\mathcal{P})$. We first construct a map $\phi : V(\mathcal{G}) \rightarrow V(\mathcal{G}(\mathcal{P}))$ from the vertices of \mathcal{G} to the vertices of $\mathcal{G}(\mathcal{P})$. The white vertices of \mathcal{G} are just sent to the corresponding white vertices of $\mathcal{G}(\mathcal{P})$. A black vertex of \mathcal{G} is associated with a unique block of \mathcal{P} – the block that its label lies in – and we map this vertex to the black vertex of $\mathcal{G}(\mathcal{P})$ corresponding to this block. Then, for each edge (a, b) of \mathcal{G} we construct an edge $(\phi(a), \phi(b))$ of $\mathcal{G}(\mathcal{P})$.

While lengthy, the above definition is very simple. Figure 3 illustrates the process.

Proof of Proposition 3.10. In the following proof we work on the event $\Xi^{(k)}(z)$ and we omit it from the notation. The reader should therefore keep in mind that $|\tilde{G}_{ij}^{(k)}(z)| \leq C$. We will compute high order moments of the terms in all three expressions (3.72), (3.73) and (3.74). The claim then follows by a high-moment Chebyshev inequality.

We begin with (3.72). Its $2p$ -th moment is

$$\mathbb{E} \left| \sum_{i_1 \dots i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1, i_2}^{(k)} h_{ki_2} \dots h_{ki_l} \tilde{G}_{i_l, j}^{(k)} \right|^{2p} = \sum_{i_1 \dots i_{2p} l}^{(k)} \mathbb{E} [h_{ki_1} X_{i_1 i_2} \dots X_{i_l j} h_{ki_{l+1}} X_{i_{l+1} i_{l+2}} \dots X_{i_{2p} j}], \quad (3.96)$$

where X_{ij} denotes either $\tilde{G}_{ij}^{(k)}$ or its complex conjugate $(\tilde{G}_{ij}^{(k)})^*$. For us the distinction will not be important as we will be bounding the terms in the above expression by their absolute value. Since $\tilde{G}^{(k)}$ is independent of the k -th row of H we can first take the expectation over this row. We then re-arrange the above expectation into all possible partitions of the summation indices $\{1, 2, \dots, 2pl\}$. Let \mathcal{P} be such a partition. Note that due to the fact that the h_{ij} are independent centered random variables, the size of each block of \mathcal{P} is at least 2, otherwise the expectation will vanish. Therefore, (3.96) can be bounded by

$$\begin{aligned} & \mathbb{E} \left| \sum_{i_1 \dots i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1, i_2}^{(k)} h_{ki_2} \dots h_{ki_l} \tilde{G}_{i_l, j}^{(k)} \right|^{2p} \\ & \leq \sum_{\mathcal{P}} \frac{C_{\mathcal{P}}}{q^{2l(p-2|\mathcal{P}|)}} \left(\frac{1}{N^{|\mathcal{P}|}} \sum_{\{i_1, i_2, \dots, i_{2pl}\} \sim \mathcal{P}} \mathbb{E} [|X_{i_1 i_2} \dots X_{i_l j} X_{i_{l+1} i_{l+2}} \dots X_{i_{2p} j} |] \right). \end{aligned} \quad (3.97)$$

Above, the first sum is over all possible partitions of the set $\{1, 2, \dots, 2pl\}$ such that the size of each block of \mathcal{P} is at least 2, and $|\mathcal{P}|$ denotes the number of blocks in \mathcal{P} . The constant $C_{\mathcal{P}}$ depends only on the partition \mathcal{P} . The second sum is over

all the possible assignments of the indices $\{i_1, i_2, \dots, i_{2pl}\}$ which induces the same partition as \mathcal{P} of the set $\{1, 2, \dots, 2pl\}$ (we recall that $\{i_1, i_2, \dots, i_{2pl}\}$ induces a partition of the set $\{1, 2, \dots, 2pl\}$, such that a and b are in the same block if and only if $i_a = i_b$).

We first construct the graph \mathcal{G} with black vertices labelled by $1, 2, \dots, 2lp$ and one white vertex associated with the index j . We create an edges between the vertices $(i, i + 1)$ for $i = nl + 1, nl + 2, \dots, (n + 1)l - 1$ and $n = 0, 1, 2, \dots, 2p - 1$, and an edge between the black vertices nl and the white vertex j , for $n = 1, \dots, 2p$. The graph \mathcal{G} represents the unpartitioned sum (3.96).

To each partition \mathcal{P} we then get a graph $\mathcal{G} = \mathcal{G}(\mathcal{P})$ as defined above. By Lemma 3.13, we see

$$\frac{1}{N^{|\mathcal{P}|}} \sum_{\{i_1, i_2, \dots, i_{2pl}\} \sim \mathcal{P}}^{(k)} \mathbb{E}[|X_{i_1 i_2} \dots X_{i_l j} X_{i_{l+1} i_{l+2}} \dots X_{i_{2l} j}|] = \mathbb{E}[\mathcal{V}(\mathcal{G}(\mathcal{P}), k)] \leq C^{2lp} (N\eta)^{-|\mathcal{P}|/2}, \quad (3.98)$$

where we used the fact that since \mathcal{G} is a connected graph, $\mathcal{G}(\mathcal{P})$ is also a connected graph. The total number of partitions of the set $\{1, 2, \dots, 2pl\}$ is bounded by $(2pl)!$ and we get

$$\mathbb{E} \left| \sum_{i_1 \dots i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots h_{ki_l} \tilde{G}_{i_l j}^{(k)} \right|^{2p} \leq C_{l,p} \max_{\mathcal{P}} \frac{1}{q^{2lp-2|\mathcal{P}|}} \frac{1}{(N\eta)^{|\mathcal{P}|/2}} \leq C_{l,p} \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/4}} \right)^{2lp}. \quad (3.99)$$

This yields (3.72).

Next we prove (3.73). We have,

$$\mathbb{E} \left| \sum_{i_1 \dots i_l}^{(k)} \tilde{G}_{i_1 i_1}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} \dots h_{ki_l} \tilde{G}_{i_l j}^{(k)} \right|^{2p} = \sum_{i_1 \dots i_{2l}}^{(k)} \mathbb{E}[X_{i_1 i_1} h_{ki_1} \dots X_{i_l j} X_{i_{l+1}} h_{ki_{l+1}} \dots X_{i_{2l} j}], \quad (3.100)$$

where again we denote by X_{ij} either $\tilde{G}_{ij}^{(k)}$ or its complex conjugate $(\tilde{G}_{ij}^{(k)})^*$. We divide the sum up into partitions \mathcal{P} of $\{1, 2, \dots, 2lp\}$. We define the graph \mathcal{G} representing the unpartitioned sum above as follows. It has $2lp$ black vertices and two white vertices if $i \neq j$ and one white vertex if $i = j$. We create an edge between black vertices $(i, i + 1)$ for $i = ln + 1, ln + 2, \dots, (n + 1)l - 1$ and $n = 0, \dots, 2p - 1$, an edge between the white vertex i and black vertices $nl + 1$, for $n = 0, 1, \dots, 2p - 1$, and an edge between the white vertex j and black vertices nl for $n = 1, 2, \dots, 2p$. We have that (3.100) is bounded by

$$\mathbb{E} \left| \sum_{i_1 \dots i_l}^{(k)} \tilde{G}_{i_1 i_1}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots h_{ki_l} \tilde{G}_{i_l j}^{(k)} \right|^{2p} \leq \sum_{\mathcal{P}} \frac{C_{\mathcal{P}}}{q^{2lp-2|\mathcal{P}|}} \mathcal{V}(\mathcal{G}(\mathcal{P}), k) \quad (3.101)$$

where again the summation is over all partitions on $\llbracket 1, 2lp \rrbracket$ with each block of size at least 2. Applying (3.89) we obtain

$$\mathbb{E} \left| \sum_{i_1 \dots i_l}^{(k)} \tilde{G}_{i_1 i_1}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots h_{ki_l} \tilde{G}_{i_l j}^{(k)} \right|^{2p} \leq C_{l,p} \max_{\mathcal{P}} \frac{1}{q^{2lp-2|\mathcal{P}|}} \frac{1}{(N\eta)^{|\mathcal{P}|/2}} \leq C_{l,p} \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/4}} \right)^{2lp} \quad (3.102)$$

and we obtain (3.73).

Lastly, we prove (3.74). The $2p$ -th moment is

$$\mathbb{E} \left| \sum_{i_1 \dots i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots \tilde{G}_{i_{l-1} i_l}^{(k)} h_{ki_l} \right|^{2p} = \sum_{i_1 \dots i_{2l}}^{(k)} \mathbb{E}[h_{ki_1} X_{i_1 i_2} \dots h_{ki_l} h_{ki_{l+1}} X_{i_{l+1} i_{l+2}} \dots X_{i_{2l-1} i_{2l}} h_{ki_{2l}}]. \quad (3.103)$$

Again we divide the sum up into partitions on $\{1, 2, \dots, 2lp\}$. We associate the following graph \mathcal{G} to the unpartitioned sum. \mathcal{G} has $2lp$ black vertices and no white vertices. There is an edge between $(i, i + 1)$ for $i = nl + 1, \dots, (n + 1)l - 1$ and $n = 0, \dots, 2p - 1$. Construct the graph $\mathcal{G}(\mathcal{P})$ as above. We have

$$\mathbb{E} \left| \sum_{i_1 \dots i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots \tilde{G}_{i_{l-1} i_l}^{(k)} h_{ki_l} \right|^{2p} \leq \sum_{\mathcal{P}} \frac{C_{\mathcal{P}}}{q^{2lp-2|\mathcal{P}|}} \mathcal{V}(\mathcal{G}(\mathcal{P}), k). \quad (3.104)$$

Let $\mathcal{G}_1, \dots, \mathcal{G}_r$ be the connected components of $\mathcal{G}(\mathcal{P})$. Since \mathcal{G} has $2p$ connected components we have $r \leq \min\{|\mathcal{P}|, 2p\}$. Therefore by (3.88), we have

$$\mathcal{V}(\mathcal{G}(\mathcal{P}), k) = \prod_i \mathcal{V}(\mathcal{G}_i, k) \leq C^{2p(l-1)} \prod_i \frac{1}{(N\eta)^{(|\mathcal{G}_i|-1)/2}} \leq \frac{C^{2p(l-1)}}{(N\eta)^{(|\mathcal{P}|-r)/2}}, \quad (3.105)$$

where $|\mathcal{G}_i|$ is the number of vertices of graph \mathcal{G}_i . We then see that (3.104) is bounded by

$$\begin{aligned} \mathbb{E} \left| \sum_{i_1, \dots, i_l}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \dots \tilde{G}_{i_{l-1} i_l}^{(k)} h_{ki_l} \right|^{2p} &\leq \sum_{\mathcal{P}} \frac{C_{\mathcal{P}}}{q^{2pl-2|\mathcal{P}|}} \frac{C^{2p(l-1)}}{(N\eta)^{(|\mathcal{P}|-r)/2}} \\ &\leq C_{l,p} \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/4}} \right)^{2p(l-2)}. \end{aligned} \quad (3.106)$$

This completes the proof. \square

Proof of Proposition 3.11. The bounds (3.75), (3.76) and (3.77) follow from more general statements [22, Lemma A.1]. We only prove (3.78) here. The proof is similar to that of (3.74), but with a more careful analysis of graphs.

The $2p$ -th moment of the lefthand side of (3.78) is bounded by

$$\begin{aligned} \mathbb{E} \left| \sum_{i_1, i_2, i_3}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \tilde{G}_{i_2 i_3}^{(k)} h_{ki_3} \right|^{2p} \\ \leq \sum_{\mathcal{P}} \frac{C_{\mathcal{P}}}{q^{6p-2|\mathcal{P}|}} \left(\frac{1}{N^{|\mathcal{P}|}} \sum_{\{i_1, i_2, \dots, i_{6p}\} \sim \mathcal{P}} \mathbb{E}[X_{i_1 i_2} X_{i_2 i_3} X_{i_4 i_5} \dots X_{i_{6p-2} i_{6p-1}} X_{i_{6p-1} i_{6p}}] \right), \end{aligned} \quad (3.107)$$

where the notations are the same as in (3.97), and the sum over \mathcal{P} runs through all the partitions of $\{1, 2, \dots, 6p\}$ where the size of each block is at least two.

We construct the graph \mathcal{G} corresponding to the unpartitioned sum in (3.107). The graph \mathcal{G} has no white vertices and black vertices labelled by $1, 2, \dots, 6p$. We add edges $(3n+1, 3n+2)$ and $(3n+2, 3n+3)$ for $n = 0, 1, 2, \dots, 2p-1$. The graph \mathcal{G} consists of $2p$ connected components and each component is a chain of three vertices and two edges.

For each partition \mathcal{P} we construct the graph $\mathcal{G}(\mathcal{P})$ as before. Let $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_r$ be the connected components of $\mathcal{G}(\mathcal{P})$. Consider the inverse images of the components in the original graph \mathcal{G} . By a relabelling, let the inverse image of each of the components $\mathcal{G}_1, \dots, \mathcal{G}_{r_1}$ consist only of a single chain of \mathcal{G} . The inverse image of the components $\mathcal{G}_{r_1+1}, \dots, \mathcal{G}_r$ consists of at least two chains. We have the relation

$$r_1 + 2(r - r_1) \leq 2p, \quad (3.108)$$

since there are in total $2p$ chains in \mathcal{G} .

Each chain consists of three vertices and we know that each block of \mathcal{P} is of size at least two. Therefore, for any $1 \leq i \leq r_1$, the component \mathcal{G}_i consists of a single vertex which corresponds to one block of the partition \mathcal{P} consisting of exactly three numbers. Since the other blocks of \mathcal{P} are of at least size two, we have the relation

$$3r_1 + 2(|\mathcal{P}| - r_1) \leq 6p. \quad (3.109)$$

Adding the inequalities (3.108) and (3.109) yields,

$$|\mathcal{P}| + r \leq 4p. \quad (3.110)$$

The same argument as in the proof of (3.73) yields

$$\mathbb{E} \left| \sum_{i_1, i_2, i_3}^{(k)} h_{ki_1} \tilde{G}_{i_1 i_2}^{(k)} h_{ki_2} \tilde{G}_{i_2 i_3}^{(k)} h_{ki_3} \right|^{2p} \leq C_{l,p} \max_{\mathcal{P}} \frac{1}{q^{6p-2|\mathcal{P}|}} \frac{1}{(N\eta)^{(|\mathcal{P}|-r)/2}}. \quad (3.111)$$

If $|\mathcal{P}| \leq 2p$ then the righthand side of above expression is bounded by $C_{l,p}q^{-2p}$ (recall $r \leq |\mathcal{P}|$). If $|\mathcal{P}| \geq 2p$ then we apply (3.110) and obtain

$$\frac{1}{q^{6p-2|\mathcal{P}|}} \frac{1}{(N\eta)^{(|\mathcal{P}|-r)/2}} \leq \frac{1}{q^{2p+(4p-2|\mathcal{P}|)}} \frac{1}{(N\eta)^{(2|\mathcal{P}|-4p)/2}} \leq \left(\frac{1}{q} + \frac{1}{\sqrt{N\eta}} \right)^{2p}. \quad (3.112)$$

This finishes the proof. \square

4. Universality of eigenvalue statistics

In this section we prove Theorem 2.9. We define the following matrix valued process H_t for any random Laplacian-type matrix H . The off-diagonal entries are given by $H_{ij}(t) = h_{ij}(t)$, $i \neq j$ where $h_{ij}(t)$ are independent Ornstein–Uhlenbeck process with initial data h_{ij} satisfying

$$dh_{ij}(t) = \frac{dB_{ij}(t)}{\sqrt{N}} - \frac{1}{2}h_{ij}(t)dt, \quad 1 \leq i < j \leq N+1, \quad (4.1)$$

where $\{B_{ij}\}_{1 \leq i < j \leq N+1}$ are independent standard Brownian motions. The diagonal elements satisfy $H_{kk}(t) = -\sum_i^{(k)} h_{ik}(t)$. We have $H_0 = H$ and the equality in law

$$H_t \stackrel{d}{=} e^{-t/2}H + (1 - e^{-t})^{1/2}W, \quad (4.2)$$

where W is a Gaussian Laplacian matrix as defined in (2.35). Furthermore, W is independent from H .

The matrices H_t all have the trivial eigenvalue 0 and corresponding eigenvector \mathbf{e} . This will destroy the eigenvalue universality at 0 and so we must remove the trivial eigenvalue. We denote the nontrivial eigenvalues of H_t as $\lambda_1(t) \leq \lambda_2(t) \leq \dots \leq \lambda_N(t)$ and corresponding eigenvectors $u_1(t), u_2(t), \dots, u_N(t)$. We will sometimes omit the time parameter for simplicity of notation. We introduce the following Green's function of the N nontrivial eigenvalues and eigenvectors. It is defined by

$$\hat{G}(z, H_t) := \sum_{i=1}^N \frac{u_i(t)u_i(t)^*}{\lambda_i(t) - z} = R(R^*H_tR - z)^{-1}R^*, \quad (4.3)$$

where R is the following matrix $(N+1) \times N$ dimensional matrix (see Section 2.4). First fix any $(N+1) \times (N+1)$ orthogonal matrix \tilde{R} with last column $\mathbf{e} = (N+1)^{-1/2}(1, \dots, 1)^*$ (the specific choice does not matter). We then define R as the first N columns of \tilde{R} .

Remark 4.1. Clearly $\hat{G}(z, H_t)$ differs from the usual Green's function by only a rank one matrix,

$$\hat{G}(z, H_t) - (H_t - z)^{-1} = \frac{1}{z}\mathbf{e}\mathbf{e}^*. \quad (4.4)$$

We will need the following lemma.

Lemma 4.2. *The nontrivial eigenvalues of H_t are distributed as the eigenvalues of the $N \times N$ matrix*

$$\hat{H}_t := (e^{-t/2}R^*HR + (1 - e^{-t})^{1/2}(R^*DR + gI_N)) + (1 - e^{-t})^{1/2}GOE \quad (4.5)$$

where GOE represents an independent $N \times N$ matrix drawn from the Gaussian orthogonal ensemble, $D = \text{diag}\{d_1, d_2, \dots, d_{N+1}\}$ is a diagonal matrix, with d_i 's independent Gaussian random variables with variance $(N+1)/N$, and g is a Gaussian random variable with variance $1/N$. Denote the matrix

$$\hat{A}_t := R^*(e^{-t/2}H + (1 - e^{-t})^{1/2}D)R + (1 - e^{-t})^{1/2}gI_N. \quad (4.6)$$

We then have with overwhelming probability uniformly for $z \in \mathcal{D}_L$,

$$|m_{\hat{A}_t}(z) - m_{\text{fc}}(z)| \leq \xi\Psi, \quad (4.7)$$

given that $t \leq \xi/q$, where the control parameter Ψ is introduced in (2.21).

Proof. The first statement holds by Proposition 2.10. For the second statement, consider the auxillary matrices

$$A^{(1)} := e^{-t/2}H + (1 - e^{-t})^{1/2}D, \quad A^{(2)} := e^{-t/2}H + (1 - e^{-t})^{1/2}(D + gI_{N+1}). \quad (4.8)$$

Since $\hat{A}_t = R^* A^{(2)} R$, which is the upper left $N \times N$ minor of $\hat{R}^* A^{(2)} R$, from the deterministic bound [6, Lemma 6.5], the Stieltjes transforms of empirical eigenvalue distributions of $A^{(2)}$ and \hat{A}_t differ by at most $(N\eta)^{-1}$. Moreover, the Stieltjes transform of the empirical eigenvalue distribution of $A^{(2)}$ is a shift of that of $A^{(1)}$:

$$m_{A^{(2)}}(z) = m_{A^{(1)}}(z - (1 - e^{-t})^{1/2}g). \quad (4.9)$$

Therefore, to prove (4.7), we need to prove that uniformly for $z \in \mathcal{D}_L$ and $t \leq \xi/q$ we have with overwhelming probability,

$$|m_{A^{(1)}}(z - (1 - e^{-t})^{1/2}g) - m_{\text{fc}}(z)| \leq \xi\Psi. \quad (4.10)$$

For this, we claim that the proof of Theorem 2.5 holds for $A^{(1)}$. We summarize here the changes. We need to modify Proposition 3.6. Instead of (3.31) we see that we obtain, by the same proof using the McDiarmid inequality, (the extra term $(t/N)^{1/2}$ on the righthand side results from the d_k 's) the bound

$$\left| \frac{1}{N} \sum_k \frac{1}{\sum_i^{(k)} e^{-t/2}h_{ik} + (1 - e^{-t})^{1/2}d_k - z - m_{\text{fc}}^{(1)}(z)} - m_{\text{fc}}^{(1)}(z) \right| \leq C\xi \left(\frac{1}{q} + \frac{\sqrt{t}}{\sqrt{N}} \right), \quad (4.11)$$

where $m_{\text{fc}}^{(1)}(z)$ is the free convolution of the semicircle with a Gaussian of variance

$$\mathbb{E} \left[\left(\sum_i^{(k)} e^{-t/2}h_{ik} + (1 - e^{-t})^{1/2}d_k \right)^2 \right] = 1 + \frac{1 - e^{-t}}{N}. \quad (4.12)$$

More explicitly, $m_{\text{fc}}^{(1)}$ is the unique solution to

$$m_{\text{fc}}^{(1)}(z) = \int \frac{\rho_G(x) dx}{(1 + (1 - e^{-t})/N)^{1/2}x - z - m_{\text{fc}}^{(1)}(z)}. \quad (4.13)$$

It's easy to check that the bounds of Lemma 3.4 also hold for $m_{\text{fc}}^{(1)}$ after changing the constants a little.

We can then run through the remainder of the proof of Theorem 2.5, simply by replacing h_{ij} by $e^{-t/2}h_{ij}$, and G , $G^{(k)}$ and $\tilde{G}^{(k)}$ by the corresponding quantities associated to $A^{(1)}$. The remainder of the proof is unchanged with the exception of the estimate (3.77). For the matrix $A^{(1)}$ we instead have the following estimate

$$\mathbb{1}_{\Xi^{(k)}(z)} \sum_{i,j}^{(k)} \tilde{G}_{ij}^{(k)} (e^{-t}h_{ki}h_{kj} - \delta_{ij}N^{-1}) \leq \xi \left(\frac{1}{q} + \frac{1}{(N\eta)^{1/2}} \right) + Ct, \quad (4.14)$$

where the constant C depends only on c_* as in (3.11). The result of the above argument following the proof of Theorem 2.5 is the following local law for $A^{(1)}$. We have the estimate,

$$|m_{A^{(1)}}(z) - m_{\text{fc}}^{(1)}(z)| \leq \xi(\Psi + (t/N)^{1/2}) + Ct \quad (4.15)$$

with overwhelming probability uniformly for $z \in \mathcal{D}_L$. In order to conclude (4.10), we need an estimate on $|m_{\text{fc}}^{(1)}(z - (1 - e^{-t})^{1/2}g) - m_{\text{fc}}(z)|$. For $m_{\text{fc}}^{(1)}$, by (4.13), with overwhelming probability uniformly for $z \in \mathcal{D}_L$, we have

$$\begin{aligned} & m_{\text{fc}}^{(1)}(z - (1 - e^{-t})^{1/2}g) \\ &= \int_{-\xi}^{\xi} \frac{\rho_G(x) dx}{(1 + (1 - e^{-t})/N)^{1/2}x + (1 - e^{-t})^{1/2}g - z - m_{\text{fc}}^{(1)}(z - (1 - e^{-t})^{1/2}g)} + O(N^{-D}) \\ &= \int \frac{\rho_G(x) dx}{x - z - m_{\text{fc}}^{(1)}(z - (1 - e^{-t})^{1/2}g)} + O\left(\frac{\xi\sqrt{t}}{\sqrt{N}}\right). \end{aligned} \quad (4.16)$$

In the second quality we used the fact that with overwhelming probability $|g| \leq \xi/\sqrt{N}$ and that for $z \in \mathcal{D}_L$ and $-\xi \leq x \leq \xi$ the denominator is uniformly bounded away from 0. Using (4.16) and the defining relation (2.15) of $m_{\text{fc}}(z)$ will yield the following estimate,

$$|m_{\text{fc}}^{(1)}(z - (1 - e^{-t})^{1/2}g) - m_{\text{fc}}(z)| \leq C \frac{\xi \sqrt{t}}{\sqrt{N}}. \quad (4.17)$$

The proof is similar to the proof of Proposition 3.9; the argument is essentially the same as that of the proof of Lemma 3.6 of [37] and we do not provide all the details. By a continuity argument, one can first show (4.17) for $z \in \mathcal{D}_L$ with $\text{Im}[z] \geq 2$. Since $m_{\text{fc}}^{(1)}$ and $m_{\text{fc}}(z)$ are both Lipschitz and satisfy a similar self-consistent equation, one can prove (4.17) for z with slightly smaller imaginary part. By repeating this process, one can finally conclude (4.17) for all $z \in \mathcal{D}_L$.

Combining (4.15) and (4.17) yields

$$|m_{\hat{A}}(z) - m_{\text{fc}}(z)| \leq \xi(\Psi + (t/N)^{1/2}) + Ct, \quad (4.18)$$

and (4.7) follows since $t \leq \xi/q$. \square

We need the following notion of a derivative of a smooth function F on the space of $(N+1) \times (N+1)$ real symmetric matrices that respects the structure of Laplacian-type matrices. For $1 \leq i < j \leq N+1$ we define the matrices $X_{ij} = E^{ij} + E^{ji} - E^{ii} - E^{jj}$, where E^{ij} is the base matrix given by $(E^{ij})_{kl} = \delta_{ik}\delta_{jl}$. For indices $1 \leq i < j \leq N+1$ we define

$$\hat{\partial}_{ij}F(H) := \lim_{h \rightarrow 0} \frac{F(H + hX_{ij}) - F(H)}{h}. \quad (4.19)$$

Let H be a Laplacian-type matrix as in (2.10). We define the following deformation of H . Given indices $1 \leq i < j \leq N+1$ and number $0 \leq \theta^{ij} \leq 1$, the deformed matrix $\theta^{ij}H$ is defined by replacing h_{kl} by $\theta_{kl}^{ij}h_{kl}$, for $1 \leq k < l \leq N+1$ where $\theta_{kl}^{ij} = 1$ unless $\{k, l\} = \{i, j\}$ in which case $\theta_{ij}^{ij} = \theta_{ji}^{ij} = \theta^{ij}$.

The proof of Theorem 2.9 consists of two steps. First we prove that for small t , the eigenvalue statistics of the nontrivial eigenvalues of H and H_t are the same by a continuity argument. In the second step, we prove the universality of eigenvalue statistics for random Laplacian-type matrices with small Gaussian components; i.e. we prove eigenvalue universality for H_t with small t . Thus, Theorem 2.9 follows.

4.1. Short time comparison

In this section we prove that the eigenvalue statistics of H and H_t are the same for small times t .

Proposition 4.3. *Let H be a random Laplacian-type matrix as in Definition 2.2. Let H_t be defined as in (4.2). Choose $t = N^{-1+\epsilon}$, for sufficiently small $\epsilon \leq \beta/2$, where β is from (2.12). Then the gap statistics and local correlation functions of the nontrivial eigenvalues of H_0 and H_t are the same. More precisely, we first have*

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}^n} O(\alpha_1, \dots, \alpha_n) \left\{ \rho_{H_0}^{(n)} \left(E + \frac{\alpha_1}{N}, \dots, E + \frac{\alpha_n}{N} \right) - \rho_{H_t}^{(n)} \left(E + \frac{\alpha_1}{N}, \dots, E + \frac{\alpha_n}{N} \right) \right\} d\alpha_1 \dots d\alpha_n = 0, \quad (4.20)$$

for any test function $O \in C_c^\infty(\mathbb{R}^n)$. Fix $\kappa > 0$ and $k \in \llbracket kN, (1 - \kappa)N \rrbracket$. Then,

$$\lim_{N \rightarrow \infty} \mathbb{E}^{(H_t)} [O(N(\lambda_{k+1} - \lambda_k), \dots, N(\lambda_{n+k} - \lambda_k))] - \mathbb{E}^{(H_0)} [O(N(\lambda_{k+1} - \lambda_k), \dots, N(\lambda_{n+k} - \lambda_k))] = 0, \quad (4.21)$$

for $O \in C_c^\infty(\mathbb{R}^n)$.

We require the following estimate on the resolvent entries.

Proposition 4.4. *For any $s \geq 0$, we have with overwhelming probability the bound*

$$\max_{ij} \sup_{\theta^{kl}} |\hat{G}_{ij}(z, \theta^{kl}H_s)| = O(1) \quad (4.22)$$

for $z \in \{z = E + i\eta : |E| \leq L, \xi^3 N^{-1} \leq \eta \leq L\}$. Fix $\gamma > 0$. We have with overwhelming probability uniformly in s

$$\max_{ij} \sup_{\theta^{kl}} |\hat{G}_{ij}(z, \theta^{kl} H_s)| = O(\xi^3 N^\gamma) \quad (4.23)$$

for $z \in \{z = E + i\eta : |E| \leq L, N^{-1-\gamma} \leq \eta \leq L\}$.

Proof. We first prove (4.22) for the matrix H_s without any deformation. Since H_s is a random Laplacian-type matrix in the sense of Definition 2.2, we see from Theorem 2.5 and (4.4) that with overwhelming probability,

$$\max_{ij} |\hat{G}_{ij}(z, H_s)| \leq \max_{ij} |(H_s - z)_{ij}^{-1}| + \frac{1}{(N+1)|z|} = O(1) \quad (4.24)$$

uniformly for any $z \in \{z = E + i\eta : |E| \leq L, \xi^3 N^{-1} \leq \eta \leq L\}$. For the deformed matrix $\theta^{kl} H_s$, we have the resolvent identity,

$$\hat{G}_{ij}(z, \theta^{kl} H_s) = \hat{G}_{ij}(z, H_s) + (1 - \theta_{kl}^{kl}) h_{kl}(s) (\hat{G}(z, \theta^{kl} H_s) X_{kl} \hat{G}(z, H_s))_{ij}. \quad (4.25)$$

The matrix X_{kl} has exactly four nonzero entries and with overwhelming probability we have $|h_{kl}(s)| \leq \xi / \log(N)q$. For N large enough we therefore have with overwhelming probability,

$$\max_{ij} \sup_{\theta^{kl}} |\hat{G}_{ij}(z, \theta^{kl} H_s)| \leq \max_{ij} |\hat{G}_{ij}(z, H_s)| + \frac{1}{2} \max_{ij} \sup_{\theta^{kl}} |\hat{G}(z, \theta^{kl} H_s)|, \quad (4.26)$$

and (4.22) follows from rearranging the above expression. The following deterministic relation is a consequence of [5, Lemma 2.1],

$$\max_{ij} |\hat{G}_{ij}(E + i\eta/M, \theta^{kl} H_s)| \leq M \max_{ij} |\hat{G}_{ij}(E + i\eta, \theta^{kl} H_s)|. \quad (4.27)$$

We obtain (4.23) by taking $M = \xi^3 N^\gamma$. □

We will need the following lemma which is a slight modification of [11, Lemma A.2].

Lemma 4.5. *Let H be a random Laplacian-type matrix. Define H_t as in (4.2). Let F be a smooth function on the space of $(N+1) \times (N+1)$ real symmetric matrices satisfying*

$$\sup_{\substack{0 \leq s \leq t, \\ 1 \leq i < j \leq N+1}} \mathbb{E} \left[(N^2 |h_{ij}(s)|^3 + N |h_{ij}(s)|) \sup_{\theta^{ij}} |\hat{\partial}_{ij}^3 F(\theta^{ij} H_s)| \right] \leq B, \quad (4.28)$$

where the deformed matrix $\theta^{ij} H_s$ is defined as above. Then,

$$|\mathbb{E}[F(H_t)] - \mathbb{E}[F(H_0)]| \leq CtB. \quad (4.29)$$

Proof. From Itô's formula we have

$$\begin{aligned} \mathbb{E}[F(H_t)] - \mathbb{E}[F(H_0)] &= \int_0^t \mathbb{E} \left[\sum_{i < j} \hat{\partial}_{h_{ij}} F(H_s) dh_{ij}(s) + \frac{1}{2N} \hat{\partial}_{h_{ij}}^2 F(H_s) ds \right] \\ &= \int_0^t \mathbb{E} \left[\sum_{i < j} -\frac{1}{2} h_{ij}(s) \hat{\partial}_{ij} F(H_s) + \frac{1}{2N} \hat{\partial}_{ij}^2 F(H_s) \right] ds \end{aligned} \quad (4.30)$$

We expand the terms above. For some $0 \leq \theta_1^{ij}, \theta_2^{ij} \leq 1$ we have the equalities

$$\mathbb{E}[h_{ij}(s) \hat{\partial}_{ij} F(H_s)] = \mathbb{E} \left[h_{ij} \hat{\partial}_{ij} F_{h_{ij}=0} + h_{ij}^2 \hat{\partial}_{ij}^2 F_{h_{ij}=0} + \frac{1}{2} h_{ij}^3 \hat{\partial}_{ij}^3 F(\theta_1^{ij} H_s) \right], \quad (4.31)$$

and

$$\mathbb{E}[\hat{\partial}_{ij}^2 F(H_s)] = \mathbb{E}[\hat{\partial}_{ij}^2 F_{h_{ij}=0} + h_{ij} \hat{\partial}_{ij}^3 F(\theta_2^{ij} H_s)]. \quad (4.32)$$

This leads to

$$\left| \mathbb{E} \left[\sum_{i < j} -\frac{1}{2} h_{ij}(s) \hat{\partial}_{ij} F(H_s) + \frac{1}{2N} \hat{\partial}_{ij}^2 F(H_s) \right] \right| \leq N^2 \sup_{i < j} \mathbb{E} \left[\left(|h_{ij}(s)|^3 + \frac{1}{N} |h_{ij}(s)| \right) \sup_{\theta^{ij}} |\hat{\partial}_{ij}^3 F(\theta^{ij} H_s)| \right], \quad (4.33)$$

and (4.29) follows. \square

4.1.1. Comparison of correlation functions

We now prove that the locally averaged correlation functions of the nontrivial eigenvalues of the matrix H_0 coincide with those of H_t , i.e. (4.20) of Proposition 4.3. The main ingredient of the proof is the following Green's function comparison lemma. Its proof follows essentially from [32, Lemma 5.2].

Lemma 4.6. *Fix $n \in \mathbb{N}$, and let $\gamma, \delta > 0$ be sufficiently small. Then the following holds for any $\eta \in [N^{-1-\gamma}, N^{-1}]$, any sequence of positive integers k_1, k_2, \dots, k_n , any set of complex parameters $z_j^m = E_j^m \pm i\eta$, where $j \in \llbracket 1, k_m \rrbracket$, $m \in \llbracket 1, n \rrbracket$, $|E_j^m| \leq L$, and the \pm signs are arbitrary. Let $F \in C^\infty(\mathbb{R}^n)$ be a test function such that for any multi-index $m = (m_1, \dots, m_n)$ with $1 \leq |m| \leq 3$ and for any $\omega > 0$,*

$$\begin{aligned} \max \{ |\partial^m F(x)| : \|x\|_\infty \leq N^\omega \} &\leq N^{c_0 \omega}, \\ \max \{ |\partial^m F(x)| : \|x\|_\infty \leq N^2 \} &\leq N^{c_0}. \end{aligned} \quad (4.34)$$

Then, defining $\hat{G}_1 := \hat{G}(z, H_t)$ and $\hat{G}_0 := \hat{G}(z, H_0)$ we have for any $t > 0$,

$$\left| \mathbb{E} F \left(N^{-k_1} \text{Tr} \prod_{j=1}^{k_1} \hat{G}_1(z_j^1), \dots, N^{-k_n} \text{Tr} \prod_{j=1}^{k_n} \hat{G}_1(z_j^n) \right) - \mathbb{E} F(\hat{G}_1 \rightarrow \hat{G}_0) \right| \leq C_1 (\xi N^\gamma)^{C_2} t N/q. \quad (4.35)$$

Here, $\mathbb{E} F(\hat{G}_1 \rightarrow \hat{G}_0)$ is the expression obtained from the leftmost expression above by replacing \hat{G}_1 with \hat{G}_0 . The constants C_1 and C_2 depend on $n, k_1, \dots, k_n, m_1, \dots, m_n$, and the constants in (4.34).

Proof. For simplicity of notation, we only prove (4.35) for $n = 1$ and $k_1 = 1$ and write z instead of z^1 ; i.e., we prove that

$$\left| \mathbb{E} [F(N^{-1} \text{Tr} \hat{G}(z, H_t))] - \mathbb{E} [F(N^{-1} \text{Tr} \hat{G}(z, H_0))] \right| \leq C_1 (\xi N^\gamma)^{C_2} t N/q. \quad (4.36)$$

By Lemma 4.5, it is sufficient to prove that

$$\sup_{\substack{0 \leq s \leq t, \\ 1 \leq i < j \leq N+1}} \mathbb{E} \left[\left(N^2 |h_{ij}(s)|^3 + N |h_{ij}(s)| \right) \sup_{\theta^{ij}} |\hat{\partial}_{ij}^3 F(N^{-1} \text{Tr} \hat{G}(z, \theta^{ij} H_s)) \right] \leq C_1 (\xi N^\gamma)^{C_2} N/q. \quad (4.37)$$

An explicit computation yields

$$\hat{\partial}_{ij}^3 F(N^{-1} \text{Tr} \hat{G}) = (N^{-1} \hat{\partial}_{ij} \text{Tr} \hat{G})^3 F''' + 3(N^{-1} \hat{\partial}_{ij} \text{Tr} \hat{G})(N^{-1} \hat{\partial}_{ij}^2 \text{Tr} \hat{G}) F'' + (N^{-1} \hat{\partial}_{ij}^3 \text{Tr} \hat{G}) F'. \quad (4.38)$$

The derivative of the trace is

$$\hat{\partial}_{ij}^r \text{Tr} \hat{G} = (-1)^r r! \text{Tr} [\hat{G} (X_{ij} \hat{G})^r]. \quad (4.39)$$

The matrix X_{ij} consists of four nonzero entries and so $(\hat{\partial}_{ij})^l \text{Tr} \hat{G}$ is a sum of $r! 4^r$ of terms which are products of resolvent entries. By Proposition 4.4 we have with overwhelming probability for any z with $\text{Im}[z] \geq N^{-1-\gamma}$ that

$$\sup_{\theta^{kl}} |N^{-1} (\hat{\partial}_{ij})^r \text{Tr} \hat{G}(z, \theta^{ij} H_s)| \leq 4^r r! (\xi^3 N^\gamma)^{r+1}, \quad r = 1, 2, 3. \quad (4.40)$$

We denote \mathcal{A} the event that (4.40) holds. On \mathcal{A} we have

$$\sup_{\theta^{ij}} |\hat{\partial}_{ij}^3 F(N^{-1} \text{Tr} \hat{G}(z, \theta^{ij} H_s))| \leq C (\xi^3 N^\gamma)^{6+c_0}. \quad (4.41)$$

On the complement \mathcal{A}^c we have the deterministic bound

$$\sup_{\theta^{kl}} |N^{-1}(\partial_{ij})^r \operatorname{Tr} \hat{G}(z, \theta^{ij} H_s)| \leq r!(4/\operatorname{Im}[z])^r \leq 4^r r! (N^{1+\gamma})^{r+1}, \quad r = 1, 2, 3. \quad (4.42)$$

We deduce (4.37), yielding the lemma. \square

The following lemma is essentially [30, Theorem 6.4]. It transforms the statement about the Green's function of Lemma 4.6 to a statement about the local correlation functions.

Lemma 4.7. *Consider two random Laplacian-type matrices H_0 and H_1 with Green's functions $\hat{G}_0(z)$ and $\hat{G}_1(z)$. Suppose that for all choices of the parameters (i.e., n, γ, δ, k_i 's, z_i 's, etc.) and functions F the estimate (4.35) holds. Then for all smooth observables O*

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}^n} O(\alpha_1, \dots, \alpha_n) \left\{ \rho_{H_0}^{(n)} \left(E + \frac{\alpha_1}{N}, \dots, E + \frac{\alpha_n}{N} \right) - \rho_{H_1}^{(n)} \left(E + \frac{\alpha_1}{N}, \dots, E + \frac{\alpha_n}{N} \right) \right\} d\alpha_1 \dots d\alpha_n = 0. \quad (4.43)$$

We have therefore proven (4.20) of Proposition 4.3.

4.1.2. Comparison of gap statistics

In this subsection, we prove (4.21) of Proposition 4.3, the comparison of the gap statistics for random Laplacian-type matrices, which is extremely similar to the argument given in detail in [32, Sections 4, 5]. The key difference is that we do not have a delocalization estimate for all N eigenvectors but only for the eigenvectors corresponding to eigenvalues belonging to the interval $[-L, L]$.

The first step is to obtain the following level repulsion estimate for random Laplacian-type matrices.

Proposition 4.8. *Let H be a random Laplacian-type matrix. Fix a small $\kappa > 0$ and fix an index $i \in \llbracket \kappa N, (1 - \kappa)N \rrbracket$. There is a $\tau_0 > 0$ so that for all fixed $0 < \tau < \tau_0$ we have for large enough N ,*

$$\mathbb{P}(|\lambda_i(H) - \lambda_{i+1}(H)| \leq N^{-1-\tau}) \leq N^{-\tau/2}. \quad (4.44)$$

Proof. The proof is similar to the proof of [32, Theorem 4.1]. We detail only what changes. Denote the matrix

$$R_i(H) := \sum_{j: j \neq i, j \leq N} \frac{1}{\lambda_i(H) - \lambda_j(H)} u_j(H) u_j(H)^* = \oint_{|z - \lambda_i(H)| = \omega} \frac{\hat{G}(z, H)}{\lambda_i(H) - z} dz, \quad (4.45)$$

where ω is chosen such that the contour $|z - \lambda_i(H)| = \omega$ encloses only $\lambda_i(H)$. We define the quantity

$$Q_i(H) := \frac{1}{N^2} \operatorname{Tr}(R_i(H)^2) = \frac{1}{N^2} \sum_{j: j \neq i, j \leq N} \frac{1}{(\lambda_j(H) - \lambda_i(H))^2}. \quad (4.46)$$

This quantity plays an important role in [39,40], where it was observed that it captures quantitatively the derivatives of $\lambda_i(H)$. The function $Q_i(H)$ is not well-defined on the space of real symmetric matrices (it diverges when $\lambda_i(H)$ is not a single eigenvalue) and so we compose it with a cutoff function $\chi = \chi_N$. The function χ is chosen so that: (1) it is smooth with its first three derivatives bounded; (2) on the interval $[0, N^{2\tau}]$ we have that $|\chi(x) - x| \leq 1$, and for $x \geq N^{2\tau}$, $\chi(x) = N^{2\tau}$. The function $\chi(Q_i(H))$ is then a well-defined smooth function on the space of real symmetric matrices.

The proof of Proposition 4.8 consists of three steps. The first step is to obtain the estimate

$$\mathbb{E}[\chi(Q_i(H_t))] \leq CN^{3\tau/2}, \quad (4.47)$$

for some $t = N^{-1+\epsilon}$ with ϵ sufficiently small. This estimate follows from [35, Theorem 3.6] and Lemma 4.2.

In the second step, we compare $\mathbb{E}[\chi(Q_i(H_t))]$ for $t = N^{-1+\epsilon}$ and $\mathbb{E}[\chi(Q_i(H_0))]$ using the continuity Lemma 4.5. By choosing ϵ sufficiently small we claim that

$$|\mathbb{E}[\chi(Q_i(H_t))] - \mathbb{E}[\chi(Q_i(H_0))]| \leq 1. \quad (4.48)$$

Taking $F(H) := \chi(Q_i(H))$ in Lemma 4.5 we see that it suffices to prove that

$$\sup_{\substack{0 \leq s \leq t, \\ 1 \leq a < b \leq N+1}} \mathbb{E} \left[(N^2 |h_{ab}(s)|^3 + N |h_{ab}(s)|) \sup_{\theta^{ab}} |\hat{\partial}_{ab}^3 \chi(Q_i(\theta^{ab} H_s))| \right] \leq C_1 (\xi N^\tau)^{C_2} N/q. \quad (4.49)$$

The proof of this bound is similar to the computation completed in [32, Proposition 4.6]. We outline the differences. The formulas [32, (4.16)–(4.18)] remain valid after replacing V by the X_{ab} . The end result is that $\hat{\partial}_{ab}^3 Q_i(H)$ is a finite sum of terms of the form

$$\begin{aligned} & \frac{1}{N^2} \sum_{j:j \neq i} \frac{\Omega}{(\lambda_i - \lambda_j)^5}, \quad \Omega = V_{ii}^3, V_{ii}^2 V_{jj}, V_{ii} V_{jj}^2, V_{jj}^3. \\ & \frac{1}{N^2} \sum_{\substack{j_1, j_2: \\ j_1, j_2 \neq i}} \frac{\Omega}{(\lambda_i - \lambda_{j_1})^4 (\lambda_i - \lambda_{j_2})}, \quad \Omega = V_{ii} V_{j_1 j_2}^2, V_{j_1 j_1} V_{j_2}^2, V_{j_1 j_1} V_{j_1 j_2}^2, V_{ii} V_{j_2}^2. \\ & \frac{1}{N^2} \sum_{\substack{j_1, j_2: \\ j_1, j_2 \neq i}} \frac{\Omega}{(\lambda_i - \lambda_{j_1})^3 (\lambda_i - \lambda_{j_2})^2}, \quad \Omega = V_{j_2 j_2} V_{j_1 j_2}^2, V_{ii} V_{j_1}^2, V_{ii} V_{j_1 j_2}^2, V_{j_1 j_1} V_{j_1 j_2}^2. \\ & \frac{1}{N^2} \sum_{\substack{j_1, j_2, j_3: \\ j_1, j_2, j_3 \neq i}} \frac{V_{j_1 j_2} V_{j_2 j_3} V_{j_3 j_1}}{(\lambda_i - \lambda_{j_1})^3 (\lambda_i - \lambda_{j_2}) (\lambda_i - \lambda_{j_3})}, \quad \sum_{\substack{j_1, j_2, j_3: \\ j_1, j_2, j_3 \neq i}} \frac{V_{j_1 j_2} V_{j_2 j_3} V_{j_3 j_1}}{(\lambda_i - \lambda_{j_1})^2 (\lambda_i - \lambda_{j_2})^2 (\lambda_i - \lambda_{j_3})}, \end{aligned} \quad (4.50)$$

where $V_{jk} := u_j^* X_{ab} u_k$. The only change in controlling the above quantities arises from the fact that we do not have the delocalization of eigenvectors corresponding to eigenvalues outside the interval $[-L, L]$. However, the contribution for eigenvalues far away from $\lambda_i(\theta^{ab} H_s)$ is negligible due to the fact that the denominator is bounded below. For the sake of exposition, we show how the existing proof of [32, Section 4] is modified to bound the term

$$\frac{1}{N^2} \sum_{j:j \neq i} \frac{V_{jj}^3}{(\lambda_i - \lambda_j)^5}. \quad (4.51)$$

On the event $\{Q_i(\theta^{ab} H_s) \leq N^{2\tau}\}$ we have

$$\sum_{j:j \neq i} \frac{1}{|\lambda_i - \lambda_j|^k} \leq \left(\sum_{j:j \neq i} \frac{1}{|\lambda_i - \lambda_j|^2} \right)^{k/2} \leq N^{k(1+\tau)}, \quad k \geq 2. \quad (4.52)$$

Let \mathcal{A} be the event that the bound (4.22) holds. On the event \mathcal{A} we have that

$$\|u_j\|_\infty^2 \leq \xi^3/N, \quad j: |\lambda_j - \lambda_i| \leq L/2. \quad (4.53)$$

Therefore on the event $\mathcal{A} \cap \{Q_i(\theta^{ab} H_s) \leq N^{2\tau}\}$ we have

$$\begin{aligned} \frac{1}{N^2} \sum_{j:j \neq i} \frac{V_{jj}^3}{|\lambda_i - \lambda_j|^5} & \leq \frac{1}{N^2} \sum_{\substack{j:j \neq i \\ |\lambda_j - \lambda_i| \leq L/2}} \frac{V_{jj}^3}{|\lambda_i - \lambda_j|^5} + \frac{1}{N^2} \sum_{\substack{j:j \neq i \\ |\lambda_j - \lambda_i| \geq L/2}} \frac{V_{jj}^3}{|\lambda_i - \lambda_j|^5} \\ & \leq \frac{\xi^9}{N^5} \sum_{j:j \neq i} \frac{1}{|\lambda_i - \lambda_j|^5} + \frac{1}{N^2} \frac{2}{L} \sum_{j:j \neq i} |u_j^* X_{ab} u_j|^3 \\ & \leq \xi^9 N^{5\tau} + \frac{C}{N^2} \leq C_1 (\xi N^\tau)^{C_2} \end{aligned} \quad (4.54)$$

The other terms of (4.50) are handled similarly. The bound (4.49) is subsequently obtained by a similar argument as in Lemma 4.6. We then choose τ and ξ sufficiently small such that $C_1 (\xi N^\tau)^{C_2} \leq N^{\beta/4}$, where β is from (2.12), and $t = N^{-1+\epsilon}$ with $\epsilon \leq \beta/2$, then we have $t C_1 (\xi N^\tau)^{C_2} N/q \leq N^{-\beta/4}$ and (4.48) follows.

Combining (4.47) and (4.48) yields

$$\mathbb{E}[\chi(Q_i(H_0))] \leq CN^{3\tau/2}. \quad (4.55)$$

The estimate (4.44) follows from Markov's inequality and the definition of χ . \square

Proof of (4.21) of Proposition 4.3. For any test function $O \in C^\infty(\mathbb{R}^n)$ with bounded first three derivatives, it suffices to prove the stronger claim:

$$\lim_{N \rightarrow \infty} \left| \mathbb{E}[O(N\lambda_i(0), \dots, N\lambda_{i+n}(0))] - \mathbb{E}[O(N\lambda_i(t), \dots, N\lambda_{i+n}(t))] \right| = 0 \quad (4.56)$$

for $t \leq N^{-1+\epsilon}$. For simplicity of notation we only prove (4.56) in the case $n = 1$:

$$\lim_{N \rightarrow \infty} \left| \mathbb{E}[O(N\lambda_i(t))] - \mathbb{E}[O(N\lambda_i(0))] \right| = 0. \quad (4.57)$$

The general case is analogous.

Take a cutoff function ρ such that $\rho(x) = 1$ for $x \leq N^{2\tau}$ and $\rho(x) = 0$ for $x \geq 2N^{2\tau}$, where $\tau > 0$ is a small constant. We can choose ρ so that its first three derivatives are bounded. We have by Proposition 4.8 that

$$\mathbb{P}(Q_i(H_s) \geq N^{2\tau}) \leq N^{-\tau/2}, \quad s = 0, t. \quad (4.58)$$

Since $O \in C^\infty(\mathbb{R})$ is bounded, we have

$$\begin{aligned} & \left| \mathbb{E}[O(N\lambda_i(0))] - \mathbb{E}[O(N\lambda_i(t))] \right| \\ & \leq \left| \mathbb{E}[O(N\lambda_i(0))\rho_M(Q_i(H_0))] - \mathbb{E}[O(N\lambda_i(t))\rho_M(Q_i(H_t))] \right| \\ & \quad + \|O\|_\infty \mathbb{P}(Q_i(H_0) \geq N^{2\tau}) + \|O\|_\infty (\mathbb{P}(Q_i(H_t) \geq N^{2\tau})) \\ & \leq \left| \mathbb{E}[O(N\lambda_i(0))\rho_M(Q_i(H_0))] - \mathbb{E}[O(N\lambda_i(t))\rho_M(Q_i(H_t))] \right| + \frac{2\|O\|_\infty}{N^{\tau/2}}. \end{aligned} \quad (4.59)$$

Notice that $O(N\lambda_i(H))\rho_M(Q_i(H))$ is a well-defined smooth function on the space of symmetric matrices. We apply Lemma 4.5 with the function $F(H) = O(N\lambda_i(H))\rho_M(Q_i(H))$. An argument similar to the derivation of (4.49) yields

$$\sup_{\theta^{ab}} \left| \hat{\partial}_{ab}^3 O(N\lambda_i(H))\rho_M(Q_i(H)) \right| \leq C_1 (\xi N^\tau)^{C_2}, \quad (4.60)$$

with overwhelming probability. Thus by Lemma 4.5,

$$\left| \mathbb{E}[O(N\lambda_i(0))\rho_M(Q_i(H_0))] - \mathbb{E}[O(N\lambda_i(t))\rho_M(Q_i(H_t))] \right| \leq C_1 (\xi N^\tau)^{C_2} tN/q. \quad (4.61)$$

The equation (4.57) follows if we choose τ and ξ sufficiently small such that $C_1 (\xi N^\tau)^{C_2} \leq N^{\beta/4}$, where β is from (2.12), and $t = N^{-1+\epsilon}$ with $\epsilon \leq \beta/2$, then we have $tC_1 (\xi N^\tau)^{C_2} N/q \leq e^{-\beta/4}$. And (4.57) follows. \square

We have therefore proven (4.21) of Proposition 4.3. This finishes the proof of Proposition 4.3.

4.2. Universality of H_t

The goal of this section is the following proposition.

Proposition 4.9. *Let H be a random Laplacian-type matrix. We define H_t as in (4.2). Let $t = N^{-1+\epsilon}$, for sufficiently small $\epsilon \leq \beta/2$, where β is from (2.12). There exists a constant c so that the following holds. Firstly, we have gap universality for H_t . For any $\kappa > 0$ and index $i \in \llbracket \kappa N, (1 - \kappa)N \rrbracket$ we have*

$$\begin{aligned} & \left| \mathbb{E}^{(H_t)} \left[O(N\rho_{\text{fc}}(\gamma_i)(\lambda_i - \lambda_{i+1}), \dots, N\rho_{\text{fc}}(\gamma_i)(\lambda_i - \lambda_{i+n})) \right] \right. \\ & \quad \left. - \mathbb{E}^{(GOE)} \left[O(N\rho_{\text{sc}}(\gamma_{i,\text{sc}})(\lambda_i - \lambda_{i+1}), \dots, N\rho_{\text{sc}}(\gamma_{i,\text{sc}})(\lambda_i - \lambda_{i+n})) \right] \right| \leq CN^{-c}, \end{aligned} \quad (4.62)$$

for any test function $O \in C_c^\infty(\mathbb{R}^n)$.

Secondly, the averaged n -point correlation functions of the nontrivial eigenvalues of H_t are universal. For any $\delta > 0$ and $E \in [-L, L]$, and $b \geq N^{-1+\delta}$ we have

$$\left| \int_{E-b}^{E+b} \int_{\mathbb{R}^n} O(\alpha_1, \dots, \alpha_n) \left\{ \frac{1}{\rho_{\text{fc}}(E)^n} \rho_{H_t}^{(n)} \left(E' + \frac{\alpha_1}{N\rho_{\text{fc}}(E)}, \dots, E' + \frac{\alpha_n}{N\rho_{\text{fc}}(E)} \right) - \frac{1}{\rho_{\text{sc}}(E)^n} \rho_{\text{GOE}}^{(n)} \left(E' + \frac{\alpha_1}{N\rho_{\text{sc}}(E)}, \dots, E' + \frac{\alpha_n}{N\rho_{\text{sc}}(E)} \right) \right\} d\alpha_1 \dots d\alpha_n \frac{dE'}{2b} \right| \leq N^{-c}, \quad (4.63)$$

where the test function $O \in C_c^\infty(\mathbb{R}^n)$.

The proof of Proposition 4.9 is a modification of [32, Section 3]. We do not provide all of the details.

We first require a minor modification of the main result of [35, Theorem 2.5], which states that for any deterministic matrix A the local eigenvalue statistics of $A + \sqrt{t}GOE$ are universal for $t \gtrsim N^{-1}$, under some mild regularity conditions of its eigenvalues. To state it more precisely, we introduce some notation. For the deterministic $N \times N$ matrix A we define

$$A_t := A + \vartheta_t G, \quad \vartheta_t := \sqrt{\frac{1-e^{-t}}{2}} = O(t^{1/2}), \quad (4.64)$$

where G is a standard GOE matrix. We denote by ρ_t the free convolution of the empirical eigenvalue distribution of A and the semicircle law with variance ϑ_t^2 , and m_t the Stieltjes transform of ρ_t . Then $m_0(z)$ is just the Stieltjes transform of empirical eigenvalue distribution of A . The density ρ_t is analytic on its support for any $t > 0$. The function m_t solves the equation

$$m_t(z) = m_0(z + \vartheta_t^2 m_t(z)) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\lambda_i(A) - z - \vartheta_t^2 m_t(z)}, \quad (4.65)$$

where $\lambda_1(A) \leq \lambda_2(A) \leq \dots \leq \lambda_N(A)$ are the eigenvalues of A . We denote the classical eigenvalues of ρ_{fc} , ρ_{sc} and ρ_t by γ_i , $\gamma_{i,\text{sc}}$ and $\gamma_{i,t}$ respectively. They are defined by

$$\int_{-\infty}^{\gamma_i} \rho_{\text{fc}}(x) dx = \frac{i - \frac{1}{2}}{N}, \quad \int_{-\infty}^{\gamma_{i,\text{sc}}} \rho_{\text{sc}}(x) dx = \frac{i - \frac{1}{2}}{N}, \quad \int_{-\infty}^{\gamma_{i,t}} \rho_t(x) dx = \frac{i - \frac{1}{2}}{N}. \quad (4.66)$$

The following is a minor modification of [35, Theorem 2.5].

Theorem 4.10. Fix small parameter $\epsilon > 0$ and large parameter $L > 0$. Suppose that there are constants $C_1, C_2 > 0$ such that

$$C_1 \leq \text{Im}[m_0(E + i\eta)] \leq C_2 \quad (4.67)$$

for all $E \in (-L, L)$ and $N^{-1+\epsilon/3} \leq \eta \leq L$. There is a constant c_ϵ so that the following holds. Suppose that $\|A\| \leq N^{c_\epsilon}$. Let i be such that $\lambda_i(A) \in (-L/2, L/2)$ and index $j \in \llbracket \kappa N, (1-\kappa)N \rrbracket$ for some $\kappa > 0$. Then for $t = N^{-1+\epsilon}$, there exists a small constant $c > 0$, which depends on ϵ and κ , such that for indices $i_1, \dots, i_n \in \mathbb{N}$ and $i_k \leq N^c$,

$$\left| \mathbb{E}^{(A_t)} \left[O(\rho_t(\gamma_{i_1,t})N(\lambda_{i_1} - \lambda_{i_1+i_1}), \dots, \rho_t(\gamma_{i_n,t})N(\lambda_{i_n} - \lambda_{i_n+i_n})) \right] - \mathbb{E}^{(GOE)} \left[O(\rho_{\text{sc}}(\gamma_j)N(\lambda_j - \lambda_{j+i_1}), \dots, \rho_{\text{sc}}(\gamma_j)N(\lambda_j - \lambda_{j+i_n})) \right] \right| \leq CN^{-c}. \quad (4.68)$$

The only difference between the above theorem and that stated in [35, Theorem 2.5] is that the spectrum of the matrix A is required to have a lower bound in [35, Theorem 2.5]. However, the only place the lower bound is used is to apply [35, Lemma 7.17] in order to conclude [35, Theorem 3.5]; that is, so that the error in the rigidity estimates on the eigenvalues of A_t is smaller than N^δ/N for any $\delta > 0$. It is clear that if instead the lower bound on A is replaced by the bound $\|A\| \leq N^{\delta_1}$ then there is a universal $C > 0$ so that the rigidity estimates of [35, Theorem 3.5] hold with an error of $N^{C\delta}/N$. The proof of the main result of [35] on universality only requires that the rigidity estimates hold with an error of N^{δ_2}/N where $\delta_2 > 0$ is some small constant. Hence Theorem 4.10 holds.

Let \mathcal{A} be the set of $N \times N$ real symmetric matrices such that

$$\mathcal{A} := \{A : |m_0(z) - m_{\text{fc}}(z)| \leq \xi \Psi, z \in \mathcal{D}_L\} \cap \{\|A\| \leq \xi\}, \quad (4.69)$$

where the $\nu > 0$ in the definition of ξ is taken small enough.

Lemma 4.11. *Let $A \in \mathcal{A}$. We have for $t \leq q/N$, (in fact we only need that $t = o(1)$)*

$$|m_t(z) - m_{\text{fc}}(z)| \leq C(\xi\Psi + t^{1/2}) \quad (4.70)$$

uniformly for

$$\begin{aligned} z \in & \{z = E + i\eta : |E| \leq 2\xi + L - 1, 2 \leq \eta \leq L - 1\} \\ & \cup \{z = E + i\eta : |E| \leq L - 1, \xi^3 N^{-1} \leq \eta \leq 2\} \\ & \cup \{z = E + i\eta : 2\xi + 1 \leq |E| \leq 2\xi + L - 1, \xi^3 N^{-1} \leq \eta \leq 2\}. \end{aligned} \quad (4.71)$$

The proof is identical to [32, Lemma 3.3]. We first derive that $|\vartheta_t^2 m_t| \leq \vartheta_t = O(t^{1/2})$, and then use the definition of the set \mathcal{A} to get

$$m_t(z) = m_0(z + \vartheta_t^2 m_t(z)) = m_{\text{fc}}(z + \vartheta_t^2 m_t(z)) + O(\xi\Psi) = m_{\text{fc}}(z) + O(\xi\Psi + t^{1/2}), \quad (4.72)$$

where we have used that m_{fc} is Lipschitz from Lemma 3.4.

We now need the following analogue of [32, Lemma 3.4].

Lemma 4.12. *For any $t \leq \min\{q/N, N^{-2/3}\}$, any real symmetric matrix $A \in \mathcal{A}$ and $E \in (-L/2, L/2)$ we have*

$$|\rho_t(E) - \rho_{\text{fc}}(E)| \leq C\xi^3/(Nt). \quad (4.73)$$

Moreover for any index i such that $\lambda_i(A) \in (-L/2, L/2)$, we have

$$|\rho_t(\gamma_{i,t}) - \rho_{\text{fc}}(\gamma_i)| \leq C\xi^3/(Nt). \quad (4.74)$$

Proof. The proof is similar to that of [32, Lemma 3.4]. We have the bounds

$$\rho_t(E) \leq C, \quad |\rho'_t(E)| \leq C/t, \quad (4.75)$$

uniformly for $E \in [-L, L]$. For (4.73), we can write the difference as

$$\begin{aligned} |\rho_t(E) - \rho_{\text{fc}}(E)| & \leq |\rho_t(E) - \text{Im}[m_t(E + i\eta)]/\pi| \\ & \quad + |\text{Im}[m_t(E + i\eta)]/\pi - \text{Im}[m_{\text{fc}}(E + i\eta)]/\pi| + |\text{Im}[m_{\text{fc}}(E + i\eta)]/\pi - \rho_{\text{fc}}(E)|. \end{aligned} \quad (4.76)$$

The statement follows by a similar argument as in [32, Lemma 3.4] and taking $\eta = \xi^3/N$.

The proof of (4.74) is analogous to the proof of [32, Lemma 3.4], and the main ingredient is the rigidity estimate

$$|\gamma_{i,t} - \gamma_i| \leq C(\xi^2 q^{-1} + \xi t^{1/2}), \quad (4.77)$$

which follows the same argument as the proof of Corollary 2.6 and the definition of \mathcal{A} . To prove (4.74) we have

$$|\rho_t(\gamma_{i,t}) - \rho_{\text{fc}}(\gamma_i)| \leq |\rho_t(\gamma_{i,t}) - \rho_{\text{fc}}(\gamma_{i,t})| + |\rho_{\text{fc}}(\gamma_{i,t}) - \rho_{\text{fc}}(\gamma_i)| \leq C\xi^3/(Nt), \quad (4.78)$$

given that $t \leq \min\{q/N, N^{-2/3}\}$, where we have used that ρ_{fc} is Lipschitz. \square

Proof of Proposition 4.9. By Lemma 4.2 we need to prove universality for the matrix \hat{H}_t defined there. We have that

$$\hat{H}_t = \hat{A}_t + (1 - e^{-t})^{1/2} GOE. \quad (4.79)$$

By Lemma 4.2 we have that $\hat{A}_t \in \mathcal{A}$ with overwhelming probability given that $t \leq \xi/q$. With this and Lemma 4.12 established, the remainder of the proof is identical to the proof of [32, Theorem 3.1]. \square

Proof of Theorem 2.9. Theorem 2.9 follows from combining Proposition 4.3, which states that the local statistics of eigenvalues of H and H_t are the same for $t = N^{-1+\epsilon}$ with $\epsilon \leq \beta/2$, and Proposition 4.9, which states that the local statistics of eigenvalues of H_t for $t = N^{-1+\epsilon}$ are the same as those of Gaussian orthogonal ensembles. Therefore the eigenvalue statistics of H agree with those of Gaussian orthogonal ensembles. \square

Acknowledgements

Part of the work of B.L. was conducted during a stay at the Erwin Schrödinger Institute. B.L. would like to thank ESI for their hospitality. The authors are grateful to Philippe Sosoe for helpful comments on a draft of this paper. The authors would also like to thank the referee for helpful comments, which contributed to improving the quality of the publication.

References

- [1] O. H. Ajanki, L. Erdős and T. Krüger. Local spectral statistics of Gaussian matrices with correlated entries. *J. Stat. Phys.* **163** (2) (2016) 280–302. MR3478311 <https://doi.org/10.1007/s10955-016-1479-y>
- [2] J. Alt. The local semicircle law for random matrices with a fourfold symmetry. *J. Math. Phys.* **56** (10) (2015) 103301. MR3406427 <https://doi.org/10.1063/1.4932606>
- [3] G. W. Anderson, A. Guionnet and O. Zeitouni. *An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics* **118**. Cambridge University Press, Cambridge, 2010. MR2760897
- [4] R. Bauerschmidt, J. Huang, A. Knowles and H.-T. Yau. Bulk eigenvalue statistics for random regular graphs. *Ann. Probab.* **45** (6A) (2017) 3626–3663. MR3729611 <https://doi.org/10.1214/16-AOP1145>
- [5] R. Bauerschmidt, A. Knowles and H.-T. Yau. Local semicircle law for random regular graphs. *Comm. Pure Appl. Math.* **70** (10) (2017) 1898–1960. MR3688032 <https://doi.org/10.1002/cpa.21709>
- [6] F. Benaych-Georges, A. Guionnet and C. Male. Central limit theorems for linear statistics of heavy tailed random matrices. *Comm. Math. Phys.* **329** (2) (2014) 641–686. MR3210147 <https://doi.org/10.1007/s00220-014-1975-3>
- [7] F. Benaych-Georges and R. R. Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* **227** (1) (2011) 494–521. MR2782201 <https://doi.org/10.1016/j.aim.2011.02.007>
- [8] P. Biane. On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.* **46** (3) (1997) 705–718. MR1488333 <https://doi.org/10.1512/iumj.1997.46.1467>
- [9] A. Bloemendal, L. Erdős, A. Knowles, H.-T. Yau and J. Yin. Isotropic local laws for sample covariance and generalized Wigner matrices. *Electron. J. Probab.* **19** (33) (2014) 1–53. MR3183577 <https://doi.org/10.1214/ejp.v19-3054>
- [10] P. Bourgade, L. Erdős, H.-T. Yau and J. Yin. Fixed energy universality for generalized Wigner matrices. *Comm. Pure Appl. Math.* **69** (10) (2016) 1815–1881. MR3541852 <https://doi.org/10.1002/cpa.21624>
- [11] P. Bourgade and H.-T. Yau. The eigenvector moment flow and local quantum unique ergodicity. *Comm. Math. Phys.* **350** (1) (2017) 231–278. MR3606475 <https://doi.org/10.1007/s00220-016-2627-6>
- [12] W. Bryc, A. Dembo and T. Jiang. Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.* **34** (1) (2006) 1–38. MR2206341 <https://doi.org/10.1214/009117905000000495>
- [13] F. Chung, L. Lu and V. Vu. Eigenvalues of random power law graphs. *Ann. Comb.* **7** (1) (2003) 21–33. MR1984643 <https://doi.org/10.1007/s000260300002>
- [14] F. Chung, L. Lu and V. Vu. The spectra of random graphs with given expected degrees. *Internet Math.* **1** (3) (2004) 257–275. MR2111009
- [15] F. Chung and M. Radcliffe. On the spectra of general random graphs. *Electron. J. Combin.* **18** (1) (2011) 215. MR2853072
- [16] F. R. K. Chung. *Spectral Graph Theory. CBMS Regional Conference Series in Mathematics* **92**. American Mathematical Society, Providence, RI, 1997. Published for the Conference Board of the Mathematical Sciences, Washington, DC. MR1421568
- [17] A. Coja-Oghlan. On the Laplacian eigenvalues of $G_{n,p}$. *Combin. Probab. Comput.* **16** (6) (2007) 923–946. MR2351691 <https://doi.org/10.1017/S0963548307008693>
- [18] X. Ding and T. Jiang. Spectral distributions of adjacency and Laplacian matrices of random graphs. *Ann. Appl. Probab.* **20** (6) (2010) 2086–2117. MR2759729 <https://doi.org/10.1214/10-AAP677>
- [19] I. Dumitriu and S. Pal. Sparse regular random graphs: Spectral density and eigenvectors. *Ann. Probab.* **40** (5) (2012) 2197–2235. MR3025715 <https://doi.org/10.1214/11-AOP673>
- [20] L. Erdős, A. Knowles, H.-T. Yau and J. Yin. Spectral statistics of Erdős–Rényi graphs II: Eigenvalue spacing and the extreme eigenvalues. *Comm. Math. Phys.* **314** (3) (2012) 587–640. MR2964770 <https://doi.org/10.1007/s00220-012-1527-7>
- [21] L. Erdős, A. Knowles, H.-T. Yau and J. Yin. The local semicircle law for a general class of random matrices. *Electron. J. Probab.* **18** (59) (2013) 1–58. MR3068390 <https://doi.org/10.1214/EJP.v18-2473>
- [22] L. Erdős, A. Knowles, H.-T. Yau and J. Yin. Spectral statistics of Erdős–Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** (3B) (2013) 2279–2375. MR3098073 <https://doi.org/10.1214/11-AOP734>
- [23] L. Erdős, S. Péché, J. A. Ramirez, B. Schlein and H.-T. Yau. Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63** (7) (2010) 895–925. MR2662426 <https://doi.org/10.1002/cpa.20317>
- [24] L. Erdős, J. Ramírez, B. Schlein, T. Tao, V. Vu and H.-T. Yau. Bulk universality for Wigner Hermitian matrices with subexponential decay. *Math. Res. Lett.* **17** (4) (2010) 667–674. MR2661171 <https://doi.org/10.4310/MRL.2010.v17.n4.a7>
- [25] L. Erdős, B. Schlein and H.-T. Yau. Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* **287** (2) (2009) 641–655. MR2481753 <https://doi.org/10.1007/s00220-008-0636-9>
- [26] L. Erdős, B. Schlein and H.-T. Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.* **37** (3) (2009) 815–852. MR2537522 <https://doi.org/10.1214/08-AOP421>
- [27] L. Erdős, B. Schlein and H.-T. Yau. Universality of random matrices and local relaxation flow. *Invent. Math.* **185** (1) (2011) 75–119. MR2810797 <https://doi.org/10.1007/s00222-010-0302-7>
- [28] L. Erdős and K. Schnelli. Universality for random matrix flows with time-dependent density. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** (4) (2017) 1606–1656. MR3729630 <https://doi.org/10.1214/16-AIHP765>
- [29] L. Erdős and H.-T. Yau. Gap universality of generalized Wigner and β -ensembles. *J. Eur. Math. Soc. (JEMS)* **17** (8) (2015) 1927–2036. MR3372074 <https://doi.org/10.4171/JEMS/548>
- [30] L. Erdős, H.-T. Yau and J. Yin. Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields* **154** (1–2) (2012) 341–407. MR2981427 <https://doi.org/10.1007/s00440-011-0390-3>

- [31] L. Erdős, H.-T. Yau and J. Yin. Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math.* **229** (3) (2012) 1435–1515. MR2871147 <https://doi.org/10.1016/j.aim.2011.12.010>
- [32] J. Huang, B. Landon and H.-T. Yau Bulk universality for sparse random matrices. Preprint, 2015. Available at arXiv:1504.05170. MR3429490 <https://doi.org/10.1063/1.4936139>
- [33] T. Jiang. Empirical distributions of Laplacian matrices of large dilute random graphs. *Random Matrices Theory Appl.* **1** (3) (2012) 1250004. MR2967963 <https://doi.org/10.1142/S2010326312500049>
- [34] T. Jiang. Low eigenvalues of Laplacian matrices of large random graphs. *Probab. Theory Related Fields* **153** (3–4) (2012) 671–690. MR2948689 <https://doi.org/10.1007/s00440-011-0357-4>
- [35] B. Landon and H.-T. Yau. Convergence of local statistics of Dyson Brownian motion. *Comm. Math. Phys.* **355** (3) (2017) 949–1000. MR3687212 <https://doi.org/10.1007/s00220-017-2955-1>
- [36] J. O. Lee and K. Schnelli. Local deformed semicircle law and complete delocalization for Wigner matrices with random potential. *J. Math. Phys.* **54** (10) (2013) 103504. MR3134604 <https://doi.org/10.1063/1.4823718>
- [37] J. O. Lee, K. Schnelli, B. Stetler and H.-T. Yau. Bulk universality for deformed Wigner matrices. *Ann. Probab.* **44** (3) (2016) 2349–2425. MR3502606 <https://doi.org/10.1214/15-AOP1023>
- [38] L. Lu and X. Peng. Spectra of edge-independent random graphs. *Electron. J. Combin.* **20** (4) (2013) 27. MR3158266
- [39] T. Tao and V. Vu. Random matrices: Universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.* **298** (2) (2010) 549–572. MR2669449 <https://doi.org/10.1007/s00220-010-1044-5>
- [40] T. Tao and V. Vu. Random matrices: Universality of local eigenvalue statistics. *Acta Math.* **206** (1) (2011) 127–204. MR2784665 <https://doi.org/10.1007/s11511-011-0061-3>
- [41] E. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.* **62** (1955) 548–564. MR0077805 <https://doi.org/10.2307/1970079>