Geom. Funct. Anal. Vol. 30 (2020) 693–769
https://doi.org/10.1007/s00039-020-00538-0
Published online July 17, 2020

GAFA Geometric And Functional Analysis

EDGE RIGIDITY AND UNIVERSALITY OF RANDOM REGULAR GRAPHS OF INTERMEDIATE DEGREE

Roland Bauerschmidt, Jiaoyang Huang, Antti Knowles and Horng-Tzer Yau

Abstract. For random d-regular graphs on N vertices with $1 \ll d \ll N^{2/3}$, we develop a $d^{-1/2}$ expansion of the local eigenvalue distribution about the Kesten–McKay law up to order d^{-3}. This result is valid up to the edge of the spectrum. It implies that the eigenvalues of such random regular graphs are more rigid than those of Erdős–Rényi graphs of the same average degree. As a first application, for $1 \ll d \ll N^{2/3}$, we show that all nontrivial eigenvalues of the adjacency matrix are with very high probability bounded in absolute value by $(2 + \mathrm{o}(1))\sqrt{d-1}$. As a second application, for $N^{2/9} \ll d \ll N^{1/3}$, we prove that the extremal eigenvalues are concentrated at scale $N^{-2/3}$ and their fluctuations are governed by Tracy–Widom statistics. Thus, in the same regime of d, 52% of all d-regular graphs have second-largest eigenvalue strictly less than $2\sqrt{d-1}$.

1 Introduction

1.1 Main results. Let $\mathbb{P}$ be the uniform probability measure on the set of d-regular graphs on N vertices. We identify a graph with its adjacency matrix $A = (A_{ij}) \in \{0,1\}^{N\times N}$, defined as $A_{ij} = 1$ if and only if i and j are adjacent. Thus, $\mathbb{P}$ is the uniform probability measure on the set of Hermitian matrices $A \in \{0,1\}^{N\times N}$ satisfying $\sum_{j=1}^{N} A_{ij} = d$ and $A_{ii} = 0$ for all $i = 1, \dots, N$.

Since A is d-regular, it is immediate that A has a trivial eigenvalue d with associated eigenvector $(1, 1, \dots, 1)^*$. Moreover, by the Perron-Frobenius theorem, all other eigenvalues are bounded in absolute value by d. For convenience, we shall consider the normalized adjacency matrix

$$H := (d-1)^{-1/2} A. \tag{1.1}$$

We denote its eigenvalues by $\lambda_1 = d/\sqrt{d-1} \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N \geqslant -d/\sqrt{d-1}$.

Unless stated otherwise, all quantities depend on the fundamental parameter N, and we omit this dependence from our notation. For the following statements, for deterministic N-dependent quantities X and Y we write

$$X \ll Y \quad \text{if} \quad X = \mathrm{O}_{\mathfrak{c}}(N^{-\mathfrak{c}} Y) \text{ for some fixed } \mathfrak{c} > 0. \tag{1.2}$$

(See the conventions in Section 1.3 below.)

Our first main result is about the locations of the nontrivial extremal eigenvalues, λ_2 and λ_N.

Theorem 1.1. *Fix $\mathfrak{c} > 0$. For $1 \ll d \ll N^{2/3}$ and large enough N, with probability $1 - N^{-1/\mathfrak{c}}$ we have*

$$|\lambda_2 - 2|,\ |\lambda_N + 2| \leqslant N^{\mathfrak{c}} \left(\frac{1}{d^3} + \frac{1}{N^{2/3}} + \frac{d^2}{N^{4/3}} \right). \tag{1.3}$$

An immediate consequence is the following optimal upper bound on the nontrivial eigenvalues. It was conjectured for instance in [V08, Conjecture 5.3] and [Vu14, Conjecture 7.3].

COROLLARY 1.2. *For $1 \ll d \ll N^{2/3}$, all nontrivial eigenvalues of the random d-regular graph are with very high probability bounded in absolute value by $(2 + \mathrm{o}(1))\sqrt{d-1}$.*

Our second main result is about the limiting distribution of the extremal eigenvalues.

Theorem 1.3. *For $N^{2/9} \ll d \ll N^{1/3}$ the distribution of $N^{2/3}(\lambda_2 - 2)$ converges to the Tracy–Widom$_1$ distribution, the limiting distribution of the largest eigenvalue of a GOE matrix. The analogous statement holds for $-N^{2/3}(\lambda_N + 2)$.*

Universality for the edge statistics of Wigner matrices (the statement that the distribution of the extremal eigenvalues converge to the Tracy–Widom law) was first established by the moment method [Sos99] under certain symmetry assumptions on the distribution of the matrix elements. The moment method was further developed in [Pec09, FS10] and [Sod10]. A different approach to edge universality for Wigner matrices based on the direct comparison with corresponding Gaussian ensembles was developed in [TV10, EYY12]. Edge universality for sparse Erdős–Rényi graphs was proven first in the regime $pN \gg N^{2/3}$ in the works [EKYY13, EKYY12] and then extended to the regime $pN \gg N^{1/3}$ in [LS18]. For smaller values of the average degree pN, edge universality no longer holds: it was proved in [HLY17, HK20] that, in the regime $1 \ll pN \ll N^{1/3}$, the second-largest eigenvalue has Gaussian fluctuations instead of Tracy–Widom fluctuations. These Gaussian fluctuations result from degree fluctuations which are absent in regular graphs. Our Theorem 1.3 implies that the eigenvalues of random regular graphs are indeed more rigid than those of Erdős–Rényi graphs of the same average degree. For random regular graphs, it is expected that (1.3) is not optimal for small d, and in fact it is conjectured that the extremal eigenvalues continue to have Tracy–Widom fluctuations down to degree $d \geqslant 3$.

As emphasized in [Sar04], the Tracy–Widom$_1$ distribution has positive measure on the set $\{x : x < 0\}$; in fact it has about 52% of its mass on negative values. Therefore Theorem 1.3 implies the existence of many d-regular graphs whose second eigenvalue is *less than* $2\sqrt{d-1}$, provided that N and d obey the conditions of Theorem 1.3.

COROLLARY 1.4. *For d large enough and $d^3 \ll N \ll d^{9/2}$, 52% of d-regular graphs on N vertices have second-largest eigenvalue bounded by $2\sqrt{d-1}$. An analogous statement holds for the smallest eigenvalue.*

In [LPS88], a d-regular graph whose largest nontrivial eigenvalue is bounded in absolute value by $2\sqrt{d-1}$ is called a Ramanujan graph. Corollary 1.4 states that for d large enough and $d^3 \ll N \ll d^{9/2}$, precisely 52% of d-regular graphs have largest (respectively smallest) nontrivial eigenvalue bounded from above by $2\sqrt{d-1}$ (respectively from below by $-2\sqrt{d-1}$). Hence, at least 4% have all nontrivial eigenvalues bounded by $2\sqrt{d-1}$ in absolute value. In fact, Theorem 1.3 and its proof can be extended to show that in the regime $N^{2/9} \ll d \ll N^{1/3}$ the largest and smallest nontrivial eigenvalues converge in distribution to *independent* Tracy–Widom$_1$ distributions; see Remark 9.9 and Theorem 9.10 below. As a consequence, we have the following result.

COROLLARY 1.5. *For d large enough and $d^3 \ll N \ll d^{9/2}$, 27% of d-regular graphs on N vertices have all nontrivial eigenvalues bounded in absolute value by $2\sqrt{d-1}$.*

The conjecture [Sar04, MN08] that a positive fraction of regular graphs of fixed d is Ramanujan remains open. Explicit constructions of Ramanujan graphs with $d = p + 1$ for some prime and prime powers p were introduced in [LPS88, Mar88] (see also [Sar04]); a construction that applies in the bipartite case for all degrees is given in [MSS15, MSS18] and see [Coh16] for polynomial time algorithm of this construction.

The following results on the extremal eigenvalues of random regular graphs are known. For fixed degree $d \geqslant 3$, Friedman [Fri08] proved that $|\lambda_2 - 2| + |\lambda_N + 2| = o(1)$ with high probability. This result was recently reproved using an alternative method in [Bor15]; see also [Pud15]. For d tending to infinity with the number of vertices, it was proved in [FKS89, DJPP13] that the nontrivial extremal eigenvalues are $O(\sqrt{d})$ for the permutation model of random regular graphs. Recently, in [CGJ18, TY19] the bound $O(\sqrt{d})$ was established with high probability for the uniform model of random regular graphs for all $1 \ll d \leqslant N/2$. Previous results in this direction include [KSVW01].

1.2 Related results. In random matrix theory, the bulk spectral statistics of Wigner matrices are well understood; see in particular [Joh01, ESY11, ERSY10, EPRSY10, EYY12, TV11, EY15, BEYY16, EKYY12, EY17]. For Erdős–Rényi random graphs and random regular graphs with growing average degrees, the bulk spectral statistics were analysed in [EKYY13, EKYY12, HLY15, HL15, BHKY17], and complete eigenvector delocalization for logarithmically growing average degree was proved in [EKYY12, HKM19, BKY17]. In the same regime, edge rigidity of Erdős–Rényi graphs was proved in [BBK17, ADK19]. Similar results have also been proved for more general degree distributions [AC15]. These types of results are false for Erdős–Rényi graphs with bounded average degree, whereas random regular graphs

are expected to have random matrix statistics even for bounded degree graphs; see [BHY19] for the proof of complete eigenvector delocalization in this regime. For a review of other results for discrete random matrices, see also [V08].

Macroscopic eigenvalue statistics for random regular graphs of fixed degree have been studied using the techniques of Poisson approximation of short cycles [DJPP13, Joh15] and (non-rigorously) using the replica method [MPL14]. These results show that the macroscopic eigenvalue statistics for random regular graphs of fixed degree are different from those of a Gaussian matrix. However, this is not predicted to be the case for the *local* eigenvalue statistics. Spectral properties of regular directed graphs have also been studied recently [Cok15, Coo15].

For the eigenvectors of random regular graphs with $d \in [N^{\mathfrak{c}}, N^{2/3-\mathfrak{c}}]$, the asymptotic normality was proved in [BHY17]; see also the prior results for Wigner matrices [KY13, TV12, BY16]. For random regular graphs of fixed degree, a Gaussian wave correlation structure for the eigenvectors was predicted in [Elo09] and partially confirmed in [BS19].

In the non-Hermitian setting, the limit of the empirical eigenvalue distribution of random matrices with i.i.d. entries is governed by the circular law. The circular law for non-Hermitian random matrices with i.i.d. entries with certain moment conditions was verified in [Ede97, PZ10, GT10], and the paper [TV10] established the circular law under the weakest moment condition. In the directed d-regular graph setting, it was conjectured [BC12] that for any fixed degree d, the empirical eigenvalue distribution converges to the oriented Kesten–McKay distribution,

$$\frac{1}{\pi}\frac{d^2(d-1)}{(d^2-(x^2+y^2))^2}\mathbf{1}_{|z|\leqslant\sqrt{d}}\,\mathrm{d}x\,\mathrm{d}y.$$

Up to rescaling by $\sqrt{d}$, this measure tends to the circular law as d tends to infinity. In the regime that d grows with the size of the graph, the circular law for the directed d-regular was established in [Coo17, BCZ18, LLTTY18].

1.3 Notation and structure of paper. We use the convention $\mathbb{N} = \{0, 1, 2, \dots\}$. We usually omit the argument N from our notation, with the convention that most quantities are allowed to depend on N and all of our estimates are uniform in them. Estimates are not uniform in quantities that are explicitly constant or fixed. By definition, a random variable F is a function $F(A)$ of the adjacency matrix A. We shall often omit the explicit argument A from our notation, and simply write F for a random variable evaluated at A. For $n \in \mathbb{N}$ we use the notation $[\![n]\!] := \{1, 2, \dots, n\}$. We use the usual big O notation $\mathrm{O}(\cdot)$, and if the implicit constant depends on a parameter α we indicate it by writing $\mathrm{O}_\alpha(\cdot)$. We use the letter $\mathfrak{c}$ to denote a generic small positive constant. For N-dependent random variables X and $Y \geqslant 0$ we write

$$X \prec Y \quad \text{if} \quad \mathbb{P}[|X| > N^{\mathfrak{c}}Y] = \mathrm{O}_{\mathfrak{c}}(N^{-1/\mathfrak{c}}) \text{ for all } \mathfrak{c} > 0. \tag{1.4}$$

If $X \prec Y$ then we also write $X = \mathrm{O}_\prec(Y)$. If the the implicit constant in $\mathrm{O}_{\mathfrak{c}}$ is the same for a family of random variables, we say that the $\prec$ is uniform in that family.

All of our uses of $\prec$ will be uniform in the matrix indices and the spectral parameter of the Green's function.

The rest of the paper is devoted to the proofs of Theorems 1.1 and 1.3 . In Section 2 we introduce the Green's function, which is the main tool in the proof of Theorem 1.1. In Section 3 we discuss switchings of graphs, which are the key operations on graphs that we use to generate self-consistent equations. In Section 4, we introduce a family of polynomials of the Green's functions entries that underlies our proof of Theorem 1.1, and derive some basic estimates on them. In Section 5 we derive the self-consistent equation for the Green's function in expectation. In Section 6 we relate the self-consistent equation derived in the previous section with the Kesten–McKay law. In Section 7, we upgrade the self-consistent equation from Section 5 to an equation in high probability. In Section 8 we use the self-consistent equation from Section 7 to derive a local law around the spectral edges, and as a consequence deduce Theorem 1.1. Finally, in Section 9 we use the rigidity estimates of Theorem 9 to conclude edge universality and Theorem 1.3.

2 Green's Function

We consider the adjacency matrix A restricted to the subspace orthogonal to the vector $\mathbf{1} := (1, \dots, 1)^*$. More precisely, let $P_\perp : \mathbb{R}^N \to \mathbb{R}^N$ be the orthogonal projection onto $\mathbf{1}^\perp$, explicitly given by $P_\perp = I - \mathbf{1}\mathbf{1}^*/N$ where I is the $N \times N$ identity matrix. Since H is the normalized adjacency matrix of a regular graph, the matrices H and $P_\perp$ commute: $HP_\perp = P_\perp H$.

For a spectral parameter $z \in \mathbb{C}_+ := \{z \in \mathbb{C} : \mathrm{Im}[z] > 0\}$ we define the *Green's function* by

$$G(z) := P_\perp (H - z)^{-1} P_\perp. \tag{2.1}$$

Thus, G and $(H - z)^{-1}$ agree on the image of $P_\perp$, which is the subspace of $\mathbb{R}^N$ perpendicular to $\mathbf{1}$. The Green's function satisfies the relation

$$G(z)H = HG(z) = zG(z) + P_\perp = zG(z) + (I - \mathbf{1}\mathbf{1}^*/N). \tag{2.2}$$

Moreover,

$$\sum_i G_{ij}(z) = \sum_j G_{ij}(z) = 0. \tag{2.3}$$

Here, and throughout the following, sums over indices run over $i \in [\![N]\!]$. We denote the normalized trace of G, which is also the Stieltjes transform of the empirical spectral measure of $H|_{\mathbf{1}^\perp}$, by

$$m(z) := \frac{1}{N} \sum_i G_{ii}(z). \tag{2.4}$$

We refer to $m(z)$ simply as the *Stieltjes transform*. Our goal is to approximate $m(z)$ by $m_d(z)$, the Stieltjes transform of the Kesten–McKay law [Kes59],

$$m_d(z) := \int_{\mathbb{R}} \frac{\rho_d(x)}{x-z}\mathrm{d}x, \qquad \rho_d(x) := \left(1 + \frac{1}{d-1} - \frac{x^2}{d}\right)^{-1} \frac{\sqrt{[4-x^2]_+}}{2\pi}.$$

The Kesten–McKay law ρ_d is the spectral measure at any vertex of the infinite d-regular tree (see, for example, [AW15] or [BHY19, Section 5]). It has support $[-2,2]$ in our normalization. The Stieltjes transform $m_d(z)$ is explicitly given by

$$m_d(z) = -\left(z + \frac{d}{d-1} m_{\mathrm{sc}}(z)\right)^{-1}, \tag{2.5}$$

where $m_{\mathrm{sc}}(z)$ is the Stieltjes transform of the Wigner semicircle law,

$$m_{\mathrm{sc}}(z) := \int_{\mathbb{R}} \frac{\rho_{\mathrm{sc}}(x)}{x-z}\mathrm{d}x, \qquad \rho_{\mathrm{sc}}(x) := \frac{\sqrt{[4-x^2]_+}}{2\pi}, \tag{2.6}$$

satisfying the self-consistent equation

$$1 + z m_{\mathrm{sc}}(z) + m_{\mathrm{sc}}^2(z) = 0. \tag{2.7}$$

Later we shall use that, alternatively, $m_d(z)$ can be characterized by the self-consistent equation

$$\begin{aligned} P_\infty(z, m_d(z)) = 0, \quad P_\infty(z,w) &= 1 + zw + \frac{d}{d-1} w^2 \\ &+ \sum_{k \geqslant 2} \frac{(-2)^{k-1}(2k-3)!!}{k!} \frac{d}{(d-1)^k} w^{2k}. \end{aligned} \tag{2.8}$$

Indeed, from (2.5) and (2.7) we get

$$\frac{1}{m_d(z)} = \frac{1}{m_{\mathrm{sc}}(z)} - \frac{m_{\mathrm{sc}}(z)}{d-1}, \tag{2.9}$$

from which we obtain, for large enough d,

$$\begin{aligned} m_{\mathrm{sc}}(z) = \frac{d-1}{2m_d(z)} \left(\sqrt{1 + 4\frac{m_d(z)^2}{d-1}} - 1\right) &= m_d(z) \\ &+ \frac{d-1}{2m_d(z)} \sum_{k \geqslant 2} (-1)^{k-1} \frac{2^k (2k-3)!!}{k!} \frac{m_d(z)^{2k}}{(d-1)^k}. \end{aligned}$$

Plugging this into $0 = 1 + z m_d(z) + \frac{d}{d-1} m_d(z) m_{\mathrm{sc}}(z)$, as follows from (2.5), yields (2.8).

We fix a large $\mathfrak{K} > 0$ and define the spectral domain

$$\mathbf{D} := \{z = E + \mathrm{i}\eta : -\mathfrak{K} \leqslant E \leqslant \mathfrak{K}, N^{-1+1/\mathfrak{K}} \leqslant \eta \leqslant \mathfrak{K}\}. \tag{2.10}$$

Here, and throughout the following, we use the notation

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

for the real and imaginary parts of z. The local semicircle law for random regular graphs [BKY17] shows that $m(z)$ is approximated by $m_d(z)$ to order $1/\sqrt{d}$ (up to logarithmic corrections), at least away from the edges ± 2 of the spectral measure. The next result follows from [BKY17, Theorem 1.1]. In fact, [BKY17, Theorem 1.1] gives a much better estimate for Λ_{d} for z away from the edge ± 2.

Proposition 2.1 ([BKY17, Theorem 1.1]). *With the deterministic control parameters*

$$\Lambda_{\mathrm{o}}(z) := \frac{1}{\sqrt{N\eta}} + \frac{1}{\sqrt{d}} + \frac{d^{3/2}}{N}, \quad \Lambda_{\mathrm{d}}(z) := \left(\frac{1}{\sqrt{N\eta}} + \frac{1}{\sqrt{d}} + \frac{d^{3/2}}{N} \right)^{1/2}, \tag{2.11}$$

we have, for $1 \ll d \ll N^{2/3}$ *and all* $z \in \mathbf{D}$,

$$\max_{i \neq j} |G_{ij}(z)| \prec \Lambda_{\mathrm{o}}(z), \quad \max_{i} |G_{ii}(z) - m_d(z)| \prec \Lambda_{\mathrm{d}}(z). \tag{2.12}$$

One by-product of our proof is an improved estimate for the Green's function entries close to the edges ± 2. The bound $1/\sqrt{d}$ is the best one can expect, since if $A_{ij} = 1$, the off-diagonal Green's function entry G_{ij} is of order $1/\sqrt{d}$. In Proposition 8.4 below, we show that near the spectral edges the estimate (2.12) in fact holds with the smaller control parameters

$$\Lambda_{\mathrm{o}}(z) = \Lambda_{\mathrm{d}}(z) = \frac{1}{\sqrt{N\eta}} + \frac{1}{\sqrt{d}} + \frac{d^{3/2}}{N}. \tag{2.13}$$

Averaging over the index i, the estimate (2.12) implies an estimate on the Stieltjes transform $m(z)$. Using additional cancelations from this average, in this paper we shall derive a more precise estimate (see Theorem 8.1) from which we obtain our results about the extremal eigenvalues.

Throughout this paper, we consistently omit the spectral parameter z from our notation in quantities such as G and m, unless it is needed to avoid confusion.

3 Switchings and Exchangeability

Our analysis makes use of switchings for regular graphs and also makes some use of the invariance under the permutation of vertices. We use ideas related to those introduced in [BHKY17], to which we also refer for references to other uses of switchings.

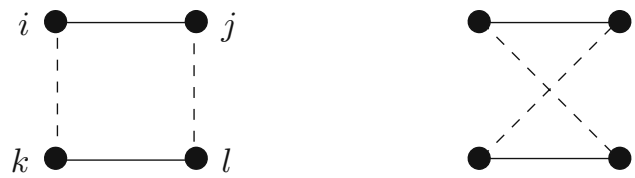

Figure 1: A simple switching is given by replacing the solid edges by the dashed edges.

3.1 Switchings. As in [BHKY17], we define the signed adjacency matrices

$$(\Delta_{ij})_{ab} := \delta_{ia}\delta_{jb} + \delta_{ib}\delta_{ja}, \qquad \xi_{ij}^{kl} := \Delta_{ij} + \Delta_{kl} - \Delta_{ik} - \Delta_{jl}, \tag{3.1}$$

corresponding to an edge at ij respectively to a switching of the edges ij and kl; see Figure 1. Clearly we have $\xi_{ij}^{kl}\mathbf{1} = 0$ and $\xi_{ik}^{jl} = -\xi_{ij}^{kl}$. For any indices i, j, k, l, we denote the indicator function that the edges ij and kl are *switchable* (i.e. the edges ij and kl are present and the switching again results in a simple regular graph) by

$$\chi_{ij}^{kl}(A) = A_{ij}A_{kl}(1 - A_{ik})(1 - A_{jl}). \tag{3.2}$$

In this section, we use switchings to estimate terms of the form

$$\mathbb{E}\left[\prod_{a=1}^{b} A_{i_a j_a} F(A)\right], \tag{3.3}$$

where F is any function which depends on the random graph A, and possibly on the indices $i_1, j_1, \cdots, i_b, j_b$. (Later, we shall take F to be a polynomial of the Green's function entries $\{G_{ij}\}_{i,j\in[\![N]\!]}$ and the Stieltjes transform m.)

PROPOSITION 3.1. *If the indices $i_1, j_1, k_1, l_1 \cdots, i_b, j_b, k_b, l_b$ are distinct, we have the identity*

$$\mathbb{E}\left[F(A)\prod_{a=1}^{b}\chi_{i_a j_a}^{k_a l_a}(A)\right] = \mathbb{E}\left[F\left(A + \sum_{a=1}^{b}\xi_{i_a j_a}^{k_a l_a}\right)\prod_{a=1}^{b}\chi_{i_a k_a}^{j_a l_a}(A)\right], \tag{3.4}$$

where the indicator function χ is as defined in (3.2).

Proof. Define the sets of graphs

$$\begin{aligned}\mathcal{G}_1 &= \left\{A : \prod_{a=1}^{b}\chi_{i_a j_a}^{k_a l_a}(A) = 1\right\},\\ \mathcal{G}_2 &= \left\{A : \prod_{a=1}^{b}\chi_{i_a k_a}^{j_a l_a}(A) = 1\right\}.\end{aligned} \tag{3.5}$$

By our assumption that the indices $i_1, j_1, k_1, l_1 \cdots, i_b, j_b, k_b, l_b$ are distinct, there is a simple bijection between $\mathcal{G}_1$ and $\mathcal{G}_2$, namely

$$A \in \mathcal{G}_2 \mapsto A + \sum_{a=1}^{b} \xi_{i_a j_a}^{k_a l_a} \in \mathcal{G}_1. \tag{3.6}$$

Since $\mathbb{P}$ is the uniform probability measure on d-regular graphs, the claim follows from (3.6). □

In the switching in (3.4), the indicator function $\chi_{ij}^{kl}(A)$ enforces that the matrix $A + \xi_{ij}^{kl}$ is again the adjacency matrix of a simple graph. Note that without this indicator function, such as in the following corollary and elsewhere throughtout our proof, $A + \xi_{ij}^{kl}$ is not necessarily the adjacency matrix of a simple graph, just a real symmetric matrix. This does however not affect our arguments, which should be viewed as operating with general symmetric matrices instead of adjacency matrices of simple graphs.

COROLLARY 3.2. *Fix indices $i, j \in [\![N]\!]$ and $i \neq j$. Let $F \equiv F_{ij}$ be a random variable possibly depending on ij. With the random control parameter*

$$\mathcal{C}_{ij}(F, A) := |F(A)| + \max_{kl} |F(A + \xi_{ij}^{kl})|$$

we have the integration by parts formula

$$\begin{aligned}\mathbb{E}[A_{ij} F(A)] = \frac{d}{N}\mathbb{E}[F(A)] + \frac{1}{Nd}\sum_{kl} \mathbb{E}\left[A_{ik}A_{jl}(F(A + \xi_{ij}^{kl}) - F(A))\right] \\ + \mathrm{O}\left(\frac{d\mathbb{E}[A_{ij}\mathcal{C}_{ij}(F, A)]}{N}\right).\end{aligned} \tag{3.7}$$

Proof. Since $\sum_{kl} A_{kl} = dN$, we have

$$\begin{aligned}&\mathbb{E}[A_{ij}F(A)] \\ &= \frac{1}{Nd}\sum_{kl}\mathbb{E}[A_{ij}A_{kl}F(A)] = \frac{1}{Nd}\sum_{\substack{kl:ijkl\\ \text{distinct}}}\mathbb{E}[\chi_{ij}^{kl}(A)F(A)] \\ &\quad + \mathrm{O}\left(\frac{1}{Nd}\sum_{\substack{kl:ijkl\\ \text{not distinct}}}\mathbb{E}[A_{ij}A_{kl}|F(A)|] + \frac{1}{Nd}\sum_{kl}\mathbb{E}[A_{ij}A_{kl}(A_{ik}+A_{jl})|F(A)|]\right) \\ &= \frac{1}{Nd}\sum_{\substack{kl:ijkl\\ \text{distinct}}}\mathbb{E}[\chi_{ij}^{kl}(A)F(A)] + \mathrm{O}\left(\frac{d\mathbb{E}[A_{ij}\mathcal{C}_{ij}(F,A)]}{N}\right),\end{aligned} \tag{3.8}$$

where we used that the row sums and column sums of A are d. By (3.4), the first term on the right-hand side of (3.8) equals

$$\begin{aligned}\frac{1}{Nd}\sum_{\substack{kl:ijkl\\ \text{distinct}}}\mathbb{E}[\chi_{ik}^{jl}(A)F(A+\xi_{ij}^{kl})] &= \frac{1}{Nd}\sum_{kl}\mathbb{E}[A_{ik}A_{jl}F(A+\xi_{ij}^{kl})]\\ &\quad+\mathrm{O}\left(\frac{1}{Nd}\sum_{\substack{kl:ijkl\\ \text{not distinct}}}\mathbb{E}[A_{ik}A_{jl}|F(A+\xi_{ij}^{kl})|]+\frac{1}{Nd}\sum_{kl}\mathbb{E}[A_{ik}A_{jl}(A_{ij}+A_{kl})|F(A+\xi_{ij}^{kl})|]\right)\\ &= \frac{1}{Nd}\sum_{kl}\mathbb{E}[A_{ik}A_{jl}F(A+\xi_{ij}^{kl})]+\mathrm{O}\left(\frac{d\mathbb{E}[A_{ij}\mathcal{C}_{ij}(F,A)]}{N}\right),\end{aligned} \tag{3.9}$$

where we used that $\mathcal{C}_{ij}(F,A)$ is independent of indices k,l and we can sum over them. The claim (3.7) follows from combining (3.8), (3.9) and the fact $\sum_{kl}A_{ik}A_{jl}=d^2$. □

For $b,c\geqslant 0$ and multi-indices $\mathbf{i}\in[\![N]\!]^b$ and $\mathbf{j}\in[\![N]\!]^c$, we denote by $\mathbf{ij}\in[\![N]\!]^{b+c}$ their concatenation. We shall often need the following random control parameter.

Definition 3.3. *For fixed $b,c\geqslant 0$ we denote the b-tuples $\mathbf{i}=(i_1,\dots,i_b),\mathbf{j}=(j_1,\dots,j_b)$, $\mathbf{k}=(k_1,\dots,k_b),\mathbf{l}=(l_1,\dots,l_b)$ and the c-tuple $\mathbf{m}=(m_1,\dots,m_c)$. Let $F=\{F_{\mathbf{ijm}}\}$ be a family of random variables indexed by $\mathbf{ijm}$. Define the random control parameter*

$$\mathcal{C}(F,A):=\max_{\mathbf{m}}\max_{\mathbf{ijkl}}\left(|F_{\mathbf{ijm}}(A)|+\left|F_{\mathbf{ijm}}\left(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}\right)\right|\right). \tag{3.10}$$

Corollary 3.4. *Let $F_{\mathbf{ijm}}$ be as in Definition 3.3. We have the integration by parts formula*

$$\begin{aligned}&\frac{1}{N^{b+c}d^b}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_aj_a}F_{\mathbf{ijm}}(A)\right]\\ &\quad=\frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\mathbf{ijkl}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_ak_a}A_{j_al_a}\left(F_{\mathbf{ijm}}\left(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}\right)-F_{\mathbf{ijm}}(A)\right)\right]\\ &\qquad+\frac{1}{N^{2b+c}}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[F_{\mathbf{ijm}}(A)\right]+\mathrm{O}\left(\frac{d\mathbb{E}[\mathcal{C}(F,A)]}{N}\right).\end{aligned} \tag{3.11}$$

Proof. Since the row and column sums of A equal d, by introducing new indices $k_1, l_1, \cdots, k_b, l_b$, we rewrite the left-hand side of (3.11) as

$$\begin{aligned}
\frac{1}{N^{b+c}d^b}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_aj_a}F_{\mathbf{ijm}}(A)\right] &= \frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\mathbf{ijkl}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_aj_a}A_{k_al_a}F_{\mathbf{ijm}}(A)\right] \\
&= \frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\substack{\mathbf{ijkl}\\ \text{distinct}}}\mathbb{E}\left[\prod_{a=1}^{b}\chi_{i_aj_a}^{k_al_a}(A)F_{\mathbf{ijm}}(A)\right] \\
&\quad + \frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\substack{\mathbf{ijkl}\\ \text{not distinct}}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_aj_a}A_{k_al_a}F_{\mathbf{ijm}}(A)\right] \\
&\quad + \frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\substack{\mathbf{ijkl}\\ \text{distinct}}}\mathbb{E}\left[\left(\prod_{a=1}^{b}A_{i_aj_a}A_{k_al_a}-\prod_{a=1}^{b}\chi_{i_aj_a}^{k_al_a}(A)\right)F_{\mathbf{ijm}}\right].
\end{aligned} \tag{3.12}$$

The second last term of the last right-hand side in (3.12) can be estimated by

$$\begin{aligned}
&\left|\frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\substack{\mathbf{ijkl}\\ \text{not distinct}}}\prod_{a=1}^{b}A_{i_aj_a}A_{k_al_a}F_{\mathbf{ijm}}(A)\right| \\
&\quad\leqslant \frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\substack{\mathbf{ijkl}\\ \text{not distinct}}}\prod_{a=1}^{b}A_{i_aj_a}A_{k_al_a}\mathcal{C}(F,A) \\
&\quad\leqslant \frac{1}{N^{2b}d^{2b}}\sum_{\substack{\mathbf{ijkl}\\ \text{not distinct}}}\prod_{a=1}^{b}A_{i_aj_a}A_{k_al_a}\mathcal{C}(F,A)\leqslant\frac{1}{N}\mathcal{C}(F,A),
\end{aligned} \tag{3.13}$$

where in the last line we used that $\sum_i A_{ij}=\sum_j A_{ij}=d$, that $A_{ii}=0$, and hence that at least one factor $1/N$ remains because of the constraint in the sum that **ijkl** be not distinct. Similarly, using $|A_{ij}A_{kl}-\chi_{ij}^{kl}(A)|=A_{ij}A_{kl}[A_{ik}+A_{jl}-A_{ik}A_{jl}]\leqslant A_{ij}A_{kl}[A_{ik}+A_{jl}]$, we have

$$\left|\frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\mathbf{ijkl}}\left[\left(\prod_{a=1}^{b}A_{i_aj_a}A_{k_al_a}-\prod_{a=1}^{b}\chi_{i_aj_a}^{k_al_a}(A)\right)F_{\mathbf{ijm}}(A)\right]\right|\leqslant\frac{d}{N}\mathcal{C}(F,A). \tag{3.14}$$

By plugging the estimates (3.13) and (3.14) into (3.12), we get

$$\begin{aligned}&\frac{1}{N^{b+c}d^b}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_aj_a}F_{\mathbf{ijm}}(A)\right]\\&\quad=\frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\substack{\mathbf{ijkl}\\\text{distinct}}}\mathbb{E}\left[\prod_{a=1}^{b}\chi_{i_aj_a}^{k_al_a}(A)F_{\mathbf{ijm}}(A)\right]+\mathrm{O}\left(\frac{d\mathbb{E}[\mathcal{C}(F,A)]}{N}\right).\end{aligned}\tag{3.15}$$

By Proposition 3.1 and an estimate analogous to the one above, we have

$$\begin{aligned}&\frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\substack{\mathbf{ijkl}\\\text{distinct}}}\mathbb{E}\left[\prod_{a=1}^{b}\chi_{i_aj_a}^{k_al_a}(A)F_{\mathbf{ijm}}(A)\right]\\&\quad=\frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\substack{\mathbf{ijkl}\\\text{distinct}}}\mathbb{E}\left[\prod_{a=1}^{b}\chi_{i_ak_a}^{j_al_a}(A)F_{\mathbf{ijm}}\left(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}\right)\right]\\&\quad=\frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\mathbf{ijkl}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_ak_a}A_{j_al_a}F_{\mathbf{ijm}}\left(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}\right)\right]+\mathrm{O}\left(\frac{d\mathbb{E}[\mathcal{C}(F,A)]}{N}\right).\end{aligned}\tag{3.16}$$

The claim now follows from combining (3.15) and (3.16). □

4 Polynomials in Green's Function Entries

In this section we collect some estimates on the Green's function G and the Stieltjes transform of the spectral measure m. These will be used repeatedly in the rest of the paper. We also introduce polynomials in the Green's function entries, and record some of their basic properties. We work under the following assumption throughout this section. Recall the Stieltjes transform m_d of the Kesten–McKay law from (2.5).

Assumption 4.1. We assume that $1 \ll d \ll N^{2/3}$ and that there are deterministic z-dependent control parameters $\Lambda_{\mathrm{o}}, \Lambda_{\mathrm{d}} \in [d^{-1/2}, 1]$ such that

$$\max_i|G_{ii}-m_d| \prec \Lambda_{\mathrm{d}}, \quad \max_{i\neq j}|G_{ij}| \prec \Lambda_{\mathrm{o}},\tag{4.1}$$

for all $z \in \mathbf{D}$ defined in (2.10).

Note that, by Proposition 2.1, we know that Assumption 4.1 holds at least when Λ_{o} and Λ_{d} are given by (2.11).

By our definition, the Green's function $G = P_\perp(H-z)^{-1}P_\perp$ is symmetric and satisfies (2.3). The *Ward identity* states that the Green's function G satisfies

$$\frac{1}{N}\sum_j|G_{ij}|^2=\frac{\mathrm{Im}[G_{ii}]}{N\eta}, \qquad \frac{1}{N}\sum_{ij}|G_{ij}|^2=\frac{\mathrm{Im}[m]}{\eta};\tag{4.2}$$

it can be proved using the resolvent identity on $G - G^*$. Here recall that m is the Stieltjes transform (2.4) of the empirical spectral measure. We record the following basic result, which we shall use tacitly throughout the rest of the paper.

LEMMA 4.2. *Suppose that* $1 \ll d \ll N^{2/3}$.

(i) *For any* $\mathfrak{c} > 0$, *with probability at least* $1 - \mathrm{O}_{\mathfrak{c}}(N^{-1/\mathfrak{c}})$ *we have for all* $z \in \mathbf{D}$

$$\max_{xy} |G_{xy}| \leqslant 2. \tag{4.3}$$

(ii) *Denoting by* $u_\alpha(i)$ *the* i*-th component of the* α*-th normalized eigenvector of* $P_\perp H P_\perp$, *we have the delocalization estimate*

$$\max_{\alpha} \max_{i} |u_\alpha(i)| \prec 1/\sqrt{N}. \tag{4.4}$$

Proof. The claim (i) follows from Proposition 2.1, the estimate $|m_{\mathrm{sc}}| \leqslant 1$, and a simple N^{-3}-net argument in $\mathbf{D}$ combined with a union bound to obtain a simultenous estimate for all $z \in \mathbf{D}$. The claim (ii) follows from [BKY17, Corollary 1.2]. □

REMARK 4.3. More explicitly, Lemma 4.2(i) says that $\max_{xy}|G_{xy}(A)| \leqslant 2$ with probability at least $1 - \mathrm{O}_{\mathfrak{c}}(N^{-1/\mathfrak{c}})$. As a consequence, for any fixed $b \in \mathbb{N}$ we find using a simple resolvent expansion that

$$\max_{xy} \left| G_{xy}\left(A + \sum_{a=1}^{b} \xi_{i_a j_a}^{k_a l_a} \right) \right| \leqslant 2 + \mathrm{O}(d^{-1/2})$$

with probability at least $1 - \mathrm{O}_{\mathfrak{c}}(N^{-1/\mathfrak{c}})$. Moreover, by a similar argument, using (4.1), for the off-diagonal entries we have the estimate

$$\max_{x \neq y} \left| G_{xy}\left(A + \sum_{a=1}^{b} \xi_{i_a j_a}^{k_a l_a} \right) \right| \prec \Lambda_{\mathrm{o}}.$$

We define the discrete and continuous derivatives for any indices i, j, k, l,

$$D_{ij}^{kl} F(A) := F(A + \xi_{ij}^{kl}) - F(A), \quad \partial_{ij}^{kl} F(A) := \sqrt{d-1}\, \partial_t F(A + t\xi_{ij}^{kl})\Big|_{t=0}, \tag{4.5}$$

where the matrix ξ_{ij}^{kl} was defined in (3.1). Note that ∂_{ij}^{kl} is the directional derivative in the direction ξ_{ij}^{kl} of the rescaled variable $H = A/\sqrt{d-1}$. For the discrete derivative operator D_{ij}^{kl}, we have the discrete product rule

$$D_{ij}^{kl}(FG) = (D_{ij}^{kl} F)G + F(D_{ij}^{kl} G) + (D_{ij}^{kl} F)(D_{ij}^{kl} G), \tag{4.6}$$

and the Taylor expansion with remainder gives

$$D_{ij}^{kl} F(A) = \sum_{n=1}^{\mathfrak{b}-1} \frac{1}{n!} \left(\frac{\partial_{ij}^{kl}}{\sqrt{d-1}} \right)^n F(A) + \frac{1}{\mathfrak{b}!} \left(\frac{\partial_{ij}^{kl}}{\sqrt{d-1}} \right)^{\mathfrak{b}} F(A + \theta \xi_{ij}^{kl}), \tag{4.7}$$

for some $0 \leqslant \theta \leqslant 1$.

For any indices i, j, k, l (which might be not distinct), the derivatives of the Green's function entries G_{ij} and the Stieltjes transform m are given by

$$\partial_{ij}^{kl} G_{ij} = -G_{ii}G_{jj} - G_{ij}G_{ij} - G_{ik}G_{lj} - G_{il}G_{kj} + G_{il}G_{jj} + G_{ii}G_{kj} + G_{ik}G_{ij} + G_{ij}G_{lj}, \tag{4.8}$$

$$\begin{aligned} \partial_{ij}^{kl} m &= \frac{2}{N} \sum_{a=1}^{N} (-G_{ia}G_{ja} - G_{ka}G_{la} + G_{ia}G_{ka} + G_{ja}G_{la}) \\ &= \frac{2}{N} (-(G^2)_{ij} - (G^2)_{kl} + (G^2)_{ik} + (G^2)_{jl}). \end{aligned} \tag{4.9}$$

A central object in our proof is the following notion of a polynomial in the entries of the Green's function.

DEFINITION 4.4. (i) *Let $F = F(\{x_{st}\}_{s,t=1}^{r})$ be a polynomial in the r^2 abstract variables $\{x_{st}\}_{s,t=1}^{r}$. We denote its degree by* $\deg(F)$. *For $\mathbf{i} \in [\![N]\!]^r$, we define its* evaluation *on the Green's function by*

$$F_{\mathbf{i}} = F(\{G_{i_s i_t}\}_{s,t=1}^{r}), \tag{4.10}$$

and say that $F_{\mathbf{i}}$ is a polynomial in the Green's function entries $\{G_{i_s i_t}\}_{s,t=1}^{r}$. *By a slight abuse of notation, we sometimes abbreviate F instead of $F_{\mathbf{i}}$ for the polynomials in the Green's function entries when there is no risk of confusion.*

(ii) *Let $F = F(\{x_{st}\}_{s,t=1}^{r})$ be a monomial in r^2 variables. Then the number of off-diagonal entries of F is the total degree of variables x_{st} with $s \neq t$. If the number of off-diagonal entries of F is zero then we define $\chi_F = 1$, otherwise we define $\chi_F = 0$.*

(iii) *For U a polynomial in two variables, by a slight abuse of notation we often abbreviate $U = U(m, \bar{m})$ for the polynomial in the Stieltjes transform m and its complex conjugate $\bar{m}$. For $r, \bar{r} \in \mathbb{N}$ we abbreviate $U^{(r,\bar{r})} := \partial_m^r \partial_{\bar{m}}^{\bar{r}} U$.*

CLAIM 4.5. *Using (4.3) and (4.4), we have for any indices $a, b \in [\![N]\!]$*

$$\left| \frac{1}{N^{k-1}} (G^k)_{ab} \right| \prec \frac{\operatorname{Im}[m]}{(N\eta)^{k-1}}, \quad \frac{1}{N^{2k-1}} (|G|^{2k})_{ab} \prec \frac{\operatorname{Im}[m]}{(N\eta)^{2k-1}}. \tag{4.11}$$

In particular,

$$\frac{1}{N} \sum_{j=1}^{N} |G_{aj}G_{jb}| \prec \frac{\operatorname{Im}[m]}{N\eta}, \tag{4.12}$$

and for distinct indices i, j, k, l, $(\partial_{ij}^{kl})^2 G_{ij}$ is a cubic polynomial in the Green's function with at least one off-diagonal factor,

$$|(\partial_{ij}^{kl})^2 G_{ij}| \prec \Lambda_{\mathrm{o}}. \tag{4.13}$$

We also have the following estimates for the derivatives of the Stieltjes transform m: for any integer $s \geqslant 1$,

$$|(\partial_{ij}^{kl})^s m| \prec \frac{\mathrm{Im}[m]}{N\eta}, \tag{4.14}$$

and for any fixed polynomial U,

$$\left|(\partial_{ij}^{kl})^s U(m,\bar{m})\right| \prec \max_{r+\bar{r}\geqslant 1} |U^{(r,\bar{r})}(m,\bar{m})| \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^r. \tag{4.15}$$

Proof. (4.11) follows directly from the spectral decomposition and the delocalization of the eigenvectors (4.4):

$$\begin{aligned} \left|\frac{1}{N^{k-1}}(G^k)_{ab}\right| &= \left|\frac{1}{N^{k-1}}\sum_\alpha \frac{u_\alpha(a)u_\alpha(b)}{(\lambda_\alpha - z)^k}\right| \\ &\prec \frac{1}{(N\eta)^{k-2}}\frac{1}{N^2}\sum_\alpha \frac{1}{|\lambda_\alpha - z|^2} = \frac{\mathrm{Im}[m]}{(N\eta)^{k-1}}, \\ \left|\frac{1}{N^{2k-1}}(|G|^{2k})_{ab}\right| &= \frac{1}{N^{2k-1}}\left|\sum_\alpha \frac{u_\alpha(a)u_\alpha(b)}{|\lambda_\alpha - z|^{2k}}\right| \\ &\prec \frac{1}{(N\eta)^{2k-2}}\frac{1}{N^2}\sum_\alpha \frac{1}{|\lambda_\alpha - z|^2} = \frac{\mathrm{Im}[m]}{(N\eta)^{2k-1}}. \end{aligned} \tag{4.16}$$

The claim (4.12) follows from Young's inequality and (4.11) by taking $k = 2$,

$$\frac{1}{N}\sum_{j=1}^N |G_{aj}G_{jb}| \leqslant \frac{1}{N}\sum_{j=1}^N (|G_{aj}|^2 + |G_{jb}|^2) = \frac{1}{N}\left((|G|^2)_{aa} + |G|^2_{bb}\right) \prec \frac{\mathrm{Im}[m]}{N\eta}. \tag{4.17}$$

For (4.13), we notice that from (4.8), one can directly verify that none of the derivatives $\partial_{ij}, \partial_{kl}, \partial_{ik}, \partial_{jl}$ produces a diagonal term, and (4.13) follows from (4.3).

For (4.14), one can check using (4.9) that $(\partial_{ij}^{kl})^s m$ is a sum of terms in the following form

$$\frac{1}{N}\sum_{a=1}^N G_{ai_1}G_{i_2i_3}\cdots G_{i_{2s}a} \tag{4.18}$$

where $i_1, i_2, \cdots, i_{2s} \in \{i,j,k,l\}$. Thanks to (4.3) and (4.12), we can bound (4.18) as

$$\left|\frac{1}{N}\sum_{a=1}^N G_{ai_1}G_{i_2i_3}\cdots G_{i_{2s}a}\right| \prec \left|\frac{1}{N}\sum_{a=1}^N G_{ai_1}G_{i_{2s}a}\right| \prec \frac{\mathrm{Im}[m]}{N\eta}, \tag{4.19}$$

and the claim (4.14) follows. For (4.15), the derivative $(\partial_{ij}^{kl})^s U(m)$ is a sum of terms of the form

$$U^{(r,\bar{r})}(m,\bar{m})(\partial_{ij}^{kl})^{s_1}m(\partial_{ij}^{kl})^{s_2}m\cdots(\partial_{ij}^{kl})^{s_r}m(\partial_{ij}^{kl})^{\bar{s}_1}\bar{m}(\partial_{ij}^{kl})^{\bar{s}_2}\bar{m}\cdots(\partial_{ij}^{kl})^{\bar{s}_{\bar{r}}}\bar{m}, \tag{4.20}$$

where $r+\bar{r} \geqslant 1$, $s_1, s_2, \cdots, s_r, \bar{s}_1, \bar{s}_2, \cdots, \bar{s}_{\bar{r}} \geqslant 1$ and $s_1 + \cdots + s_r + \bar{s}_1 + \cdots + \bar{s}_{\bar{r}} = s$. Thanks to (4.14) we have

$$|(4.20)| \prec |U^{(r,\bar{r})}(m,\bar{m})| \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{r+\bar{r}}. \tag{4.21}$$

The claim (4.15) follows from (4.21). □

Claim 4.6. *Let U be a fixed polynomial of the Stieltjes transform m and its complex conjugate $\bar{m}$. For any indices i, j, k, l, a, b and fixed positive integer $\mathfrak{b} > 0$, we have*

$$D_{ij}^{kl} G_{ab}(A) = \sum_{n=1}^{\mathfrak{b}-1} \frac{(-1)^n}{(d-1)^{n/2}} (G(\xi_{ij}^{kl} G)^n)_{ab} + \mathrm{O}_\prec(d^{-\mathfrak{b}/2}) = \mathrm{O}_\prec(d^{-1/2}), \tag{4.22}$$

$$D_{ij}^{kl} m(A) = \mathrm{O}_\prec\left(\frac{\mathrm{Im}[m]}{d^{1/2} N\eta}\right), \tag{4.23}$$

$$D_{ij}^{kl} U(A) = \mathrm{O}_\prec\left(\frac{1}{d^{1/2}} \max_{s+\bar{s}\geqslant 1} |U^{(s,\bar{s})}(m,\bar{m})| \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right). \tag{4.24}$$

Proof. By the Taylor expansion (4.7),

$$D_{ij}^{kl} G_{ab}(A) = \sum_{n=1}^{\mathfrak{b}-1} \frac{1}{n!(d-1)^{n/2}} (\partial_{ij}^{kl})^n G_{ab} + \frac{1}{\mathfrak{b}!(d-1)^{\mathfrak{b}/2}} (\partial_{ij}^{kl})^{\mathfrak{b}} G_{ab}(A + \theta \xi_{ij}^{kl}), \tag{4.25}$$

for some random $\theta \in [0,1]$. Thanks to Lemma 4.2 and Remark 4.3,

$$(\partial_{ij}^{kl})^n G_{ab} = (-1)^n n! (G(\xi_{ij}^{kl} G)^n)_{ab} \prec 1, \quad (\partial_{ij}^{kl})^{\mathfrak{b}} G_{ab}(A + \theta \xi_{ij}^{kl}) \prec 1. \tag{4.26}$$

The expression (4.22) follows from the bound (4.26). The estimate (4.23) follows from averaging (4.25) and using (4.19). The estimate (4.24) follows from (4.23) and the discrete product rule (4.6). □

As an application of Claim 4.6, we have the following estimate, which says essentially that when acting with discrete derivatives on a product of the form FU, where U is a polynomial in m and F a polynomial in the Green's function entries (recall Definition 4.4), the main contribution is given by differentiating F.

Claim 4.7. *Let F be a fixed polynomial of Green's function entries $\{G_{ij}\}_{i,j\in[\![N]\!]}$, and U a fixed polynomial in the Stieltjes transform m and its complex conjugate $\bar{m}$. For any indices $b \geqslant 1$, $i_1, j_1, k_1, l_1, \cdots, i_b, j_b, k_b, l_b$ and positive integer $\mathfrak{b} \geqslant 1$, we have*

$$\begin{aligned} FU\left(A + \sum_{a=1}^{b} \xi_{i_a j_a}^{k_a l_a}\right) - FU(A) &= \left(\sum_{n=1}^{\mathfrak{b}-1} \frac{1}{n!(d-1)^{n/2}} \left(\sum_{a=1}^{b} \partial_{i_a j_a}^{k_a l_a}\right)^n F(A)\right) U(A) \\ &+ \mathrm{O}_\prec\left(\frac{|U(A)|}{d^{\mathfrak{b}/2}} + \frac{|F(A + \sum_{a=1}^{b} \xi_{i_a j_a}^{k_a l_a})|}{d^{1/2}} \max_{s+\bar{s}\geqslant 1} |U^{(s,\bar{s})}(m,\bar{m})| \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right). \end{aligned} \tag{4.27}$$

Proof. We denote $\xi = \sum_{a=1}^{b} \xi_{i_a j_b}^{k_a l_a}$, and rewrite the left-hand side of (4.27) as

$$FU(A+\xi) - FU(A) = (F(A+\xi) - F(A))\, U(A) + F(A+\xi)\,(U(A+\xi) - U(A)). \tag{4.28}$$

Since F is a polynomial of Green's function entries $\{G_{ij}\}_{i,j\in[\![N]\!]}$, by the same argument as for (4.22), we have

$$F(A+\xi) - F(A) = \sum_{n=1}^{\mathfrak{b}-1} \frac{1}{n!(d-1)^{n/2}} \left(\sum_{a=1}^{b} \partial_{i_a j_a}^{k_a l_a} \right)^n F(A) + \mathrm{O}_{\prec}\left(d^{-\mathfrak{b}/2}\right). \tag{4.29}$$

Similarly, since U is a polynomial of the Stieltjes transform m and its complex conjugate $\bar{m}$, by (4.24), we have

$$U(A+\xi) - U(A) = \mathrm{O}_{\prec}\left(\frac{1}{d^{1/2}} \max_{s+\bar{s}\geqslant 1} |U^{(s,\bar{s})}(m,\bar{m})| \left(\frac{\mathrm{Im}[m]}{N\eta} \right)^{s+\bar{s}} \right). \tag{4.30}$$

The claim (4.27) follows from plugging (4.29) and (4.30) into (4.28). □

We conclude this section with an elementary result for the operator $\prec$, which we shall use tacitly throughout the following sections.

Lemma 4.8. *Suppose that A and B are nonnegative random variables satisfying $A \leqslant N^C$ and $B \geqslant N^{-C}$ for some constant $C > 0$. Then $A \prec B$ implies $\mathbb{E}[A] \prec \mathbb{E}[B]$.*

5 Self-consistent Equation in Expectation

In this section we derive the self-consistent equation *in expectation* for the Stieltjes transform m; in Section 7 below we shall extend this self-consistent equation to a high probability estimate.

Proposition 5.1. *Suppose that Assumption 4.1 holds. For every fixed integer $\mathfrak{a} \geqslant 1$, there exists a polynomial, depending on d and $\mathfrak{a}$ but not N, and whose degree depends on $\mathfrak{a}$ only,*

$$P_{\mathfrak{a}}(z,w) = 1 + zw + Q_{\mathfrak{a}}(w), \tag{5.1}$$

where

$$Q_{\mathfrak{a}}(w) = \frac{dw^2}{d-1} + \frac{1}{d}\left(a_3 w^3 + a_4 w^4 + \cdots\right), \tag{5.2}$$

is a polynomial with bounded coefficients $a_3, a_4, \ldots$ such that, for any $z \in \mathbf{D}$,

$$\mathbb{E}[P_{\mathfrak{a}}(z,m)] \prec \frac{1}{d^{\mathfrak{a}/2}} + \frac{\mathbb{E}[\mathrm{Im}[m]]}{N\eta} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}. \tag{5.3}$$

REMARK 5.2. An expansion *in expectation* similar to Proposition 5.1 (with different coefficients) would be possible for the Erdős–Rényi graph in the same regime of expected degree. The essential difference between the random regular and the Erdős–Rényi graph is in the high moment estimates in Section 7, using which we convert the expansion in expectation to one in high probability. For the random regular graph, there are fundamental cancellations arising from the degree constraint, which imply stronger concentration than possible for the Erdős–Rényi graph; these cancellations are manifest only in the high probability expansion. We emphasize that without concentration, the self-consistent equation in expectation does not lead to a closed equation for m (or its expectation), and hence does not provide useful spectral information.

We shall show that the estimate (5.3) results from the switching invariance of random regular graphs. It may be viewed as an approximate *Schwinger–Dyson Equation* for the random regular graph ensemble; in statistical mechanics and field theory, such equations are typically derived by integration by parts.

Before giving the proof of Proposition 5.1, we explain the mechanism behind it. Starting from (2.2) we obtain

$$1 + zm = \sum_{ij} \frac{A_{ij} G_{ij}}{N(d-1)^{1/2}} + (\text{error}),$$

where we use (error) to denote a small error that we do not keep track of in this sketch. Taking the expectation and using Corollary 3.2, recalling that $\sum_i G_{ij} = 0$ and recalling the notation (4.5), we get

$$\mathbb{E}[1 + zm] = \frac{1}{N^2 d(d-1)^{1/2}} \sum_{ijkl} \mathbb{E}\big[A_{ik} A_{jl} D_{ij}^{kl} G_{ji}\big] + (\text{error}).$$

Using the integration by parts formula from Corollary 3.4 we therefore obtain

$$\begin{aligned} \mathbb{E}[1 + zm] = {} & \frac{d}{N^4 (d-1)^{1/2}} \sum_{ijkl} \mathbb{E}\big[D_{ij}^{kl} G_{ji}\big] \\ & + \frac{1}{N^4 d^3 (d-1)^{1/2}} \sum_{ijklrstu} \mathbb{E}\Big[A_{ir} A_{ks} A_{jt} A_{lu} \Big((D_{ij}^{kl} G_{ji})(A + \xi_{ik}^{rs} \\ & + \xi_{jl}^{tu}) - (D_{ij}^{kl} G_{ji})(A) \Big)\Big] + (\text{error}). \end{aligned} \tag{5.4}$$

For the first term on the right-hand side of (5.4), we use the Taylor expansion (4.7) to expand $D_{ij}^{kl} G_{ji}$ as a polynomial in the entries of G, up to a small error. The leading term is $-(d-1)^{-1/2} G_{jj} G_{ii}$, and it yields the first term of (5.2). The other terms are polynomials that either contain off-diagonal entries of G or a higher order. Similarly, for the second term of (5.4), we keep on reapplying inductively the integration by parts formula from Corollary 3.4, and expand all discrete derivatives using Taylor's formula (4.7).

This procedure results in a proliferation of terms that contain products of factors of the form

$$\frac{1}{d^{o/2}}\frac{1}{N^{b+c}d^b}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_aj_a}F_{\mathbf{ijm}}\right] \tag{5.5}$$

where F is a polynomial in the entries of G. Each application of Corollary 3.4 yields a main term (second term on the right-hand side of (3.11)) with no factors of A and another term (first term on the right-hand side of (3.11)) that is (after Taylor expansion of the discrete derivatives) of *higher order*, in the sense of both the degree of $F_{\mathbf{ijm}}$ and the power of $d^{-1/2}$ in front of it. Any main term that contains one or more off-diagonal can be shown to either vanish or be small enough. Hence, we only need to keep track of terms of the form (5.5) with $b = 0$ and $\chi_F = 1$ (recall Definition 4.4).

Such a term is in general not a polynomial in m; consider for instance the term $\frac{1}{N}\sum_i \mathbb{E}[G_{ii}^2]$. An important ingredient of our argument is to rewrite such a term as a corresponding polynomial in m, up to higher order terms; for instance,

$$\frac{1}{N}\sum_i \mathbb{E}[G_{ii}^2] = \frac{1}{N^2}\sum_{ij}\mathbb{E}[G_{ii}G_{jj}] + (\text{higher order}) + (\text{error}),$$

whereby $\frac{1}{N^2}\sum_{ij}\mathbb{E}[G_{ii}G_{jj}] = \mathbb{E}[m^2]$. To explain how this works, consider some polynomial X_i in the Green function entries (think of e.g. $X_i = G_{ii}$). We want to replace $\mathbb{E}[G_{ii}X_i]$ with $\mathbb{E}[G_{jj}X_i]$ up to higher order terms and small errors. From (2.2) we get the equations

$$1 - \frac{1}{N} + zG_{ii} = (HG)_{ii}, \qquad 1 - \frac{1}{N} + zG_{jj} = (HG)_{jj}.$$

Multiplying the first by $G_{jj}X_i$ and the second by $G_{ii}X_i$ and taking the difference, we obtain

$$\begin{aligned}(G_{ii} - G_{jj})X_i &= \big(G_{ii}(HG)_{jj} - G_{jj}(HG)_{ii}\big)X_i + (\text{error})\\ &= \frac{1}{(d-1)^{1/2}}\sum_k \big(A_{jk}G_{ii}G_{kj} - A_{ik}G_{ki}G_{jj}\big)X_i + (\text{error}).\end{aligned}$$

We take the expectation, average over ij, and apply the integration by parts formula of Corollary 3.4 twice. Recalling that $\sum_k G_{ik} = 0$, we therefore obtain

$$\begin{aligned}&\frac{1}{N^2}\sum_{ij}\mathbb{E}[(G_{ii} - G_{jj})X_i]\\ &\quad= \frac{d}{(d-1)^{1/2}N^5}\sum_{ijkrs}\mathbb{E}\big[D_{jk}^{rs}(G_{ii}G_{kj}X_i) - D_{ik}^{rs}(G_{ki}G_{jj}X_i)\big]\\ &\qquad+(\text{higher order}) + (\text{error}).\end{aligned}$$

We expand the discrete derivatives using (4.7). This yields

$$\begin{aligned}
&\frac{1}{N^2}\sum_{ij}\mathbb{E}[(G_{ii}-G_{jj})X_i] \\
&\quad=\frac{d}{(d-1)N^5}\sum_{ijkrs}\mathbb{E}\big[\partial_{jk}^{rs}(G_{ii}G_{kj}X_i)-\partial_{ik}^{rs}(G_{ki}G_{jj}X_i)\big]+(\text{higher order})+(\text{error}) \\
&\quad=\frac{d}{(d-1)N^5}\sum_{ijkrs}\mathbb{E}\big[-G_{ii}G_{kk}G_{jj}X_i+G_{kk}G_{ii}G_{jj}X_i\big]+(\text{higher order})+(\text{error}),
\end{aligned}$$

where we used that all other terms are small because they contain off-diagonal terms. Thus, the leading order terms cancel exactly. We may therefore continue this expansion iteratively on the terms (higher order), which will stop after a finite number of steps, as the order, and hence the power of $d^{-1/2}$, increases at each iteration.

We conclude this informal discussion by noting that the algorithm sketched above that generates the polynomial $Q_{\mathfrak{a}}$, while explicit, is quite complicated and tracking the actual coefficients of (5.2) that it generates is difficult. This issue will be addressed in the next section, where, instead of tracking the coefficients explicitly, we characterize them indirectly by showing that up to a small error term the Stieltjes transform of the Kesten–McKay law m_d is a root of $P_{\mathfrak{a}}(z,\cdot)$, which will imply that the coefficients of $P_{\mathfrak{a}}$ are close to those of (2.8). This concludes the outline of the proof of Proposition 5.1.

The rest of this section is devoted to the proof of Proposition 5.1. Throughout, we fix an integer $\mathfrak{a}\geqslant 1$. The spectral parameter z is always taken in the set $\mathbf{D}$, and our estimates are uniform in z. We shall always work under Assumption 4.1.

5.1 Estimates for moments of the Green's function. We begin with a definition of a family of fundamental terms that form the backbone of our expansion. They are classified by the order of the variable $d^{-1/2}$ and the degree of the polynomial in the entries of G. Both quantities are important to keep track of. The former because it will allow us to stop the recursive application of identities yielding high order terms after a fixed number, $\mathfrak{a}$, of steps, up to an error term of order $d^{-\mathfrak{a}/2}$. The latter is important to ensure that the polynomial in m that we shall ultimately generate will have a large enough degree.

For the following statements, we recall that for multi-indices $\mathbf{i}\in[\![N]\!]^b$ and $\mathbf{j}\in[\![N]\!]^c$, we denote by $\mathbf{ij}\in[\![N]\!]^{b+c}$ their concatenation (and analogously for $\mathbf{ijk}$). Together with this notation, we recall from Definition 4.4 that, for a polynomial F in $(2b+c)^2$ variables and $\mathbf{i},\mathbf{j}\in[\![N]\!]^b$ and $\mathbf{m}\in[\![N]\!]^c$, we write $F_{\mathbf{ijm}}$ for its evaluation in the Green's function entries, and that for a polynomial U in $m,\bar{m}$, we abbreviate $U(m,\bar{m})$ by U.

Definition 5.3.

(i) *For $\mathfrak{o} \in \mathbb{N}$, we define the expressions*

$$\mathcal{T}_{\mathfrak{o}}(F,U) := \frac{1}{d^{\mathfrak{o}/2}} \frac{1}{N^{b+c} d^b} \sum_{\mathbf{m}} \sum_{\mathbf{ij}} \mathbb{E}\left[\prod_{a=1}^{b} A_{i_a j_a} F_{\mathbf{ijm}}(A) U(A) \right]. \tag{5.6}$$

(ii) *For a given polynomial $U = U(m, \bar{m})$ and integers $\mathfrak{o}, \mathfrak{d} \in \mathbb{N}$, we use the symbol*

$$\mathcal{T}_{\mathfrak{o},\mathfrak{d}}(U)$$

to denote a finite linear combination of terms of the form $\alpha \mathcal{T}_{\tilde{\mathfrak{o}}}(\tilde{F}, U)$, where $\tilde{\mathfrak{o}} \geqslant \mathfrak{o}$, $\deg(\tilde{F}) \geqslant \mathfrak{d}$, and $\alpha = C\big(d/(d-1)\big)^r$ for some $r \in \mathbb{Z}/2$ and a deterministic constant $C \in \mathbb{R}$.

Note that $\mathcal{T}_{\mathfrak{o}}$ depends on $\mathfrak{o} \in \mathbb{N}$ only by a multiplicative factor $d^{-\mathfrak{o}/2}$. From (4.3), we have $|F_{\mathbf{ijm}}| \prec 1$, and thus by Lemma 4.8 we find the a priori estimate

$$|\mathcal{T}_{\mathfrak{o}}(F,U)| \prec \frac{1}{d^{\mathfrak{o}/2}} \frac{1}{N^{b+c} d^b} \sum_{\mathbf{m}} \sum_{\mathbf{ij}} \mathbb{E}\left[\prod_{a=1}^{b} A_{i_a j_a} |U(A)| \right] = \frac{\mathbb{E}[|U(A)|]}{d^{\mathfrak{o}/2}}. \tag{5.7}$$

Proposition 5.1 will follow from much more precise estimates that follow from an inductive application of the following proposition, which extracts the leading term from (5.6) and shows that it can be expressed as a monomial in the trace $m = \frac{1}{N} \operatorname{Tr} G$ rather than the individual Green's function entries. To prove Proposition 5.1, we only need the special case $U = 1$; we allow for a general U for later use in Section 7.

Aside from the notation introduced in Definition 5.3, recall that $\mathcal{C}(F, A)$ was defined in Definition 3.3, and that $F_{\mathbf{i}}$, U, and χ_F were defined in Definition 4.4. Also recall Λ_o from (4.1).

Proposition 5.4. *Fix $\mathfrak{o} \in \mathbb{N}$. Let F be a fixed monic monomial in $(2b+c)^2$ abstract variables, with degree $\deg(F)$. Then*

$$\begin{aligned} \frac{1}{d^{\mathfrak{o}/2}} \frac{1}{N^{b+c} d^b} \sum_{\mathbf{m}} \sum_{\mathbf{ij}} \mathbb{E}\left[\prod_{a=1}^{b} A_{i_a j_a} F_{\mathbf{ijm}} U \right] &= \frac{\chi_F \mathbb{E}[m^{\deg(F)} U]}{d^{\mathfrak{o}/2}} + \mathcal{T}_{\mathfrak{o}+1, \deg(F)+1}(U) \\ &+ \frac{1}{d^{\mathfrak{o}/2}} \mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta} + \frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}} + \frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N} + \Lambda_{\mathrm{o}}^{2-\chi_F} \max_{s+\bar{s} \geqslant 1} \mathbb{E}\left[|U^{(s,\bar{s})}| \left(\frac{\mathrm{Im}[m]}{N\eta} \right)^{s+\bar{s}} \right] \right). \end{aligned} \tag{5.8}$$

To prove Proposition 5.4, we shall use the following claims. The first claim states that the averages of monomials with more than one off-diagonal Green's function terms are subleading.

Claim 5.5. *Let F be a fixed monomial in b^2 abstract variables with at least two off-diagonal entries. Then*

$$\left| \frac{1}{N^b} \sum_{\mathbf{i}} \mathbb{E}\left[F_{\mathbf{i}} U \right] \right| \prec \frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}. \tag{5.9}$$

Proof. Using $|G_{xy}| \prec 1$ and $U \prec 1$ from (4.3) to bound all except the two off-diagonal factors of G in $F_{\mathbf{i}}$ and then using the Cauchy–Schwarz inequality, we have

$$\left|\frac{1}{N^b}\sum_{\mathbf{i}}\mathbb{E}\left[F_{\mathbf{i}}U\right]\right| \prec \frac{1}{N^b}\sum_{\mathbf{i}}\mathbb{E}\left[|G_{i_1i_2}|^2|U|\right] = \frac{1}{N^2}\sum_{i_1,i_2}\mathbb{E}\left[|G_{i_1i_2}|^2|U|\right] = \frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}, \tag{5.10}$$

where the last equality is the Ward identity (4.2). □

The following claim separates the leading order term of $\mathcal{T}_{\mathfrak{o}}(F,U)$ plus other terms of higher order and much small error terms. It says that, to leading order, each factor of A in (5.6) can be replaced with its expectation d/N.

Claim 5.6. *Fix $\mathfrak{o} \in \mathbb{N}$. Let F be a fixed monomial in $(2b+c)^2$ abstract variables and let U be a fixed polynomial in m. Then*

$$\begin{aligned}
&\frac{1}{d^{\mathfrak{o}/2}}\frac{1}{N^{b+c}d^b}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_aj_a}F_{\mathbf{ijm}}U\right] \\
&\quad= \frac{1}{d^{\mathfrak{o}/2}}\frac{1}{N^{2b+c}}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[F_{\mathbf{ijm}}U\right] + \mathcal{T}_{\mathfrak{o}+1,\deg(F)+1}(U) \\
&\qquad+ \frac{1}{d^{\mathfrak{o}/2}}\,\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}} + \frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N} + \frac{\Lambda_{\mathrm{o}}^{1-\chi_F}}{d^{1/2}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).
\end{aligned} \tag{5.11}$$

Proof. We prove the statement for $\mathfrak{o} = 0$; the general statement follows by multiplying both sides by $1/d^{\mathfrak{o}/2}$. By Corollary 3.4, the left-hand side of (5.11) is

$$\begin{aligned}
&\frac{1}{N^{2b+c}}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[F_{\mathbf{ijm}}U(A)\right] \\
&\quad+ \frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\mathbf{ijkl}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_ak_a}A_{j_al_a}\left(F_{\mathbf{ijm}}U\left(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}\right) - F_{\mathbf{ijm}}U(A)\right)\right] \\
&\quad+ \mathrm{O}_{\prec}\left(\frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}\right),
\end{aligned} \tag{5.12}$$

where we used Remark 4.3 to estimate $|F_{\mathbf{ijm}}| \prec 1$ as well as Lemma 4.8. For the second term in (5.12), we use Claim 4.7 with $F = F_{\mathbf{ijm}}$. The term resulting from the first term on the right-hand side of (4.27) gives rise to $\mathcal{T}_{\mathfrak{o}+1,\deg(F)+1}(U)$. For the error terms, we note that, by Remark 4.3 we have

$$\frac{1}{N^{2b+c}d^{2b}}\sum_{\mathbf{m}}\sum_{\mathbf{ijkl}}\prod_{a=1}^{b}A_{i_ak_a}A_{j_al_a}\left|F_{\mathbf{ijm}}\left(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}\right)\right| \prec \frac{d}{N} + \Lambda_{\mathrm{o}}^{1-\chi_F} \leqslant 2\Lambda_{\mathrm{o}}^{1-\chi_F}, \tag{5.13}$$

where in the last step we used Assumption 4.1. In summary, the error terms resulting from the application of Claim 4.7 to the second term of (5.12) are bounded by

$$\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{\Lambda_{\mathrm{o}}^{1-\chi_F}}{d^{1/2}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).$$

The proof is therefore complete. □

The following claim is a decoupling argument: when averaging over an index i that appears in a diagonal Green's function entry G_{ii} and possibly many other places as well, up to some error terms we can replace G_{ii} with $G_{i'i'}$, where i' is a new summation index that appears in no other place, over which we take the average. For example, this allows us to convert an expression of the form $\frac{1}{N}\sum_i (G_{ii})^2$ to a polynomial in m of the form m^2.

Claim 5.7. *Fix $\mathfrak{o}\in\mathbb{N}$. Let F be a fixed monomial in $(1+c)^2$ abstract variables. Then*

$$\begin{aligned}
&\frac{1}{d^{\mathfrak{o}/2}}\frac{1}{N^{2+c}}\sum_{ii'\mathbf{m}}\mathbb{E}[G_{ii}F_{i\mathbf{m}}U]\\
&\quad=\frac{1}{d^{\mathfrak{o}/2}}\frac{1}{N^{2+c}}\sum_{ii'\mathbf{m}}\mathbb{E}[G_{i'i'}F_{i\mathbf{m}}U]+\mathcal{T}_{\mathfrak{o}+1,\deg(F)+3}(U)\\
&\qquad+\frac{1}{d^{\mathfrak{o}/2}}\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}+\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}\right.\\
&\qquad\left.+\Lambda_{\mathrm{o}}^{2-\chi_F}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).
\end{aligned}\tag{5.14}$$

Proof. We prove the statement for $\mathfrak{o}=0$, the general statement follows by multiplying both sides by $1/d^{\mathfrak{o}/2}$. By the definition of the Green's function (2.2), we have

$$\left(1-\frac{1}{N}\right)=-zG_{i'i'}+\sum_{j=1}^{N}\frac{A_{i'j}G_{ji'}}{\sqrt{d-1}},\tag{5.15}$$

$$\left(1-\frac{1}{N}\right)=-zG_{ii}+\sum_{j=1}^{N}\frac{A_{ij}G_{ji}}{\sqrt{d-1}}.\tag{5.16}$$

Multiplying (5.15) and (5.16) by $G_{ii}F_{i\mathbf{m}}U$ and $G_{i'i'}F_{i\mathbf{m}}U$ respectively, averaging over the indices, and then taking the difference, we get

$$\begin{aligned}
\frac{1}{N^{2+c}}\sum_{ii'\mathbf{m}}\mathbb{E}[G_{ii}F_{i\mathbf{m}}U]&=\frac{1}{N^{2+c}}\sum_{ii'\mathbf{m}}\mathbb{E}[G_{i'i'}F_{i\mathbf{m}}U]+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[|U|]}{N}\right)\\
&+\frac{1}{N^{2+c}(d-1)^{1/2}}\sum_{ii'j\mathbf{m}}\left(\mathbb{E}[A_{i'j}G_{ji'}G_{ii}F_{i\mathbf{m}}U]-\mathbb{E}[A_{ij}G_{ji}G_{i'i'}F_{i\mathbf{m}}U]\right),
\end{aligned}\tag{5.17}$$

where we used that $|G_{ii}F_{i\mathbf{m}}| \prec 1$ and $|G_{i'i'}F_{i\mathbf{m}}| \prec 1$. We shall show that the difference of the two terms on the right-hand side of (5.17) is of order $\mathfrak{o}$ greater than 0, up to small error terms.

Using Corollary 3.4 we find

$$\begin{aligned}&\frac{1}{N^{2+c}(d-1)^{1/2}}\sum_{ii'j\mathbf{m}}\mathbb{E}[A_{i'j}G_{ji'}G_{ii}F_{i\mathbf{m}}U]\\&\quad=\frac{1}{N^{3+c}d(d-1)^{1/2}}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[A_{i'k}A_{jl}D^{kl}_{i'j}(G_{ji'}G_{ii}F_{i\mathbf{m}}U)]+\mathrm{O}_\prec\left(\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}\right),\end{aligned}\tag{5.18}$$

where we used that $\sum_{i'}G_{ji'}=0$, so that the main term in (3.11) vanishes, and from $|G_{ji'}|$ we gain an off-diagonal factor that is estimated by Λ_{o}. For the first term on the right-hand side of (5.18), by the discrete product rule (4.6) and Claim 4.6,

$$\begin{aligned}&D^{kl}_{i'j}(G_{ji'}G_{ii}F_{i\mathbf{m}}U)=D^{kl}_{i'j}(G_{ji'}G_{ii}F_{i\mathbf{m}})U+G_{ji'}G_{ii}F_{i\mathbf{m}}D^{kl}_{i'j}(U)+D^{kl}_{i'j}(G_{ji'}G_{ii}F_{i\mathbf{m}})D^{kl}_{i'j}(U)\\&\quad=D^{kl}_{i'j}(G_{ji'}G_{ii}F_{i\mathbf{m}})U+\mathrm{O}_\prec\left(\left|G_{ji'}G_{ii}F_{i\mathbf{m}}+D^{kl}_{i'j}(G_{ji'}G_{ii}F_{i\mathbf{m}})\right|\frac{1}{\sqrt{d}}\max_{s+\bar{s}\geqslant 1}|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right).\end{aligned}\tag{5.19}$$

We notice that $G_{ji'}G_{ii}F_{i\mathbf{m}}$ contains at least $2-\chi_F$ off-diagonal entries and $D^{kl}_{i'j}(G_{ji'}G_{ii}F_{i\mathbf{m}})$ contains at least $1-\chi_F$ off-diagonal entries. Thus, by plugging (5.19) into (5.18), we get

$$\begin{aligned}&\frac{1}{N^{2+c}(d-1)^{1/2}}\sum_{ii'j\mathbf{m}}\mathbb{E}[A_{i'j}G_{ji'}G_{ii}F_{i\mathbf{m}}U]\\&\quad=\frac{1}{N^{3+c}d(d-1)^{1/2}}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[A_{i'k}A_{jl}D^{kl}_{i'j}(G_{ji'}G_{ii}F_{i\mathbf{m}})U]\\&\qquad+\mathrm{O}_\prec\left(\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}+\Lambda_o^{2-\chi_F}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right)\\&\quad=\frac{1}{N^{3+c}d}\sum_{n=1}^{\mathfrak{a}}\frac{1}{n!(d-1)^{(n+1)/2}}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[A_{i'k}A_{jl}(\partial^{kl}_{i'j})^n(G_{ji'}G_{ii}F_{i\mathbf{m}})U]\\&\qquad+\mathrm{O}_\prec\left(\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}+\Lambda_o^{2-\chi_F}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right),\end{aligned}\tag{5.20}$$

where the last step follows by Taylor expansion, as in Claim 4.7. The remaining derivative $(\partial^{kl}_{i'j})^n(G_{ji'}G_{ii}F_{i\mathbf{m}})$ is again a polynomial in $\{G_{xy}\}_{x,y\in ii'jkl\mathbf{m}}$, and thus this term is in the form $\mathcal{T}_{n-1,\deg(F)+2+n}(U)$. Treating the terms $n=1$ and $n\geqslant 2$ separately, we get

$$\frac{1}{N^{3+c}d(d-1)}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[A_{i'k}A_{jl}\partial_{i'j}^{kl}(G_{ji'}G_{ii}F_{i\mathbf{m}})U]+\mathcal{T}_{1,\deg(F)+4}(U)$$
$$+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}+\Lambda_o^{2-\chi_F}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right). \tag{5.21}$$

By Claim 5.6, the first term in (5.21) can be expanded as

$$\begin{aligned}&\frac{1}{N^{3+c}d(d-1)}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[A_{i'k}A_{jl}\partial_{i'j}^{kl}(G_{ji'}G_{ii}F_{i\mathbf{m}})U]\\&\quad=\frac{d}{N^{5+c}(d-1)}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[\partial_{i'j}^{kl}(G_{ji'}G_{ii}F_{i\mathbf{m}})U]+\mathcal{T}_{1,\deg(F)+3}(U)\\&\qquad+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}+\frac{\Lambda_{\mathrm{o}}^{1-\chi_F}}{d^{1/2}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).\end{aligned} \tag{5.22}$$

Moreover, $\partial_{i'j}^{kl}(G_{ji'}G_{ii}F_{i\mathbf{m}})=G_{ji'}\partial_{i'j}^{kl}(G_{ii}F_{i\mathbf{m}})+\partial_{i'j}^{kl}(G_{ji'})G_{ii}F_{i\mathbf{m}}$. The first term $G_{ji'}\partial_{i'j}^{kl}(G_{ii}F_{i\mathbf{m}})$ contains at least two off-diagonal Green's function entries. Thus, by the Claim 5.5, we have

$$\left|\frac{d}{N^{5+c}(d-1)}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[G_{ji'}\partial_{i'j}^{kl}(G_{ii}F_{i\mathbf{m}})U]\right|\prec\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}. \tag{5.23}$$

To analyse the second term, we write $\partial_{i'j}^{kl}(G_{ji'})=-(G\xi_{i'j}^{kl}G)_{ji'}=-G_{jj}G_{i'i'}+G_{jj}G_{li'}+G_{jk}G_{i'i'}-G_{ji'}G_{ji'}-G_{jk}G_{li'}-G_{jl}G_{ki'}+G_{ji'}G_{ki'}+G_{jl}G_{ji'}$. Since the row and column sums of G are zero, all but the first and fourth terms vanish when taking the average over the indices $i'jkl$. The fourth term has two off-diagonal Green's function entries and can therefore be estimated using Claim 5.5. Thus we have

$$\begin{aligned}&\frac{d}{N^{5+c}(d-1)}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[\partial_{i'j}^{kl}(G_{ji'})G_{ii}F_{i\mathbf{m}}U]\\&\quad=-\frac{d}{N^{5+c}(d-1)}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[G_{jj}G_{i'i'}G_{ii}F_{i\mathbf{m}}U]+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}\right).\end{aligned} \tag{5.24}$$

By combining the estimates (5.20), (5.21), (5.22) and (5.23), we get

$$\begin{aligned}&\frac{1}{N^{2+c}(d-1)^{-1/2}}\sum_{ii'j\mathbf{m}}\mathbb{E}[A_{i'j}G_{ji'}G_{ii}F_{i\mathbf{m}}U]\\&\quad=-\frac{d}{N^{5+c}(d-1)}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[G_{jj}G_{i'i'}G_{ii}F_{i\mathbf{m}}U]+\mathcal{T}_{1,\deg(F)+3}(U)\end{aligned}$$

$$+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}+\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}\right.$$
$$\left.+\Lambda_{\mathrm{o}}^{2-\chi_F}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right). \tag{5.25}$$

Analogously, repeating the above argument for the last term in (5.17), we find

$$\begin{aligned}&\frac{1}{N^{2+c}(d-1)^{1/2}}\sum_{ii'j\mathbf{m}}\mathbb{E}[A_{ij}G_{ji}G_{i'i'}F_{i\mathbf{m}}]\\&\quad=-\frac{d}{N^{5+c}(d-1)}\sum_{ii'jkl\mathbf{m}}\mathbb{E}[G_{jj}G_{i'i'}G_{ii}F_{i\mathbf{m}}U]+\mathcal{T}_{1,\deg(F)+3}(U)\\&\qquad+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}+\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}\right.\\&\qquad\left.+\Lambda_{\mathrm{o}}^{2-\chi_F}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).\end{aligned} \tag{5.26}$$

Since the first terms on the right-hand sides of (5.25) and (5.26) are the same, they cancel upon taking their difference, and our claim (5.14) follows by combining (5.17), (5.25) and (5.26). (Note that the error term on the right-hand side of (5.17) can be absorbed into the third error term of (5.14).) □

Proof of Proposition 5.4. We prove the statement for $\mathfrak{o}=0$, the general statement follows by multiplying both sides by $1/d^{\mathfrak{o}/2}$. By Claim 5.6, we have

$$\begin{aligned}&\frac{1}{N^{b+c}d^b}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[\prod_{a=1}^{b}A_{i_aj_a}F_{\mathbf{ijm}}U\right]\\&\quad=\frac{1}{N^{2b+c}}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[F_{\mathbf{ijm}}U\right]+\mathcal{T}_{1,\deg(F)+1}(U)\\&\qquad+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}+\frac{\Lambda_{\mathrm{o}}^{1-\chi_F}}{d^{1/2}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).\end{aligned} \tag{5.27}$$

We now estimate the first term on the right-hand side, distinguishing three cases.
Case 1. The monomial $F_{\mathbf{ijm}}$ has more than one off-diagonal Green's function factors. Then, by Claim 5.5,

$$\frac{1}{N^{2b+c}}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[F_{\mathbf{ijm}}U\right]\prec\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}. \tag{5.28}$$

Case 2. The monomial $F_{\mathbf{ijm}}$ contains exactly one off-diagonal Green's function factor. Then, without loss of generality, we assume that $F_{\mathbf{ijm}}=G^{r}_{m_1m_1}G_{m_1m_2}\tilde{F}_{\mathbf{ij}m_2\cdots m_c}$, where $\tilde{F}_{\mathbf{ij}m_2\cdots m_c}$ is a monomial in terms of the Green's function entries

$\{G_{xx}\}_{x\in\{i_1,j_1,\cdots,i_b,j_b,m_2,\cdots,m_c\}}$ and $r\in\mathbb{N}$. If $r=0$ then $F_{\mathbf{ijm}}$ vanishes upon taking the average over $\mathbf{m}$ since $\sum_{m_1}G_{m_1m_2}=0$. For $r\geqslant 1$, we introduce new indices $m_1^1,m_1^2,\cdots,m_1^r$, and repeatedly use Claim 5.7 to replace $G^r_{m_1m_1}$ by $G_{m_1^1m_1^1}G_{m_1^2m_1^2}\cdots G_{m_1^rm_1^r}$. This way we obtain

$$\begin{aligned}
&\frac{1}{N^{2b+c}}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[G^r_{m_1m_1}G_{m_1m_2}\tilde{F}_{\mathbf{ij}m_2\cdots m_c}U\right]\\
&\quad=\frac{1}{N^{2b+c+r}}\sum_{m_1^1,\cdots,m_1^r}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[\prod_{a=1}^{r}G_{m_1^am_1^a}G_{m_1m_2}\tilde{F}_{\mathbf{ij}m_2\cdots m_c}U\right]+\mathcal{T}_{1,\deg(F)+2}(U)\\
&\qquad+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}+\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}\right.\\
&\qquad\left.+\Lambda_{\mathrm{o}}^2\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right)\\
&\quad=\mathcal{T}_{1,\deg(F)+2}(U)\\
&\qquad+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}+\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}\right.\\
&\qquad\left.+\Lambda_{\mathrm{o}}^2\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right),
\end{aligned}\tag{5.29}$$

where in the last equality we used that $\sum_{m_1}G_{m_1m_2}=0$.

Case 3. The monomial $F_{\mathbf{ijm}}$ contains only diagonal Green's function terms. Then, by the same argument as in Case 2, we can repeatedly use Claim 5.7 to get

$$\begin{aligned}
&\frac{1}{N^{2b+c}}\sum_{\mathbf{m}}\sum_{\mathbf{ij}}\mathbb{E}\left[F_{\mathbf{ijm}}U\right]\\
&\quad=\mathbb{E}[m^{\deg(F)}U]+\mathcal{T}_{1,\deg(F)+2}(U)\\
&\qquad+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}+\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}+\frac{d^{3/2}\Lambda_{\mathrm{o}}\mathbb{E}[\mathcal{C}(U,A)]}{N}\right.\\
&\qquad\left.+\Lambda_{\mathrm{o}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).
\end{aligned}\tag{5.30}$$

By (5.27) and putting the three cases, (5.28), (5.29) and (5.30), together, the proposition follows. □

5.2 Proof of Proposition 5.1. We prove the following proposition, from which Proposition 5.1 will follow easily by taking $U=1$. The general form of Proposition 5.8 will be used in Section 7.

PROPOSITION 5.8. *Suppose that Assumption 4.1 holds. For every fixed integer* $\mathfrak{a} \geqslant 1$*, there exists a polynomial, depending on* d *and* $\mathfrak{a}$ *but not* N*, and whose degree depends on* $\mathfrak{a}$ *only,*

$$Q_{\mathfrak{a}}(w) = \frac{dw^2}{d-1} + \frac{1}{d}\left(a_3 w^3 + a_4 w^4 + \cdots\right),$$

with bounded coefficients $a_3, a_4, \ldots$ *such that, for any* $z \in \mathbf{D}$,

$$\begin{aligned} &\frac{1}{N^2 d(d-1)^{1/2}} \sum_{ijkl} \mathbb{E}[A_{ik}A_{jl}(D_{ij}^{kl}G_{ij})U] + \mathbb{E}[Q_{\mathfrak{a}}U] \\ &\quad = \mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta} + \frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}} + \frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N} + \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}} \max_{s+\bar{s}\geqslant 1} \mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right). \end{aligned} \tag{5.31}$$

Proof. By (4.22) and Lemma 4.2,

$$\begin{aligned} \frac{1}{N^2}\sum_{ijkl} \frac{\mathbb{E}[A_{ik}A_{jl}(D_{ij}^{kl}G_{ij})U]}{d(d-1)^{1/2}} &= \sum_{ijkl} \frac{\mathbb{E}[A_{ik}A_{jl}(\partial_{ij}^{kl}G_{ij})U]}{d(d-1)N^2} \\ &\quad + \sum_{ijkl} \frac{\mathbb{E}[A_{ik}A_{jl}((\partial_{ij}^{kl})^2 G_{ij})U]}{2d(d-1)^{3/2}N^2} + \mathcal{T}_{2,4}(U) + \mathrm{O}_{\prec}\left(\frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}}\right). \end{aligned} \tag{5.32}$$

For the first term on the right-hand side of (5.32), the derivative $\partial_{ij}^{kl}G_{ij}$ is given by (4.8). For the last four terms, $G_{il}G_{jj} + G_{ii}G_{kj} + G_{ik}G_{ij} + G_{ij}G_{lj}$, using that $\sum_j G_{ij} = 0$, we have

$$\sum_{ijkl} \frac{1}{d(d-1)N^2}\mathbb{E}[A_{ik}A_{jl}(G_{il}G_{jj} + G_{ii}G_{kj} + G_{ik}G_{ij} + G_{ij}G_{lj})U] = 0. \tag{5.33}$$

For the term $-G_{ij}G_{ij}$, using (4.12), we have

$$-\sum_{ijkl} \frac{1}{d(d-1)N^2}\mathbb{E}[A_{ik}A_{jl}G_{ij}G_{ij}U] = -\sum_{ij} \frac{d}{(d-1)N^2}\mathbb{E}[G_{ij}G_{ij}U] \prec \frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}. \tag{5.34}$$

For the term $-G_{il}G_{kj}$, we use $(H-z)G = P_{\perp}$ and (4.12) to get

$$\begin{aligned} -\sum_{ijkl} \frac{1}{d(d-1)N^2}\mathbb{E}[A_{ik}A_{jl}G_{il}G_{kj}U] &= -\frac{1}{d(d-1)N^2}\mathbb{E}[\mathrm{Tr}(AGAG)U] \\ &= -\frac{1}{dN^2}\mathbb{E}[\mathrm{Tr}(z^2G^2 + 2zG + P_{\perp})U] \prec \frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{dN\eta} + \frac{\mathbb{E}[|U|]}{dN} \prec \frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{dN\eta}. \end{aligned} \tag{5.35}$$

Summarizing, we can rewrite the first term on the right-hand side of (5.32) as

$$\begin{aligned}\sum_{ijkl}\frac{\mathbb{E}[A_{ik}A_{jl}(\partial_{ij}^{kl}G_{ij})U]}{d(d-1)N^2} &= -\frac{d}{d-1}\mathbb{E}[m^2U]\\ &\quad -\sum_{ijkl}\frac{\mathbb{E}[A_{ik}A_{jl}G_{ik}G_{jl}U]}{d(d-1)N^2}+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}\right).\end{aligned} \tag{5.36}$$

By Corollary 3.4, $\sum_j G_{ij}=0$ and the trivial extension of the product rule (4.6) and (4.24) to differences in the direction $\xi_{ik}^{i'k'}+\xi_{jl}^{j'l'}$, the second term on the right-hand side of (5.36) is

$$\begin{aligned}\sum_{ijkl}\frac{\mathbb{E}[A_{ik}A_{jl}G_{ik}G_{jl}U]}{d(d-1)N^2} &= \mathrm{O}_{\prec}\left(\frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}\right)\\ &\quad+\frac{1}{d^3(d-1)N^4}\sum_{ii'jj'kk'll'}\mathbb{E}\left[A_{ii'}A_{kk'}A_{jj'}A_{ll'}\left(G_{ik}G_{jl}U\left(A+\xi_{ik}^{i'k'}+\xi_{jl}^{j'l'}\right)-G_{ik}G_{jl}U(A)\right)\right]\\ &=\frac{1}{d^3(d-1)N^4}\sum_{ii'jj'kk'll'}\mathbb{E}\left[A_{ii'}A_{kk'}A_{jj'}A_{ll'}\left(G_{ik}G_{jl}\left(A+\xi_{ik}^{i'k'}+\xi_{jl}^{j'l'}\right)-G_{ik}G_{jl}(A)\right)U\right]\\ &\quad+\mathrm{O}_{\prec}\left(\frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}+\frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right)\\ &=\frac{1}{d^3(d-1)^{3/2}N^4}\sum_{ii'jj'kk'll'}\mathbb{E}\left[A_{ii'}A_{kk'}A_{jj'}A_{ll'}(\partial_{ik}^{i'k'}+\partial_{jl}^{j'l'})(G_{ik}G_{jl})U\right]\\ &\quad+\mathcal{T}_{2,4}(U)+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[|U|]}{d^{(\mathfrak{a}+1)/2}}+\frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}+\frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right),\end{aligned} \tag{5.37}$$

where in the first equality, we used Remark 4.3 and that $G_{ik}G_{jl}$ contains two off-diagonal terms, and in the second equality we used Claim 4.7. For the first term on the right-hand side of (5.37), we notice that $(\partial_{ik}^{i'k'}+\partial_{jl}^{j'l'})(G_{ik}G_{jl})$ contains at least one off-diagonal term. By Proposition 5.4 we get

$$\begin{aligned}&\frac{1}{d^3(d-1)^{3/2}N^4}\sum_{ii'jj'kk'll'}\mathbb{E}\left[A_{ii'}A_{kk'}A_{jj'}A_{ll'}(\partial_{ik}^{i'k'}+\partial_{jl}^{j'l'})(G_{ik}G_{jl})U\right]=\mathcal{T}_{2,4}(U)\\ &\quad+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}+\frac{\mathbb{E}[|U|]}{d^{(\mathfrak{a}+1)/2}}+\frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}+\frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).\end{aligned} \tag{5.38}$$

It follows by combining (5.36), (5.37), and (5.38) that

$$\begin{aligned}\sum_{ijkl}\frac{\mathbb{E}[A_{ik}A_{jl}(\partial_{ij}^{kl}G_{ij})U]}{d(d-1)N^2} &= -\frac{d}{d-1}\mathbb{E}[m^2U]+\mathcal{T}_{2,4}(U)\\ &\quad+\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta}+\frac{\mathbb{E}[|U|]}{d^{(\mathfrak{a}+1)/2}}+\frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}+\frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right).\end{aligned} \tag{5.39}$$

For the second term on the right-hand side of (5.32), we notice that $(\partial_{ij}^{kl})^2 G_{ij}$ contains at least one off-diagonal term. By Proposition 5.4 we get

$$\sum_{ijkl} \frac{\mathbb{E}[A_{ik}A_{jl}((\partial_{ij}^{kl})^2 G_{ij})U]}{2d(d-1)^{3/2}N^2} = \mathcal{T}_{2,3}(U) + \mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta} + \frac{\mathbb{E}[|U|]}{d^{(\mathfrak{a}+1)/2}} + \frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N} + \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}} \max_{s+\bar{s}\geqslant 1} \mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right). \tag{5.40}$$

We plug (5.39) and (5.40) into (5.32), which yields

$$\frac{1}{N^2}\sum_{ijkl} \frac{\mathbb{E}[A_{ik}A_{jl}(D_{ij}^{kl}G_{ij})U]}{d(d-1)^{1/2}} = -\frac{d}{d-1}\mathbb{E}[m^2 U] + \mathcal{T}_{2,3}(U) + \mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta} + \frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}} + \frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N} + \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}} \max_{s+\bar{s}\geqslant 1} \mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right). \tag{5.41}$$

Next, we apply Proposition 5.4 repeatedly to the term $\mathcal{T}_{2,3}(U)$ on the right-hand side of (5.41). For $\mathfrak{o} \geqslant 2$ and $\mathfrak{d} \geqslant 3$, Proposition 5.4 yields

$$\begin{aligned}\mathcal{T}_{\mathfrak{o},\mathfrak{d}}(U) = \frac{1}{d^{\mathfrak{o}/2}}\mathbb{E}[pU] + \mathcal{T}_{\mathfrak{o}+1,\mathfrak{d}+1}(U) \\ + \frac{1}{d}\,\mathrm{O}_{\prec}\left(\frac{\mathbb{E}[\mathrm{Im}[m]|U|]}{N\eta} + \frac{\mathbb{E}[|U|]}{d^{\mathfrak{a}/2}} + \frac{d\mathbb{E}[\mathcal{C}(U,A)]}{N}\right. \\ \left. + \Lambda_{\mathrm{o}} \max_{s+\bar{s}\geqslant 1} \mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\right),\end{aligned} \tag{5.42}$$

where p is a polynomial in m, with bounded coefficients depending only on d, of degree at least $\mathfrak{d}$. Applying (5.42) to $\mathcal{T}_{2,3}(U)$ in (5.41) and then iterating (5.42) $\mathfrak{a}$ times concludes the proof, by (5.7). □

Proof of Proposition 5.1. Let $Q_{\mathfrak{a}}(w)$ be as constructed in Proposition 5.8. By taking the normalized trace on both sides of (2.2), we have

$$1 + zm = \frac{1}{N} + \sum_{ij} \frac{A_{ij}G_{ij}}{N(d-1)^{1/2}} = \frac{1}{N} + \sum_{i\neq j} \frac{A_{ij}G_{ij}}{N(d-1)^{1/2}}, \tag{5.43}$$

where the last equality follows since $A_{ii} = 0$. By Corollary 3.2 with $F = G_{ij}$, we have for $i \neq j$, $|F(A)| + \max_{kl}|F(A + \xi_{ij}^{kl})| \prec \Lambda_{\mathrm{o}}$ (since $\Lambda_{\mathrm{o}} \geqslant d^{-1/2}$, see also Remark 4.3), and

$$\begin{aligned}\mathbb{E}[1 + zm] &= \frac{1}{N(d-1)^{1/2}} \sum_{i\neq j} \mathbb{E}[A_{ij}G_{ij}] + \frac{1}{N} \\ &= \frac{1}{N^2}\sum_{ijkl}\left(\frac{1}{d(d-1)^{1/2}}\mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}G_{ij}]\right) + \mathrm{O}_{\prec}\left(\frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\right).\end{aligned} \tag{5.44}$$

By Proposition 5.8 with $U = 1$, we have

$$\frac{1}{N^2}\sum_{ijkl}\left(\frac{1}{d(d-1)^{1/2}}\mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}G_{ij}]\right) + \mathbb{E}[Q_{\mathfrak{a}}] \prec \frac{\mathbb{E}[\mathrm{Im}[m]]}{N\eta} + \frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N}. \quad (5.45)$$

The claim (5.3) follows from combining (5.44) and (5.45), and $\Lambda_{\mathrm{o}} \geqslant 1/\sqrt{d}$,

$$\mathbb{E}\left[1 + zm + Q_{\mathfrak{a}}(m)\right] = \mathrm{O}_{\prec}\left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{\mathbb{E}[\mathrm{Im}[m]]}{N\eta} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\right). \quad (5.46)$$

This finishes the proof of Proposition 5.1. □

6 Identification of the Self-consistent Equation

The algorithm that generates the polynomial $P_{\mathfrak{a}}$ from Proposition 5.1 is explicit but quite complicated, so that explicitly tracking the resulting coefficients of $P_{\mathfrak{a}}$ is a hopeless task beyond the first few orders. In this section we characterize these coefficients (asymptotically) as those of the power series $P_\infty(z,w)$ from (2.8), characterizing the Stieltjes transform m_d of the Kesten–McKay law.

Proposition 6.1. *Uniformly in $z \in \mathbb{C}_+$, the polynomial $P_{\mathfrak{a}}(z,w) = 1 + zw + Q_{\mathfrak{a}}(w)$ constructed in Proposition 5.1 satisfies*

$$P_{\mathfrak{a}}(z, m_d(z)) = \mathrm{O}(d^{-\mathfrak{a}/2}), \quad (6.1)$$

where m_d is the Stieltjes transform of the Kesten–McKay law, given by (2.5).

Corollary 6.2. *Let $P_{\mathfrak{a}}$ be the polynomial constructed in Proposition 5.1. Then $P_{\mathfrak{a}}(z,w) - P_\infty(z,w) = Q_{\mathfrak{a}}(w) - Q_\infty(w)$ is a power series in w, which converges on the whole complex plane. Each of its coefficients is of order $\mathrm{O}(d^{-\mathfrak{a}/2})$.*

Definition 6.3. *We write $P'_{\mathfrak{a}}(z,w) := \partial_w P_{\mathfrak{a}}(z,w)$, and similarly $P_{\mathfrak{a}}^{(k)}(z,w)$ for the k-th derivative in the variable w.*

Corollary 6.4. *The polynomial $P_{\mathfrak{a}}$ constructed in Proposition 5.1 satisfies*

$$\begin{aligned} &|P'_{\mathfrak{a}}(z, m_d(z))| \asymp \sqrt{|\kappa| + \eta} + \mathrm{O}(d^{-\mathfrak{a}/2}), \quad P''_{\mathfrak{a}}(z, m_d(z)) = 2 + \mathrm{O}(d^{-1/2}), \\ &P'''_{\mathfrak{a}}(z, m_d(z)) = \mathrm{O}(1), \end{aligned} \quad (6.2)$$

where $z = 2 + \kappa + \mathrm{i}\eta$ or $z = -2 - \kappa + \mathrm{i}\eta$ for $\eta \leqslant \mathfrak{K}, -2 \leqslant \kappa \leqslant \mathfrak{K}$, where the constant $\mathfrak{K}$ is from (2.10).

6.1 Proof of Proposition 6.1. The main ingredient of the proof of the proposition is the *ideal Green's function* $\widehat{G}$ introduced in the following definition.

DEFINITION 6.5. *For $z \in \mathbb{C}_+$ define the ideal Green's function $\widehat{G}(z) = (\widehat{G}_{ij}(z))_{i,j\in[N]}$ through*

$$\widehat{G}_{ij} := m_d\,\delta_{ij} - \frac{m_d m_{\mathrm{sc}}}{\sqrt{d-1}} A_{ij}. \tag{6.3}$$

Note that while Proposition 6.1 is deterministic, $\widehat{G}$ is random. However, we remark that when i and j have distance at most 1 in the random regular graph defined by A, the ideal Green's function $\widehat{G}_{ij}$ coincides with the Green's function of the infinite d-regular tree

$$G^{\mathrm{tree}}_{ij} = m_d\left(-\frac{m_{\mathrm{sc}}}{\sqrt{d-1}}\right)^{\mathrm{dist}(i,j)}$$

(see [BHY19, Proposition 5.1]) for vertices i and j with the same distance, while it is set to be 0 for all pairs of vertices i, j with greater distance. Because it agrees with the tree Green's function locally, the ideal Green's function is a random matrix that shares key algebraic properties with the true Green's function of the random graph, while its normalized trace is equal to the deterministic m_d. As a consequence, we shall show that $m_d = \frac{1}{N}\operatorname{Tr}\widehat{G}$ satisfies the same self-consistent equation as $m = \frac{1}{N}\operatorname{Tr} G$, up to small error terms, which will imply Proposition 6.1.

LEMMA 6.6. *The ideal Green's function $\widehat{G}$ has the following properties, uniformly in $z \in \mathbb{C}_+$.*

(i) *For any $i \in [\![N]\!]$, $|\widehat{G}_{ii}| = \mathrm{O}(1)$, and for $i \neq j$, $|\widehat{G}_{ij}| = \mathrm{O}((d-1)^{-1/2})$;*
(ii) $\sum_{i=1}^N \widehat{G}_{ii} = N m_d$;
(iii) $\sum_{i=1}^N |\widehat{G}_{ij}| = \sum_{j=1}^N |\widehat{G}_{ij}| = \mathrm{O}(d^{1/2})$;
(iv) $\sum_{i=1}^N |\widehat{G}_{ij}|^2 = \sum_{j=1}^N |\widehat{G}_{ij}|^2 = \mathrm{O}(1)$.
(v) *For any $i \in [\![N]\!]$,*

$$1 = \big((H-z)\widehat{G}\big)_{ii} = -z\widehat{G}_{ii} + \sum_{j=1}^N \frac{A_{ij}\widehat{G}_{ij}}{\sqrt{d-1}}. \tag{6.4}$$

Proof. Properties (i)–(iv) are immediate consequences of the definition and $m_d, m_{\mathrm{sc}} = \mathrm{O}(1)$. The identity (6.4) follows from (2.5). □

Crucially, under perturbation by a switching ξ^{kl}_{ij} of A, the ideal Green's function $\widehat{G}$ satisfies the same resolvent expansion as the real Green's function G. The intuitive reason behind this behaviour is the following. Take two edges, ij and kl, that are switchable. Then, in the limit where $\mathrm{dist}(ij, kl)$ tends to infinity, the Green's functions $\widehat{G}$ and G^{tree}, when restricted to the vertices $ijkl$, have identical behaviour under the switching ξ^{kl}_{ij}.

To state this precisely, we introduce some notation. For an $N \times N$ matrix $M = (M_{ij})_{i,j\in[\![N]\!]}$ and $V \subset [\![N]\!]$, we denote by $M|_V := (M_{ij})_{i,j\in V}$ the submatrix induced by the set V. In particular, $A|_V$ is the adjacency matrix of the graph A restricted to the vertex set V. We also frequently abbreviate subsets of $[\![N]\!]$ as $\{i,j\} \equiv ij$, and so on.

Claim 6.7. *Let i,j,k,l be distinct indices such that $A|_{ijkl} = \Delta_{ik}+\Delta_{jl}$ (i.e. $\chi_{ik}^{jl}(A) = 1$). Then*

$$\widehat{G}(A+\xi_{ij}^{kl})|_{ijkl} = \left(\left(\widehat{G}(A)|_{ijkl}\right)^{-1} + (d-1)^{-1/2}\xi_{ij}^{kl}|_{ijkl}\right)^{-1}. \tag{6.5}$$

Proof. By the definition of $\widehat{G}(A)$, we have

$$\widehat{G}(A)|_{ijkl} = \begin{bmatrix} m_d & 0 & \frac{-m_d m_{\mathrm{sc}}}{\sqrt{d-1}} & 0 \\ 0 & m_d & 0 & \frac{-m_d m_{\mathrm{sc}}}{\sqrt{d-1}} \\ \frac{-m_d m_{\mathrm{sc}}}{\sqrt{d-1}} & 0 & m_d & 0 \\ 0 & \frac{-m_d m_{\mathrm{sc}}}{\sqrt{d-1}} & 0 & m_d \end{bmatrix} = \begin{bmatrix} \frac{1}{m_{\mathrm{sc}}} & 0 & \frac{1}{\sqrt{d-1}} & 0 \\ 0 & \frac{1}{m_{\mathrm{sc}}} & 0 & \frac{1}{\sqrt{d-1}} \\ \frac{1}{\sqrt{d-1}} & 0 & \frac{1}{m_{\mathrm{sc}}} & 0 \\ 0 & \frac{1}{\sqrt{d-1}} & 0 & \frac{1}{m_{\mathrm{sc}}} \end{bmatrix}^{-1}, \tag{6.6}$$

where in the last equality we used the identity (2.9) and $1 + zm_{\mathrm{sc}} + m_{\mathrm{sc}}^2 = 0$. Analogously,

$$\begin{aligned} \widehat{G}(A+\xi_{ij}^{kl})|_{ijkl} &= \begin{bmatrix} \frac{1}{m_{\mathrm{sc}}} & \frac{1}{\sqrt{d-1}} & 0 & 0 \\ \frac{1}{\sqrt{d-1}} & \frac{1}{m_{\mathrm{sc}}} & 0 & 0 \\ 0 & 0 & \frac{1}{m_{\mathrm{sc}}} & \frac{1}{\sqrt{d-1}} \\ 0 & 0 & \frac{1}{\sqrt{d-1}} & \frac{1}{m_{\mathrm{sc}}} \end{bmatrix}^{-1} \\ &= \left(\left(\widehat{G}(A)|_{ijkl}\right)^{-1} + (d-1)^{-1/2}\xi_{ij}^{kl}|_{ijkl}\right)^{-1}, \end{aligned} \tag{6.7}$$

where the last equality follows from (6.6). This proves (6.5). □

By definition, the ideal Green's function of course has a trivial perturbation expansion under switchings since it is an affine linear function of A. However, when perturbed by the specific direction and magnitude of a switching, its behaviour can be written in a more complicated but more useful way, which matches precisely the corresponding resolvent expansion for switchings of the true Green's function. Indeed, as a consequence of Claim 6.7, we obtain the following resolvent expansion for the transformation by switching of the ideal Green's function.

Claim 6.8. *For $b,c \in \mathbb{N}$ consider tuples $\mathbf{i} = (i_1,\dots,i_b)$, $\mathbf{j} = (j_1,\dots,j_b)$, $\mathbf{k} = (k_1,\dots,k_b)$, $\mathbf{l} = (l_1,\dots,l_b)$, $\mathbf{m} = (m_1,\dots,m_c)$ such that the indices $\mathbf{ijklm}$ are distinct and $A|_{\mathbf{ijklm}} = A|_{\mathbf{m}} + \sum_{a=1}^{b}(\Delta_{i_a k_a} + \Delta_{j_a l_a})$. Then*

$$\left(\widehat{G}\left(A + \sum_{a=1}^{b}\xi_{i_a j_a}^{k_a l_a}\right) - \widehat{G}(A)\right)\Big|_{\mathbf{ijklm}} = \left(\sum_{n\geqslant 1}\frac{1}{(d-1)^{n/2}}\widehat{G}(A)\left(-\sum_{a=1}^{b}\xi_{i_a j_a}^{k_a l_a}\widehat{G}(A)\right)^n\right)\Big|_{\mathbf{ijklm}}. \tag{6.8}$$

Proof. We shall prove that

$$\widehat{G}\left(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}\right)\Big|_{\mathbf{ijklm}}=\left(\left(\widehat{G}(A)|_{\mathbf{ijklm}}\right)^{-1}+(d-1)^{-1/2}\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}|_{\mathbf{ijklm}}\right)^{-1}, \quad (6.9)$$

and the claim (6.8) then follows from the resolvent expansion, which converges for large enough d because $\widehat{G}_{ij}=\mathrm{O}(1)$ by Lemma 6.6. By assumption, $A|_{\mathbf{ijklm}}$ is a block matrix with $b+1$ blocks, indexed by $\{i_1,j_1,k_1,l_1\},\{i_2,j_2,k_2,l_2\},\cdots,\{i_b,j_b,k_b,l_b\}$ and $\{m_1,\cdots,m_c\}$. By our definition (6.3) of ideal Green's function $\widehat{G}$, both $\widehat{G}(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_b})|_{\mathbf{ijklm}}$ and $\widehat{G}(A)|_{\mathbf{ijklm}}$ have this same block structure. Thus both sides of (6.9) vanish except for the submatrix indexed by $\{i_1,j_1,k_1,l_1\},\{i_2,j_2,k_2,l_2\},\cdots,\{i_b,j_b,k_b,l_b\}$ and $\{m_1,\cdots,m_c\}$, and (6.9) follows from (6.5). □

Claim 6.8 says that the ideal Green's function $\widehat{G}$ has exactly the same behaviour under a switching $A\mapsto A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}$ as the true Green's function G. To make this precise, we need the following definition, which is to be compared with Definitions 4.4 and 5.3 .

Definition 6.9. (i) *For any polynomial F in r^2 abstract variables and $\mathbf{i}\in[\![N]\!]^r$, denote by $\widehat{F}_{\mathbf{i}}=F(\{\widehat{G}_{i_si_t}\}_{i_s,i_t=1}^{r})$ the corresponding polynomial in the ideal Green's function entries. Hence, $\widehat{F}_{\mathbf{i}}$ is obtained from $F_{\mathbf{i}}$ by replacing every factor G_{ij} in $F_{\mathbf{i}}$ by $\widehat{G}_{ij}$.*

(ii) *For a polynomial $\widehat{F}_{\mathbf{ijm}}$ in the ideal Green's function entries and $\mathfrak{o}\in\mathbb{N}$, we define $\widehat{\mathcal{T}}_{\mathfrak{o}}(F,1)$ as the right-hand side of (5.6) with $U=1$ and $F_{\mathbf{ijm}}$ replaced with $\widehat{F}_{\mathbf{ijm}}$.*

With this definition, we can rewrite the right-hand side of (6.8) as

$$\left(\sum_{n\geqslant 1}\frac{1}{n!(d-1)^{n/2}}\left(\sum_{a=1}^{b}\partial_{i_aj_a}^{k_al_a}\right)^{n}G(A)\right)\widehat{\ }\Big|_{\mathbf{ijklm}}$$

Thus, Claim 6.8 implies the following result, which is the analogue of Claim 4.7.

Claim 6.10. *Let* **ijklm** *satisfy the assumptions of Claim 6.8, and let $F_{\mathbf{ijm}}$ be a polynomial in the Green's function entries. Then for any positive integer $\mathfrak{a}\geqslant 1$ we have*

$$\widehat{F}_{\mathbf{ijm}}\left(A+\sum_{a=1}^{b}\xi_{i_aj_a}^{k_al_a}\right)-\widehat{F}_{\mathbf{ijm}}(A)=\left(\sum_{n=1}^{\mathfrak{a}-1}\frac{1}{n!(d-1)^{n/2}}\left(\sum_{a=1}^{b}\partial_{i_aj_a}^{k_al_a}\right)^{n}F_{\mathbf{ijm}}(A)\right)\widehat{\ }+\mathrm{O}\left(\frac{1}{d^{\mathfrak{a}/2}}\right).$$

The main ingredient in the proof of Proposition 6.1 is the following result, which is the analogue of Proposition 5.4 for the ideal Green's function.

PROPOSITION 6.11. *Fix $\mathfrak{o} \in \mathbb{N}$. Let F be a fixed monic monomial in $(2b+c)^2$ abstract variables, with degree $\deg(F)$, and let $\widehat{F}_{\mathbf{ijm}}$ be the corresponding monomial in the ideal Green's function entries from Definition 6.9(i). Recall the definition of χ_F from Definition 4.4. Then*

$$\begin{aligned} &\frac{1}{d^{\mathfrak{o}/2}} \frac{1}{N^{b+c} d^b} \sum_{\mathbf{m}} \sum_{\mathbf{ij}} \mathbb{E}\left[\prod_{a=1}^{b} A_{i_a j_a} \widehat{F}_{\mathbf{ijm}}\right] \\ &\quad = \frac{\chi_F \, m_d^{\deg(F)}}{d^{\mathfrak{o}/2}} + \widehat{\mathcal{T}}_{\mathfrak{o}+1,\deg(F)+1}(1) + \frac{1}{d^{\mathfrak{o}/2}} \mathrm{O}\left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N}\right), \end{aligned} \tag{6.10}$$

where the term $\widehat{\mathcal{T}}_{\mathfrak{o}+1,\deg(F)+1}(1)$ is equal to the corresponding term $\mathcal{T}_{\mathfrak{o}+1,\deg(F)+1}(1)$ from (5.8) with G replaced by $\widehat{G}$ (see Definition 6.9(ii)).

Proof. The proof involves repeating the proof of Proposition 5.4 with $U = 1$ almost verbatim, replacing Claim 4.7 and (5.16) with Claim 6.10 with (6.4) respectively.

More precisely, the proof of Proposition 5.4 relies only on the following ingredients: $\sum_i G_{ij} = 0$, Claim 5.5, Claim 5.6, and Claim 5.7. Each of these has the following analogue for the ideal Green's function $\widehat{G}$. We replace $\sum_i G_{ij} = 0$ with Lemma 6.6(iii). For Claim 5.5, we replace (5.9) with $U = 1$ by

$$\left| \frac{1}{N^b} \sum_{\mathbf{i}} \mathbb{E}\left[\widehat{F}_{\mathbf{i}}\right] \right| = \mathrm{O}\left(\frac{1}{N}\right),$$

as follows from Lemma 6.6(iv). For Claim 5.6, we replace (5.11) with $U = 1$ by

$$\begin{aligned} &\frac{1}{d^{\mathfrak{o}/2}} \frac{1}{N^{b+c} d^b} \sum_{\mathbf{m}} \sum_{\mathbf{ij}} \mathbb{E}\left[\prod_{a=1}^{b} A_{i_a j_a} \widehat{F}_{\mathbf{ijm}}\right] \\ &\quad = \frac{1}{d^{\mathfrak{o}/2}} \frac{1}{N^{2b+c}} \sum_{\mathbf{m}} \sum_{\mathbf{ij}} \mathbb{E}\left[\widehat{F}_{\mathbf{ijm}}\right] + \widehat{\mathcal{T}}_{\mathfrak{o}+1,\deg(F)+1}(1) + \frac{1}{d^{\mathfrak{o}/2}} \mathrm{O}\left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N}\right), \end{aligned}$$

where $\widehat{\mathcal{T}}_{\mathfrak{o}+1,\deg(F)+1}(1)$ is obtained from $\mathcal{T}_{\mathfrak{o}+1,\deg(F)+1}(1)$ in (5.11) by replacing G with $\widehat{G}$. Here we used Claim 6.10 instead of Claim 4.7. Finally, for Claim 5.7, we replace (5.14) with $U = 1$ by

$$\begin{aligned} &\frac{1}{d^{\mathfrak{o}/2}} \frac{1}{N^{2+c}} \sum_{ii'\mathbf{m}} \mathbb{E}[\widehat{G}_{ii} \widehat{F}_{i\mathbf{m}}] \\ &\quad = \frac{1}{d^{\mathfrak{o}/2}} \frac{1}{N^{2+c}} \sum_{ii'\mathbf{m}} \mathbb{E}[\widehat{G}_{i'i'} \widehat{F}_{i\mathbf{m}}] + \widehat{\mathcal{T}}_{\mathfrak{o}+1,\deg(F)+3}(1) + \frac{1}{d^{\mathfrak{o}/2}} \mathrm{O}\left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N}\right), \end{aligned}$$

where $\widehat{\mathcal{T}}_{\mathfrak{o}+1,\deg(F)+3}(1)$ is obtained from $\mathcal{T}_{\mathfrak{o}+1,\deg(F)+3}(1)$ in (5.14) by replacing G with $\widehat{G}$. Here we used Claim 6.10 and (6.4) instead of Claim 4.7 and (5.15), (5.16), respectively.

This concludes the proof. □

Proof of Proposition 6.1. The polynomial $P_{\mathfrak{a}}(z,w)$ in Proposition 5.1 was constructed by repeatedly applying Proposition 5.4. Using Proposition 6.11 instead, we can repeat its proof verbatim to obtain

$$P_{\mathfrak{a}}(z,m_d) = \mathrm{O}\left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N}\right). \tag{6.11}$$

Since the left-hand side does not depend on N (recall Proposition 5.1), taking the limit $N \to \infty$ yields the claim. □

6.2 Proof of Corollaries 6.2 and 6.4

Proof of Corollary 6.2. We define the power series

$$R(w) = P_{\mathfrak{a}}(z,w) - P_\infty(z,w). \tag{6.12}$$

It follows from Proposition 6.1 that for any $z \in \mathbb{C}$,

$$R(m_d(z)) = P_{\mathfrak{a}}(z,m_d(z)) = \mathrm{O}(d^{-\mathfrak{a}/2}). \tag{6.13}$$

Next, we remark that the disk $\{w \in \mathbb{C} : |w| \leqslant \frac{d-1}{d}\}$ lies in the image $m_d(\mathbb{C})$. Indeed, from (2.7) we find that the image $m_{\mathrm{sc}}(\mathbb{C})$ is the closed unit disk in $\mathbb{C}$, and the above claim then follows from the identity $m_d = \frac{(d-1)m_{\mathrm{sc}}}{d-1-m_{\mathrm{sc}}^2}$, which follows from (2.5) and (2.7). As a consequence, we have that for any w such that $|w| \leqslant \frac{d-1}{d}$,

$$R(w) = \mathrm{O}(d^{-\mathfrak{a}/2}). \tag{6.14}$$

For any fixed degree k, the coefficients of w^k in the infinite series $R(w)$ is given by

$$\frac{1}{2\pi \mathrm{i}} \oint_{|w|=(d-1)/d} \frac{R(w)}{w^{k+1}}\,\mathrm{d}w = \mathrm{O}(d^{-\mathfrak{a}/2}), \tag{6.15}$$

since $d \geqslant 2$. This finishes the proof. □

Proof of Corollary 6.4. By Corollary 6.2, we have

$$P_{\mathfrak{a}}^{(k)}(z,m_d(z)) = P_\infty^{(k)}(z,m_d(z)) + \mathrm{O}(d^{-\mathfrak{a}/2}), \quad k = 1,2. \tag{6.16}$$

Since $P_\infty(z,m_d(z)) = 0$ for $z \in \mathbb{C}_+$, by the chain rule,

$$0 = \partial_z P_\infty(z,m_d(z)) = m_d(z) + P_\infty'(z,m_d(z))\partial_z m_d(z). \tag{6.17}$$

Rearranging the above expression gives

$$P_\infty'(z,m_d(z)) = -\frac{m_d(z)}{\partial_z m_d(z)}, \tag{6.18}$$

so that an elementary analysis of (2.5) and (2.7) yields

$$\big|P_\infty'(z,m_d(z))\big| \asymp \sqrt{|\kappa| + \eta},$$

and the first relation of (6.2) follows from combining (6.16) and (6.18). The second relation of (6.2) follows from (6.16) and

$$P_\infty''(z,m_d(z)) = 2 + \mathrm{O}(d^{-1/2}). \tag{6.19}$$

This completes the proof. □

7 Moment Estimate for the Self-consistent Equation

In order to establish the self-consistent equation for m in the sense of high probability, in this section we derive a recursive moment estimate for the high moments of $P_{\mathfrak{a}}(z,m)$, where $P_{\mathfrak{a}}$ is the polynomial constructed in Proposition 5.1. In the next section, we establish eigenvalue rigidity estimates using a careful analysis of this recursive moment estimate and an iteration argument.

PROPOSITION 7.1. *Suppose that Assumption 4.1 holds, and that* $\Lambda_{\mathrm{d}} \geqslant \Lambda_{\mathrm{o}} \geqslant 1/\sqrt{d}$. *Let* $P_{\mathfrak{a}}(z,w)$ *be the polynomial constructed in Proposition 5.1. Fix* $r \in \mathbb{N}$. *Abbreviating* $P_{\mathfrak{a}} \equiv P_{\mathfrak{a}}(z,m(z))$, *we have for any* $z \in \mathbf{D}$,

$$\begin{aligned}
\mathbb{E}[|P_{\mathfrak{a}}|^{2r}] &\prec \frac{\Lambda_{\mathrm{d}}^2}{d^{1/2}}\mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right] + \mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{(N\eta)^2}|P_{\mathfrak{a}}|^{2r-2}\right] \\
&+ \frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|^2}{(N\eta)^3}|P_{\mathfrak{a}}|^{2r-3}\right] + \left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\right)\mathbb{E}[|P_{\mathfrak{a}}|^{2r-1}] \\
&+ \max_{1\leqslant s\leqslant 2r}\mathbb{E}\left[\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^s|P_{\mathfrak{a}}|^{2r-s}\right] \\
&+ \left(\frac{\Lambda_{\mathrm{d}}}{d} + \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}} + \frac{d\Lambda_{\mathrm{o}}}{N}\right)\max_{1\leqslant s\leqslant 2r-1}\mathbb{E}\left[\left(\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}\right)^s|P_{\mathfrak{a}}|^{2r-s-1}\right].
\end{aligned} \tag{7.1}$$

The rest of this section is devoted to the proof of Proposition 7.1. We begin by writing $P_{\mathfrak{a}}(z,m) = 1 + mz + Q_{\mathfrak{a}}(m)$ (see (5.1)) and then apply the identity (see (2.2))

$$1 + zm = \frac{1}{N}\sum_{ij} H_{ij}G_{ij} + \frac{1}{N}$$

to obtain

$$\begin{aligned}
\mathbb{E}[|P_{\mathfrak{a}}|^{2r}] &= \mathbb{E}[(1+zm)P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}] + \mathbb{E}[Q_{\mathfrak{a}}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}] \\
&= \frac{1}{(d-1)^{1/2}N}\sum_{ij}\mathbb{E}[A_{ij}G_{ij}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}] + \mathbb{E}[Q_{\mathfrak{a}}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}] + \mathrm{O}\left(\frac{\mathbb{E}[|P_{\mathfrak{a}}|^{2r-1}]}{N}\right).
\end{aligned} \tag{7.2}$$

For the first term on the second line of (7.2), we use Corollary 3.2 with the random variable $F = G_{ij}P_{\mathfrak{a}}^{r-1}P_{\mathfrak{a}}^{r}$ and note that for $i \neq j$

$$|F(A)| + \max_{kl}|F(A+\xi_{ij}^{kl})| \prec \Lambda_{\mathrm{o}}\mathcal{C}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}, A),$$

where we recall the definition (3.10) and we used Remark 4.3. This yields

$$\begin{aligned}\frac{1}{(d-1)^{1/2}N}\sum_{ij}\mathbb{E}[A_{ij}G_{ij}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}] &= \frac{d}{(d-1)^{1/2}N^2}\sum_{i\neq j}\mathbb{E}[G_{ij}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}] \\ &\quad+\frac{1}{N^2d(d-1)^{1/2}}\sum_{i\neq j}\sum_{kl}\mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}(G_{ij}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})]+\mathrm{O}_{\prec}\left(\frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\mathbb{E}[\mathcal{C}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r},A)]\right) \\ &= \frac{1}{N^2d(d-1)^{1/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}(G_{ij}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})]+\mathrm{O}_{\prec}\left(\frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\mathbb{E}[\mathcal{C}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r},A)]\right),\end{aligned} \tag{7.3}$$

where in the last equality we used $\Lambda_{\mathrm{o}}\geqslant 1/\sqrt{d}$. To estimate the error term, we remark that for $U=P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}$ we have

$$\max_{s+\bar{s}\geqslant 1}\mathbb{E}\left[|U^{(s,\bar{s})}|\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^{s+\bar{s}}\right]\prec\max_{1\leqslant s\leqslant 2r-1}\mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta}+\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s|P_{\mathfrak{a}}|^{2r-1-s}\right], \tag{7.4}$$

as can be seen after some elementary algebra. Using (4.24) for $U=P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}$ we get

$$\mathcal{C}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r},A)\prec|P_{\mathfrak{a}}|^{2r-1}+\frac{1}{\sqrt{d}}\max_{1\leqslant s\leqslant 2r-1}\mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta}+\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s|P_{\mathfrak{a}}|^{2r-1-s}\right]. \tag{7.5}$$

By the discrete product rule (4.6), the first term in the last line of (7.3) equals

$$\begin{aligned}\mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}(G_{ij}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})] &= \mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}(G_{ij})P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}] \\ &\quad+\mathbb{E}[A_{ik}A_{jl}G_{ij}D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})]+\mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}(G_{ij})D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})].\end{aligned} \tag{7.6}$$

At this point we note the crucial cancellation, by Proposition 5.8, of the first term on the right-hand side of (7.6) with $\mathbb{E}[Q_{\mathfrak{a}}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}]$ from (7.2). Indeed, by Proposition 5.8 for $U=P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}$, recalling (7.4) and (7.5), we find

$$\begin{aligned}&\frac{1}{N^2d(d-1)^{1/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}(G_{ij})P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}]+\mathbb{E}[Q_{\mathfrak{a}}P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}] \\ &\quad\prec\mathbb{E}\left[\left(\frac{1}{d^{\mathfrak{a}/2}}+\frac{\operatorname{Im}[m]}{N\eta}+\frac{d}{N}\right)|P_{\mathfrak{a}}|^{2r-1}\right] \\ &\qquad+\left(\frac{d^{1/2}}{N}+\frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\right)\max_{1\leqslant s\leqslant 2r-1}\mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta}+\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s|P_{\mathfrak{a}}|^{2r-1-s}\right].\end{aligned} \tag{7.7}$$

What remains, therefore, is to estimate the contributions of the second and third terms on the right-hand side of (7.6). For the second term on the right-hand side of (7.6), we claim that

$$\begin{aligned}\frac{1}{N^2 d(d-1)^{1/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}G_{ij}D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})] &\prec \mathbb{E}\left[\frac{\Lambda_{\mathrm{o}}\Lambda_{\mathrm{d}}}{d^{1/2}}\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right]\\ &+\mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{(N\eta)^2}|P_{\mathfrak{a}}|^{2r-2}\right]+\frac{\Lambda_{\mathrm{o}}}{d}\max_{1\leqslant s\leqslant 2r-1}\mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta}+\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s|P_{\mathfrak{a}}|^{2r-1-s}\right]\\ &+\frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\mathbb{E}\left[\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2|P_{\mathfrak{a}}|^{2r-2}\right]+\frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|^2}{(N\eta)^3}|P_{\mathfrak{a}}|^{2r-3}\right].\end{aligned}\tag{7.8}$$

Essentially, this estimate will arise from the three off-diagonal Green's function entries obtained from G_{ij} and the derivative on $P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}$.

For the third term on the right-hand side of (7.6), we claim that

$$\begin{aligned}\frac{1}{N^2 d(d-1)^{1/2}}\mathbb{E}[A_{ik}A_{jl}D_{ij}^{kl}(G_{ij})D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})] &\prec \frac{\Lambda_{\mathrm{d}}^2}{d^{1/2}}\mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right]\\ &+\frac{d}{N}\mathbb{E}[|P_{\mathfrak{a}}|^{2r-1}]\\ &+\left(\frac{d^{1/2}}{N}+\frac{\Lambda_{\mathrm{d}}}{d}+\frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\right)\max_{1\leqslant s\leqslant 2r-1}\mathbb{E}\left[\left(\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}+\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s|P_{\mathfrak{a}}|^{2r-s-1}\right]\\ &+\mathbb{E}\left[\frac{\operatorname{Im}[m]}{N\eta}|P_{\mathfrak{a}}|^{2r-1}\right].\end{aligned}\tag{7.9}$$

This estimate will arise from certain special cancellations arising from the d-regular graph structure.

Proposition (7.1) follows from combining the estimates (7.2), (7.3), (7.4), (7.6), (7.7), (7.8) and (7.9). In the remainder of this section, we prove these estimates (7.8) and (7.9).

7.1 Proof of (7.8). By (4.7), left-hand side of (7.8) can be written as

$$\begin{aligned}&\frac{\mathrm{O}(1)}{N^2 d^{3/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}G_{ij}D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})]\\ &\quad=\sum_{n=1}^{\mathfrak{b}-1}\frac{\mathrm{O}(1)}{N^2 d^{(n+3)/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}G_{ij}(\partial_{ij}^{kl})^n(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})]\\ &\quad\quad+\frac{\mathrm{O}(1)}{N^2 d^{(\mathfrak{b}+3)/2}}\sum_{ijkl}\mathbb{E}\left[A_{ik}A_{jl}G_{ij}\left((\partial_{ij}^{kl})^{\mathfrak{b}}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})(A+\theta\xi_{ij}^{kl})\right)\right],\end{aligned}\tag{7.10}$$

for some random $\theta\in[0,1]$. In fact, for the terms corresponding to $n\geqslant 3$ in (7.10), we have the following simple estimate.

Claim 7.2. *For the terms in* (7.10) *with* $n \geqslant 3$,

$$\begin{aligned} &\frac{1}{N^2 d^{(n+3)/2}} \sum_{ijkl} \mathbb{E}[A_{ik}A_{jl}G_{ij}(\partial_{ij}^{kl})^n (P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})] \\ &\quad \prec \frac{\Lambda_{\mathrm{o}}}{d^{(n-1)/2}} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right]. \end{aligned} \tag{7.11}$$

Proof. Thanks to (4.15) and chain rule, we have

$$(\partial_{ij}^{kl})^n (P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}) \prec \max_{1\leqslant s\leqslant 2r-1} \left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}. \tag{7.12}$$

Using $\max_{i\neq j}|G_{ij}| \prec \Lambda_{\mathrm{o}}$ and $\max_i |G_{ii}| \prec 1$, it leads to

$$\begin{aligned} &\frac{1}{N^2 d^{(n+3)/2}} \sum_{ijkl} \mathbb{E}[A_{ik}A_{jl}G_{ij}(\partial_{ij}^{kl})^n (P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})] \\ &\quad \prec \frac{1}{N^2 d^{(n+3)/2}} \left(\max_{1\leqslant s\leqslant 2r-1} \sum_{i\neq j}\sum_{kl} \mathbb{E}\left[A_{ik}A_{jl}\Lambda_{\mathrm{o}}\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right] \right. \\ &\qquad \left. + \max_{1\leqslant s\leqslant 2r-1} \sum_{ikl} \mathbb{E}\left[A_{ik}A_{il}\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right]\right) \\ &\quad \prec \frac{\Lambda_{\mathrm{o}}}{d^{(n-1)/2}} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right], \end{aligned} \tag{7.13}$$

where in the last inequality we used that $\sum_{i\neq jkl} A_{ik}A_{jl} = N^2d^2$ and $\sum_{ikl} A_{ik}A_{il} = Nd^2$ and $\Lambda_{\mathrm{o}} \geqslant 1/N$. □

Moreover, by choosing $\mathfrak{b}$ large enough, depending on r, it follows that the second line of (7.10) is bounded by N^{-2r}.

In the following, we estimate the terms on the right-hand side of (7.10), corresponding to $n = 1$ and $n = 2$.

Claim 7.3. *For the term in* (7.10) *with* $n = 1$,

$$\frac{1}{N^2 d^2} \sum_{ijkl} \mathbb{E}[A_{ik}A_{jl}G_{ij}\partial_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})] \prec \mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{(N\eta)^2}|P_{\mathfrak{a}}|^{2r-2}\right]. \tag{7.14}$$

Proof. For the derivative $\partial_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})$, we have

$$\partial_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}) = (r-1)\partial_{ij}^{kl} m P_{\mathfrak{a}}' P_{\mathfrak{a}}^{r-2}\bar{P}_{\mathfrak{a}}^{r} + (\cdots), \tag{7.15}$$

where $(\cdots)$ denotes analogous terms with complex conjugates obtained by applying the derivatives to $\bar{P}_{\mathfrak{a}}$ instead of $P_{\mathfrak{a}}$. We estimate the error from the first term, for which we can first sum over the indices i, j, k, l. By (4.9),

$$\frac{1}{N^2d^2}\sum_{ijkl}A_{ik}A_{jl}G_{ij}\partial_{ij}^{kl}m=\frac{2}{N^3d^2}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(-(G^2)_{ij}-(G^2)_{kl}+(G^2)_{ik}+(G^2)_{jl}). \tag{7.16}$$

There are four terms on the right-hand side of (7.16). For the first term, using (4.11),

$$\frac{1}{N^3d^2}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(G^2)_{ij}=\frac{1}{N^3}\sum_{ij}G_{ij}(G^2)_{ij}=\frac{1}{N^3}\operatorname{Tr}G^3\prec\frac{\operatorname{Im}[m]}{(N\eta)^2}. \tag{7.17}$$

For the second term on the right-hand side of (7.16), using that $(H-z)G=P_\perp$ and (4.11),

$$\begin{aligned}\frac{1}{N^3d^2}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(G^2)_{kl}&=\frac{1}{N^3d^2}\operatorname{Tr}(GAGAG)\\&=\frac{1}{N^3d^2}\operatorname{Tr}(G(A-\sqrt{d-1}z)G(A-\sqrt{d-1}z)G)\\&\quad+\frac{2}{N^3d^2}\operatorname{Tr}(G(\sqrt{d-1}z)G(A-\sqrt{d-1}z)G)\\&\quad+\frac{1}{N^3d^2}\operatorname{Tr}(G(\sqrt{d-1}z)G(\sqrt{d-1}z)G)\\&=\frac{\mathrm{O}(1)}{N^3d}\operatorname{Tr}G+\frac{\mathrm{O}(1)}{N^3d}\operatorname{Tr}G^2+\frac{\mathrm{O}(1)}{N^3d}\operatorname{Tr}G^3=\mathrm{O}_\prec\left(\frac{\operatorname{Im}[m]}{d(N\eta)^2}\right).\end{aligned} \tag{7.18}$$

For the last two terms on the right-hand side of (7.16), since $\sum_j G_{ij}=0$,

$$\begin{aligned}\frac{1}{N^3d^2}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(G^2)_{ik}&=\frac{1}{N^3d}\sum_{ijk}A_{ik}G_{ij}(G^2)_{ik}=0,\\\frac{1}{N^3d^2}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(G^2)_{jl}&=\frac{1}{N^3d}\sum_{ijl}A_{jl}G_{ij}(G^2)_{jl}=0.\end{aligned} \tag{7.19}$$

By combining expressions (7.17), (7.18) and (7.19), we obtain the estimate

$$\frac{1}{N^2d^2}\sum_{ijkl}A_{ik}A_{jl}G_{ij}\partial_{ij}^{kl}m\prec\frac{\operatorname{Im}[m]}{(N\eta)^2}, \tag{7.20}$$

for (7.16), and the Claim 7.3 follows. □

Claim 7.4. *For the term in (7.10) with $n=2$,*

$$\begin{aligned}\frac{1}{N^2d^{5/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}G_{ij}(\partial_{ij}^{kl})^2(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})]&\prec\mathbb{E}\left[\frac{\Lambda_{\mathrm{o}}\Lambda_{\mathrm{d}}}{d^{1/2}}\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right]\\&+\frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\mathbb{E}\left[\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2|P_{\mathfrak{a}}|^{2r-2}\right]+\frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|^2}{(N\eta)^3}|P_{\mathfrak{a}}|^{2r-3}\right]\\&+\frac{\Lambda_{\mathrm{o}}}{d}\mathbb{E}\left[\left(\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}\right)^2|P_{\mathfrak{a}}|^{2r-3}\right].\end{aligned} \tag{7.21}$$

Proof. For the derivative $\partial_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})$, we have

$$\begin{aligned}(\partial_{ij}^{kl})^2(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}) = ((\partial_{ij}^{kl})^2 m)P_{\mathfrak{a}}' P_{\mathfrak{a}}^{r-2}\bar{P}_{\mathfrak{a}}^{r} + (\partial_{ij}^{kl}m)^2 P_{\mathfrak{a}}'' P_{\mathfrak{a}}^{r-2}\bar{P}_{\mathfrak{a}}^{r} \\ +((\partial_{ij}^{kl}m)P_{\mathfrak{a}}')^2 P_{\mathfrak{a}}^{r-3}\bar{P}_{\mathfrak{a}}^{r} + (\cdots),\end{aligned} \tag{7.22}$$

where $(\cdots)$ denotes analogous terms with complex conjugates obtained by applying the derivatives to $\bar{P}_{\mathfrak{a}}$ instead of $P_{\mathfrak{a}}$. We consider the terms separately.

For the first term in (7.22), we use the explicit formula

$$\begin{aligned}(\partial_{ij}^{kl})^2 m &= \frac{2}{N}\sum_{a=1}^{N}\partial_{ij}^{kl}(-G_{ia}G_{ja} - G_{ka}G_{la} + G_{ia}G_{ka} + G_{ja}G_{la}) \\ &= \frac{2}{N}\sum_{a=1}^{N}(G_{ii}G_{ja}^2 + \cdots),\end{aligned} \tag{7.23}$$

where $\cdots$ denotes 31 other terms obtained by applying the product rule for differentiation for ∂_{ij}^{kl}. Using that $\sum_j G_{ij} = 0$ and (4.11), the first term gives

$$\begin{aligned}\frac{1}{N^3 d^{5/2}}\sum_{ijkla} A_{ik}A_{jl}G_{ij}G_{ii}G_{ja}^2 &= \frac{1}{N^3 d^{1/2}}\sum_{ij} G_{ij}G_{ii}(G^2)_{jj}, \\ = \frac{1}{N^3 d^{1/2}}\sum_{ij} G_{ij}(G_{ii}-m)(G^2)_{jj} &\prec \frac{\Lambda_{\mathrm{o}}\Lambda_{\mathrm{d}}}{d^{1/2}}\frac{\operatorname{Im}[m]}{N\eta}.\end{aligned} \tag{7.24}$$

An analogous calculation can be performed for all terms on the right-hand side of (7.23). Indeed, every such term has three factors of G, exactly two of which have an index a, and the third remaining factor is either diagonal (in which case the same argument as above applies) or off-diagonal (in which case we gain Λ_{o} instead of Λ_{d}). Since $\Lambda_{\mathrm{d}} \geqslant \Lambda_{\mathrm{o}}$, it leads to the estimate

$$\frac{1}{N^2 d^{5/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}G_{ij}((\partial_{ij}^{kl})^2 m)P_{\mathfrak{a}}' P_{\mathfrak{a}}^{r-2}\bar{P}_{\mathfrak{a}}^{r}] \prec \mathbb{E}\left[\frac{\Lambda_{\mathrm{o}}\Lambda_{\mathrm{d}}}{d^{1/2}}\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right]. \tag{7.25}$$

For the second term in (7.22), we can directly apply (4.14) to get the bound

$$\frac{1}{N^2 d^{5/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}G_{ij}(\partial_{ij}^{kl}m)^2 P_{\mathfrak{a}}'' P_{\mathfrak{a}}^{r-2}\bar{P}_{\mathfrak{a}}^{r}] \prec \frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\mathbb{E}\left[\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2|P_{\mathfrak{a}}''||P_{\mathfrak{a}}|^{2r-2}\right], \tag{7.26}$$

which is enough by $P_{\mathfrak{a}}'' \prec 1$.

For the third term in (7.22), we have $(\partial_{ij}^{kl}m)^2 = 4((G^2)_{ij} + (G^2)_{kl} - (G^2)_{ik} - (G^2)_{jl})^2/N^2$. There are ten different terms, which we estimate one by one. For the term $(G^2)_{ij}(G^2)_{ij}$, we use (4.11) to get

$$\begin{aligned}\frac{1}{N^4d^{5/2}}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(G^2)_{ij}(G^2)_{ij} &= \frac{1}{N^4d^{1/2}}\sum_{ij}G_{ij}(G^2)_{ij}(G^2)_{ij}\\ &\prec \frac{1}{N^4d^{1/2}}\sum_{ij}\Lambda_{\mathrm{o}}\left|(G^2)_{ij}\right|^2+\frac{1}{N^4d^{1/2}}\sum_{i}\left|(G^2)_{ii}\right|^2\\ &= \frac{1}{N^4d^{1/2}}\Lambda_{\mathrm{o}}\operatorname{Tr}|G|^4+\frac{1}{Nd^{1/2}}\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\prec\frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\frac{\operatorname{Im}[m]}{(N\eta)^3}+\frac{1}{Nd^{1/2}}\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2.\end{aligned}\tag{7.27}$$

For the term $(G^2)_{ij}(G^2)_{kl}$, we use (4.11) and $(H-z)G=P_\perp$ to get

$$\begin{aligned}\frac{1}{N^4d^{5/2}}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(G^2)_{ij}(G^2)_{kl} &= \frac{1}{N^4d^{5/2}}\sum_{ij}G_{ij}(G^2)_{ij}(AG^2A)_{ij}\\ &= \frac{d-1}{N^4d^{5/2}}\sum_{ij}G_{ij}(G^2)_{ij}(z^2G^2+2zG+P_\perp)_{ij}\\ &\prec \frac{1}{N^2d^{3/2}}\sum_{ij}|G_{ij}|\frac{\operatorname{Im}[m]^2}{(N\eta)^2}\prec\frac{\Lambda_{\mathrm{o}}}{d^{3/2}}\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2,\end{aligned}\tag{7.28}$$

where in the second equality we used that $AG^2A=(d-1)(H-z+z)G^2(H-z+z)=(d-1)(z^2G^2+2zG+P_\perp)$. For the term $(G^2)_{ij}(G^2)_{ik}$, we use (4.11) and $(H-z)G=P_\perp$ to get

$$\begin{aligned}\frac{1}{N^4d^{5/2}}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(G^2)_{ij}(G^2)_{ik} &= \frac{1}{N^4d^{3/2}}\sum_{ij}G_{ij}(G^2)_{ij}(AG^2)_{ii}\\ &= \frac{(d-1)^{1/2}}{N^4d^{3/2}}\sum_{ij}G_{ij}(G^2)_{ij}(zG^2+G)_{ii}\\ &\prec \frac{1}{N^2d}\sum_{ij}|G_{ij}|\frac{\operatorname{Im}[m]^2}{(N\eta)^2}\prec\frac{\Lambda_{\mathrm{o}}}{d}\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2,\end{aligned}\tag{7.29}$$

where in the second equality we used that $AG^2=(d-1)^{1/2}(H-z+z)G^2=(d-1)^2(zG^2+G)$. The term $(G^2)_{ij}(G^2)_{jl}$ can be estimated in the same way. For the term $(G^2)_{kl}(G^2)_{kl}$, we use (4.11) and $(H-z)G=P_\perp$ to get

$$\begin{aligned}\frac{1}{N^4d^{5/2}}\sum_{ijkl}A_{ik}A_{jl}G_{ij}(G^2)_{kl}(G^2)_{kl} &= \frac{1}{N^4d^{5/2}}\sum_{kl}(AGA)_{kl}(G^2)_{kl}(G^2)_{kl}\\ &= \frac{d-1}{N^4d^{5/2}}\sum_{kl}(z^2G+zP_\perp+HP_\perp)_{kl}(G^2)_{kl}(G^2)_{kl}\\ &\prec \frac{1}{N^2d^{3/2}}\sum_{kl}|(z^2G+zP_\perp+HP_\perp)_{kl}|\frac{\operatorname{Im}[m]^2}{(N\eta)^2}\prec\frac{\Lambda_{\mathrm{o}}}{d^{3/2}}\left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2,\end{aligned}\tag{7.30}$$

where in the second equality we used that $AGA = (d-1)(H-z+z)G(H-z+z) = (d-1)(z^2G + zP_\perp + HP_\perp)$, and in the last inequality, we used $\Lambda_o \geqslant 1/\sqrt{d}$ and $\sum_{kl} |(z^2G + zP_\perp + HP_\perp)_{kl}| \prec N^2\Lambda_o + d^{1/2}N \leqslant N^2\Lambda_o$. For the term $(G^2)_{kl}(G^2)_{ik}$, we use (4.11) and $(H-z)G = P_\perp$ to get

$$\begin{aligned} \frac{1}{N^4 d^{5/2}} \sum_{ijkl} A_{ik}A_{jl}G_{ij}(G^2)_{kl}(G^2)_{ik} &= \frac{1}{N^4 d^{5/2}} \sum_{ikl} A_{ik}(AG)_{il}(G^2)_{kl}(G^2)_{ik} \\ &= \frac{(d-1)^{1/2}}{N^4 d^{5/2}} \sum_{ikl} A_{ik}(zG + P_\perp)_{il}(G^2)_{kl}(G^2)_{ik} \\ &\prec \frac{1}{N^2 d^2} \sum_{ikl} A_{ik}|(zG + P_\perp)_{il}| \frac{\mathrm{Im}[m]^2}{(N\eta)^2} \prec \frac{\Lambda_o}{d} \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^2, \end{aligned} \tag{7.31}$$

where in the second line we used $AG = (d-1)^{1/2}(H-z+z)G = (d-1)^{1/2}(zG+P_\perp)$, and in the last inequality we used $\sum_l |(zG + P_\perp)_{il}| \leqslant N\Lambda_o$. The term $(G^2)_{kl}(G^2)_{jl}$ can be estimated in the same way. For the term $(G^2)_{ik}(G^2)_{ik}$, we use that $\sum_j G_{ij} = 0$ to get

$$\frac{1}{N^4 d^{5/2}} \sum_{ijkl} A_{ik}A_{jl}G_{ij}(G^2)_{ik}(G^2)_{ik} = \frac{1}{N^4 d^{3/2}} \sum_{ijk} A_{ik}G_{ij}(G^2)_{ik}(G^2)_{ik} = 0. \tag{7.32}$$

The term $(G^2)_{jl}(G^2)_{jl}$ can be estimated in the same way. For the term $(G^2)_{ik}(G^2)_{jl}$, we use (4.11) and $(H-z)G = P_\perp$ to get

$$\begin{aligned} \frac{1}{N^4 d^{5/2}} \sum_{ijkl} A_{ik}A_{jl}G_{ij}(G^2)_{ik}(G^2)_{jl} &= \frac{1}{N^4 d^{5/2}} \sum_{ij} G_{ij}(AG^2)_{ii}(AG^2)_{jj} \\ &= \frac{(d-1)}{N^4 d^{5/2}} \sum_{ij} G_{ij}(zG^2 + G)_{ii}(zG^2 + G)_{jj} \\ &\prec \frac{1}{N^2 d^{3/2}} \sum_{ij} |G_{ij}| \frac{\mathrm{Im}[m]^2}{(N\eta)^2} \prec \frac{\Lambda_o}{d^{3/2}} \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^2. \end{aligned} \tag{7.33}$$

where in the second equality, we used $AG^2 = (d-1)^{1/2}(H-z+z)G^2 = (d-1)^{1/2}(zG^2+G)$.

We combine the above estimates together, and find that the third term in (7.22) is bounded by

$$\begin{aligned} &\frac{1}{N^2 d^{5/2}} \sum_{ijkl} \mathbb{E}[A_{ik}A_{jl}G_{ij}((\partial_{ij}^{kl} m)P'_{\mathfrak{a}})^2 P_{\mathfrak{a}}^{r-3} \bar{P}_{\mathfrak{a}}^r] \\ &\prec \frac{\Lambda_o}{d^{1/2}} \mathbb{E}\left[\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|^2}{(N\eta)^3} |P_{\mathfrak{a}}|^{2r-3}\right] + \frac{\Lambda_o}{d} \mathbb{E}\left[\left(\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|}{N\eta}\right)^2 |P_{\mathfrak{a}}|^{2r-3}\right]. \end{aligned} \tag{7.34}$$

The Claim 7.4 follows from combining (7.25), (7.26) and (7.34). □

7.2 Proof of (7.9). By (4.7), the left-hand side of (7.9) can be written as

$$\begin{aligned}
&\frac{1}{N^2 d(d-1)^{1/2}} \sum_{ijkl} \mathbb{E}[A_{ik}A_{jl} D_{ij}^{kl}(G_{ij}) D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})] \\
&\quad = \sum_{n_1=1}^{\mathfrak{b}_1-1} \frac{1}{n_1! N^2 d(d-1)^{(n_1+1)/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl} (\partial_{ij}^{kl})^{n_1}(G_{ij}) D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\
&\qquad + \frac{1}{\mathfrak{b}_1! N^2 d(d-1)^{(\mathfrak{b}_1+1)/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}\Big((\partial_{ij}^{kl})^{\mathfrak{b}_1}(G_{ij})(A+\theta\xi_{ij}^{kl})\Big) D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right],
\end{aligned} \tag{7.35}$$

for some random $\theta \in [0,1]$. As above, by choosing $\mathfrak{b}_1$ large enough, depending on r, we find that the last line of (7.35) is bounded by N^{-2r}. Moreover, for terms corresponding to $n_1 \geqslant 3$ in (7.35), we have the following simple estimate.

Claim 7.5. *For the terms $n_1 \geqslant 3$ on the right-hand side of (7.35),*

$$\begin{aligned}
&\frac{1}{N^2 d(d-1)^{(n_1+1)/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl} (\partial_{ij}^{kl})^{n_1}(G_{ij}) D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\
&\quad \prec \frac{1}{d^{n_1/2}} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right].
\end{aligned} \tag{7.36}$$

Proof. Thanks to (4.7) and (7.12), we have

$$D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}) \prec \max_{1\leqslant s\leqslant 2r-1} \left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s \frac{|P_{\mathfrak{a}}|^{2r-1-s}}{d^{1/2}}. \tag{7.37}$$

Combining with $(\partial_{ij}^{kl})^{n_1} G_{ij} \prec 1$ we obtain

$$\begin{aligned}
&\frac{1}{N^2 d(d-1)^{(n_1+1)/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl} (\partial_{ij}^{kl})^{n_1}(G_{ij}) D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\
&\quad \prec \frac{1}{N^2 d^{(n_1+4)/2}} \max_{1\leqslant s\leqslant 2r-1} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right] \\
&\quad = \frac{1}{d^{n_1/2}} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right].
\end{aligned}$$

□

In the following, we estimate the terms on the right-hand side of (7.35), corresponding to $n_1 = 1$ and $n_1 = 2$.

CLAIM 7.6. *For the term $n_1 = 1$ on the right-hand side of (7.35),*

$$\begin{aligned}\frac{1}{N^2 d(d-1)} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}\partial_{ij}^{kl} G_{ij} D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] &\prec \frac{\Lambda_{\mathrm{d}}^2}{d^{1/2}} \mathbb{E}\left[\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right] \\ &+ \frac{d}{N}\mathbb{E}[|P_{\mathfrak{a}}|^{2r-1}] \\ &+ \left(\frac{d^{1/2}}{N} + \frac{\Lambda_{\mathrm{d}}}{d} + \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\right) \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{\operatorname{Im}[m]|P_{\mathfrak{a}}'|}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-s-1}\right] \\ &+ \mathbb{E}\left[\frac{\operatorname{Im}[m]}{N\eta}|P_{\mathfrak{a}}|^{2r-1}\right].\end{aligned} \tag{7.38}$$

Proof. The derivative $\partial_{ij}^{kl} G_{ij}$ is given by

$$\begin{aligned}\partial_{ij}^{kl} G_{ij} &= -G_{ii}G_{jj} + G_{il}G_{jj} + G_{ii}G_{kj} - G_{ij}G_{ij} - G_{ik}G_{lj} - G_{il}G_{kj} \\ &\quad + G_{ik}G_{ij} + G_{ij}G_{lj}.\end{aligned} \tag{7.39}$$

We shall show that the biggest term is $-G_{ii}G_{jj}$, and that the other terms are smaller. We first estimate those terms in (7.39) which contain two off-diagonal terms, i.e., $-G_{ij}G_{ij} - G_{ik}G_{lj} - G_{il}G_{kj} + G_{ik}G_{ij} + G_{ij}G_{lj}$. For the term $G_{ij}G_{ij}$, using (7.37) to bound $D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})$,

$$\begin{aligned}&\frac{1}{N^2 d(d-1)} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{ij}G_{ij} D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\ &\prec \frac{1}{N^2 d(d-1)} \left(\sum_{i\neq j}\sum_{kl} \mathbb{E}\left[A_{ik}A_{jl}\Lambda_{\mathrm{o}}^2 |D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})|\right] + \sum_{ikl} \mathbb{E}\left[A_{ik}A_{il}|D_{ii}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})|\right]\right) \\ &\prec \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right].\end{aligned} \tag{7.40}$$

The same argument applies to the other terms with two off-diagonal indices.

We next estimate those terms in (7.39), which contain exactly one off-diagonal term, i.e., $G_{il}G_{jj} + G_{ii}G_{kj}$. For the term $G_{il}G_{jj}$,

$$\begin{aligned}&\frac{1}{N^2 d(d-1)} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj} D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\ &= \sum_{n_2=1}^{\mathfrak{b}_2-1} \frac{1}{n_2! N^2 d(d-1)^{1+n_2/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj}(\partial_{ij}^{kl})^{n_2}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\ &\quad + \frac{1}{\mathfrak{b}_2! N^2 d(d-1)^{1+\mathfrak{b}_2/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj}\left((\partial_{ij}^{kl})^{\mathfrak{b}_2}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})(A+\theta\xi_{ij}^{kl})\right)\right],\end{aligned} \tag{7.41}$$

for some random $\theta \in [0,1]$. As above, by choosing $\mathfrak{b}_2$ large enough, depending on r, we find that the last line of (7.41) is bounded by N^{-2r}. Moreover, for terms corresponding to $n_2 \geqslant 2$ in (7.35), by the same argument as (7.40) and $\max_{i\neq l}|G_{il}| \prec \Lambda_{\mathrm{o}}$, we have the simple estimate

$$\begin{aligned} &\frac{1}{N^2 d(d-1)^{1+n_2/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj}(\partial_{ij}^{kl})^{n_2}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\ &\quad \prec \frac{\Lambda_{\mathrm{o}}}{d^{n_2/2}} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right]. \end{aligned} \tag{7.42}$$

Using the above estimate in (7.41), we get

$$\begin{aligned} &\frac{1}{N^2 d(d-1)} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj}D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\ &\quad = \frac{1}{N^2 d(d-1)^{3/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj}\partial_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\ &\qquad + \mathrm{O}_{\prec}\left(\frac{\Lambda_{\mathrm{o}}}{d} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right]\right). \end{aligned} \tag{7.43}$$

The first term on the right-hand side of (7.43), arising from $n_2 = 1$, needs to be estimated more precisely. It can be written as

$$\begin{aligned} &\frac{1}{N^2 d(d-1)^{3/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj}\partial_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\ &\quad = \frac{\mathrm{O}(1)}{N^2 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj}\left((r-1)\partial_{ij}^{kl}mP_{\mathfrak{a}}'P_{\mathfrak{a}}^{r-2}\bar{P}_{\mathfrak{a}}^{r} + r\partial_{ij}^{kl}\bar{m}\bar{P}_{\mathfrak{a}}'|P_{\mathfrak{a}}|^{2r-2}\right)\right]. \end{aligned} \tag{7.44}$$

We estimate $\sum_{ikl} A_{ik}A_{jl}G_{il}\partial_{ij}^{kl}m$, and the other term $\sum_{ikl} A_{ik}A_{jl}G_{il}\partial_{ij}^{kl}\bar{m}$ can be estimated in the same way. Recalling (4.9), we have

$$\partial_{ij}^{kl}m = \frac{2}{N}(-(G^2)_{ij} - (G^2)_{kl} + (G^2)_{ik} + (G^2)_{jl}). \tag{7.45}$$

We first estimate the terms in (7.45) which do not contain the index l. For the terms $(G^2)_{ij} + (G^2)_{ik}$, using the definition of the Green's function $(H-z)G = P_\perp$ and then using (4.12),

$$\begin{aligned}\frac{1}{N^3d^{5/2}}\sum_{ijkl}A_{ik}A_{jl}G_{il}(-(G^2)_{ij}+(G^2)_{ik}) &= \frac{1}{N^3d^{5/2}}\sum_{ijk}A_{ik}(AG)_{ij}(-(G^2)_{ij}+(G^2)_{ik})\\ &= \frac{(d-1)^{1/2}}{N^3d^{5/2}}\sum_{ijk}A_{ik}(zG+P_\perp)_{ij}(-(G^2)_{ij}+(G^2)_{ik})\\ &\prec \frac{1}{N^2d^2}\sum_{ijk}A_{ik}|(zG+P_\perp)_{ij}|\frac{\mathrm{Im}[m]}{N\eta} \prec \frac{\Lambda_\mathrm{o}}{d}\frac{\mathrm{Im}[m]}{N\eta},\end{aligned} \tag{7.46}$$

where in the second equality we used that $AG = (d-1)^{1/2}(H-z+z)G = (d-1)^{1/2}(zG+P_\perp)$. For the terms which do not contain the index i, $(G^2)_{jl}-(G^2)_{lk}$, analogously,

$$\frac{1}{N^3d^{5/2}}\sum_{ijkl}A_{ik}A_{jl}G_{il}((G^2)_{jl}-(G^2)_{lk}) \prec \frac{\Lambda_\mathrm{o}}{d}\frac{\mathrm{Im}[m]}{N\eta}. \tag{7.47}$$

We plug the estimates (7.46) and (7.47) into (7.44), and get

$$\frac{1}{N^2d(d-1)^{3/2}}\sum_{ijkl}\mathbb{E}\left[A_{ik}A_{jl}G_{il}G_{jj}\partial_{ij}^{kl}(P_\mathfrak{a}^{r-1}\bar{P}_\mathfrak{a}^r)\right] \prec \frac{\Lambda_\mathrm{o}}{d}\mathbb{E}\left[\frac{\mathrm{Im}[m]|P_\mathfrak{a}'|}{N\eta}|P_\mathfrak{a}|^{2r-2}\right]. \tag{7.48}$$

The same argument applies to the term $G_{ii}G_{kj}$, and we have the same estimate as (7.48).

To estimate the term $-G_{ii}G_{jj}$ in (7.39) we use the following Claim 7.7, which concludes the proof. □

The following claim uses a cancellation that exploits that the graph is regular.

Claim 7.7. *We have*

$$\begin{aligned}&\frac{1}{N^2d(d-1)}\sum_{ijkl}\mathbb{E}\left[A_{ik}A_{jl}G_{ii}G_{jj}D_{ij}^{kl}(P_\mathfrak{a}^{r-1}\bar{P}_\mathfrak{a}^r)\right]\\ &\quad\prec \frac{\Lambda_\mathrm{d}^2}{d^{1/2}}\mathbb{E}\left[\frac{\mathrm{Im}[m]|P_\mathfrak{a}'|}{N\eta}|P_\mathfrak{a}|^{2r-2}\right]+\frac{d}{N}\mathbb{E}[|P_\mathfrak{a}|^{2r-1}]\\ &\qquad+\left(\frac{d^{1/2}}{N}+\frac{\Lambda_\mathrm{d}}{d}\right)\max_{1\leqslant s\leqslant 2r-1}\mathbb{E}\left[\left(\frac{\mathrm{Im}[m]|P_\mathfrak{a}'|}{N\eta}+\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^2\right)^s|P_\mathfrak{a}|^{2r-s-1}\right]\\ &\qquad+\mathbb{E}\left[\frac{\mathrm{Im}[m]}{N\eta}|P_\mathfrak{a}|^{2r-1}\right].\end{aligned} \tag{7.49}$$

Proof. Using $\sum_{ij}(A_{ij} - d/N) = 0$ and Corollary 3.2 with the random variable $F = m^2 P_{\mathfrak{a}}^{r-1} P_{\mathfrak{a}}^r$ we have

$$\begin{aligned} 0 &= \frac{1}{N(d-1)} \sum_{ijkl} \mathbb{E}[(A_{ij} - \frac{d}{N}) m^2 (P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r)] \\ &= \frac{1}{N^2 d(d-1)} \sum_{ijkl} \mathbb{E}[A_{ik} A_{jl} D_{ij}^{kl} (m^2 (P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r))] + \mathrm{O}_{\prec}\left(\frac{d}{N} \mathbb{E}\left[\mathcal{C}(P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r, A) \right] \right) \\ &= \frac{1}{N^2 d(d-1)} \sum_{ijkl} \mathbb{E}[A_{ik} A_{jl} m^2 D_{ij}^{kl} (P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r)] + \mathrm{O}_{\prec}\left(\frac{d}{N} \mathbb{E}\left[\mathcal{C}(P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r, A) \right] \right. \\ &\quad \left. + \max_{1 \leqslant s \leqslant 2r} \mathbb{E}\left[\left(\frac{\mathrm{Im}[m]}{N\eta} \right)^s |P_{\mathfrak{a}}|^{2r-s} \right] \right), \end{aligned} \tag{7.50}$$

where we used (4.24), the discrete derivative rule (4.6), and (7.12). The error $\mathcal{C}(P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r, A)$ is estimated in (7.5).

Therefore, subtracting (7.50) from the left-hand side of (7.49), we get

$$\begin{aligned} &\frac{1}{N^2 d(d-1)} \sum_{ijkl} \mathbb{E}\left[A_{ik} A_{jl} G_{ii} G_{jj} D_{ij}^{kl} (P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r) \right] \\ &\quad = \frac{1}{N^2 d(d-1)} \sum_{ijkl} \mathbb{E}\left[A_{ik} A_{jl} (G_{ii} G_{jj} - m^2) D_{ij}^{kl} (P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r) \right] \\ &\qquad + \mathrm{O}_{\prec}\left(\frac{d}{N} \mathbb{E}\left[\mathcal{C}(P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r, A) \right] + \max_{1 \leqslant s \leqslant 2r} \mathbb{E}\left[\left(\frac{\mathrm{Im}[m]}{N\eta} \right)^s |P_{\mathfrak{a}}|^{2r-s} \right] \right). \end{aligned} \tag{7.51}$$

Using (4.7), we rewrite the first term on the right-hand side as

$$\begin{aligned} &\frac{\mathrm{O}(1)}{N^2 d^2} \sum_{ijkl} \mathbb{E}\left[A_{ik} A_{jl} (G_{ii} G_{jj} - m^2) D_{ij}^{kl} (P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r) \right] \\ &\quad = \sum_{n_2=1}^{\mathfrak{b}_2 - 1} \frac{1}{n_2! N^2 d^{n_2/2+2}} \sum_{ijkl} \mathbb{E}\left[A_{ik} A_{jl} (G_{ii} G_{jj} - m^2) (\partial_{ij}^{kl})^{n_2} (P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r) \right] \\ &\qquad + \frac{1}{\mathfrak{b}_2! N^2 d^{\mathfrak{b}_2/2+2}} \sum_{ijkl} \mathbb{E}\left[A_{ik} A_{jl} (G_{ii} G_{jj} - m^2) \Big((\partial_{ij}^{kl})^{\mathfrak{b}_2} (P_{\mathfrak{a}}^{r-1} \bar{P}_{\mathfrak{a}}^r)(A + \theta \xi_{ij}^{kl}) \Big) \right] \end{aligned} \tag{7.52}$$

for some random $\theta \in [0,1]$. As above, by choosing $\mathfrak{b}_2$ large enough, depending on r, we find that the last line of (7.52) is bounded by N^{-2r}. Moreover, for terms corresponding to $n_2 \geqslant 2$ in (7.52), we have the following simple estimate. Using $|G_{ii} - m| \prec \Lambda_{\mathrm{d}}$, (7.12), and that $\sum_k A_{ik} = \sum_l A_{jl} = d$, we find

$$\begin{aligned}
&\frac{1}{N^2 d^{n_2/2+2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(G_{ii}G_{jj} - m^2)(\partial_{ij}^{kl})^{n_2}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\
&\quad \prec \frac{\mathrm{O}(1)}{N^2 d^{n_2/2+2}} \max_{1\leqslant s\leqslant 2r-1} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}\Lambda_{\mathrm{d}}\left(\frac{\mathrm{Im}[m]|P_{\mathfrak{a}}'|}{N\eta} + \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-s-1}\right] \\
&\quad \prec \frac{\Lambda_{\mathrm{d}}}{d^{n_2/2}} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{\mathrm{Im}[m]|P_{\mathfrak{a}}'|}{N\eta} + \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-s-1}\right].
\end{aligned} \tag{7.53}$$

For the term in (7.52) corresponding to $n_2 = 1$, the estimate is more involved. We start by writing

$$\begin{aligned}
&\frac{\mathrm{O}(1)}{N^2 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(G_{ii}G_{jj} - m^2)\partial_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r}))\right] \\
&\quad = \frac{\mathrm{O}(1)}{N^2 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(G_{ii}G_{jj} - m^2)\right. \\
&\qquad \left. \times \left(r(\partial_{ij}^{kl}\bar{m})\bar{P}_{\mathfrak{a}}'|P_{\mathfrak{a}}|^{2r-2} + (r-1)(\partial_{ij}^{kl}m)P_{\mathfrak{a}}'\bar{P}_{\mathfrak{a}}^2|P_{\mathfrak{a}}|^{2r-3}\right)\right].
\end{aligned} \tag{7.54}$$

We estimate the term $\sum_{kl} A_{ik}A_{jl}\partial_{ij}^{kl}m$; its complex conjugate $\sum_{kl} A_{ik}A_{jl}\partial_{ij}^{kl}\bar{m}$ is estimated analogously. We use (4.9) and estimate the resulting four terms one by one. For the term $-(G^2)_{ij}$,

$$\begin{aligned}
&\frac{1}{N^3 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(G_{ii}G_{jj} - m^2)(G^2)_{ij}\bar{P}_{\mathfrak{a}}'|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad = \frac{1}{N^3 d^{1/2}} \sum_{ij} \mathbb{E}\left[(G_{ii}G_{jj} - m^2)(G^2)_{ij}\bar{P}_{\mathfrak{a}}'|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad = \frac{1}{N^3 d^{1/2}} \sum_{ij} \mathbb{E}\left[((G_{ii} - m)m + (G_{jj} - m)m\right. \\
&\qquad \left. + (G_{ii} - m)(G_{jj} - m))(G^2)_{ij}\bar{P}_{\mathfrak{a}}'|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad = \frac{1}{N^3 d^{1/2}} \sum_{ij} \mathbb{E}\left[(G_{ii} - m)(G_{jj} - m)(G^2)_{ij}\bar{P}_{\mathfrak{a}}'|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad \prec \frac{\Lambda_{\mathrm{d}}^2}{d^{1/2}} \mathbb{E}\left[\frac{\mathrm{Im}[m]|P_{\mathfrak{a}}'|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right],
\end{aligned} \tag{7.55}$$

where in the third equality, we used that the first two terms vanish after summing over index j and i respectively because $\sum_j (G^2)_{ij} = 0 = \sum_i (G^2)_{ij}$, and in the last inequality, we used $|G_{ii} - m| \prec \Lambda_{\mathrm{d}}$ and (4.11). For the term $(G^2)_{ik}$,

$$\begin{aligned}
&\frac{1}{N^3 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(G_{ii}G_{jj} - m^2)(G^2)_{ik}\bar{P}'_{\mathfrak{a}}|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad = \frac{1}{N^3 d^{3/2}} \sum_{ij} \mathbb{E}\left[(G_{ii}G_{jj} - m^2)(AG^2)_{ii}\bar{P}'_{\mathfrak{a}}|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad = \frac{(d-1)^{1/2}}{N^3 d^{3/2}} \sum_{ij} \mathbb{E}\left[(G_{ii}G_{jj} - m^2)(zG^2 + G)_{ii}\bar{P}'_{\mathfrak{a}}|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad \prec \frac{\Lambda_{\mathrm{d}}}{d}\mathbb{E}\left[\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right],
\end{aligned} \tag{7.56}$$

where we used that $AG^2 = (d-1)^{1/2}(H - z + z)G^2 = (d-1)^{1/2}(zG^2 + G)$ and (4.11). We have the same estimate for the term $(G^2)_{jl}$,

$$\frac{1}{N^3 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(G_{ii}G_{jj} - m^2)(G^2)_{jl}\bar{P}'_{\mathfrak{a}}|P_{\mathfrak{a}}|^{2r-2}\right] \prec \frac{\Lambda_{\mathrm{d}}}{d}\mathbb{E}\left[\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right]. \tag{7.57}$$

For the term $-(G^2)_{kl}$, we use (4.11)

$$\begin{aligned}
&\frac{1}{N^3 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(G_{ii}G_{jj} - m^2)(G^2)_{kl}\bar{P}'_{\mathfrak{a}}|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad = \frac{1}{N^3 d^{5/2}} \sum_{ij} \mathbb{E}\left[(G_{ii}G_{jj} - m^2)(AG^2A)_{ij}\bar{P}'_{\mathfrak{a}}|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad = \frac{d-1}{N^3 d^{5/2}} \sum_{ij} \mathbb{E}\left[(G_{ii}G_{jj} - m^2)(z^2G^2 + 2zG + P_\perp)_{ij}\bar{P}'_{\mathfrak{a}}|P_{\mathfrak{a}}|^{2r-2}\right] \\
&\quad \prec \frac{\Lambda_{\mathrm{d}}}{d^{3/2}}\mathbb{E}\left[\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right],
\end{aligned} \tag{7.58}$$

where in the third equality, we used $AG^2A = (d-1)(H - z + z)G^2(H - z + z) = (d-1)(z^2G^2 + 2zG + P_\perp)$. We combine the estimates (7.55), (7.56), (7.57) and (7.58), and use that $\Lambda_{\mathrm{d}} \geqslant 1/\sqrt{d}$,

$$\frac{\mathrm{O}(1)}{N^2 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(G_{ii}G_{jj} - m^2)\partial_{ij}^{kl}(\bar{P}_{\mathfrak{a}}|P_{\mathfrak{a}}|^{2r-2}))\right] \prec \frac{\Lambda_{\mathrm{d}}^2}{d^{1/2}}\mathbb{E}\left[\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|}{N\eta}|P_{\mathfrak{a}}|^{2r-2}\right]. \tag{7.59}$$

The Claim 7.7 follows from combining (7.51), (7.53) and (7.59). □

Claim 7.8. *For the term $n_1 = 2$ on the right-hand side of (7.35),*

$$\begin{aligned}
&\frac{1}{N^2 d(d-1)^{3/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(\partial_{ij}^{kl})^2(G_{ij})D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\
&\quad \prec \frac{\Lambda_{\mathrm{o}}}{d} \max_{1 \leqslant s \leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P'_{\mathfrak{a}}|\,\mathrm{Im}[m]}{N\eta} + \left(\frac{\mathrm{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right].
\end{aligned} \tag{7.60}$$

Proof. For $i \neq j$ we use $(\partial_{ij}^{kl})^2 G_{ij} \prec \Lambda_{\mathrm{o}}$ from (4.13),

$$\begin{aligned}
&\frac{1}{N^2 d(d-1)^{3/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}(\partial_{ij}^{kl})^2(G_{ij}) D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})\right] \\
&\quad \prec \frac{\Lambda_{\mathrm{o}}}{N^2 d^{5/2}} \sum_{ijkl} \mathbb{E}\left[A_{ik}A_{jl}|D_{ij}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})|\right] + \frac{1}{N^2 d^{5/2}} \sum_{ikl} \mathbb{E}\left[A_{ik}A_{il}|D_{ii}^{kl}(P_{\mathfrak{a}}^{r-1}\bar{P}_{\mathfrak{a}}^{r})|\right] \\
&\quad \prec \frac{\Lambda_{\mathrm{o}}}{d} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{|P_{\mathfrak{a}}'|\operatorname{Im}[m]}{N\eta} + \left(\frac{\operatorname{Im}[m]}{N\eta}\right)^2\right)^s |P_{\mathfrak{a}}|^{2r-1-s}\right],
\end{aligned} \tag{7.61}$$

where we used (7.37) for the second inequality. □

8 Analysis of Self-consistent Equation and Proof of Theorem 1.1

In this section we analyse the recursive moment estimate (7.1) from Proposition 7.1, around the spectral edges ± 2, and obtain an improved estimate for the Stieltjes transform $m(z)$.

In the following, we focus on the right spectral edge; an analogous argument applies to the left edge. For a fixed integer $\mathfrak{a} \geqslant 1$ and the same large constant $\mathfrak{K} > 0$ as in (2.10), we define the spectral domain for the right edge (around the point $z = 2$) by

$$\begin{aligned}
\mathbf{D}_{\mathrm{e}} = \Bigg\{ & 2 + \kappa + \mathrm{i}\eta : 0 \leqslant \eta \leqslant \mathfrak{K}, 0 \leqslant 2 + \kappa \leqslant \mathfrak{K} \\
& |\kappa| + \eta \geqslant \frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N} + \frac{d^3}{N^2} + \frac{d^{3/2}}{N(N\eta)^{1/2}} + \frac{1}{d^{3/2}N\eta} + \frac{1}{(N\eta)^2} \Bigg\}.
\end{aligned} \tag{8.1}$$

Also recall m_d from (2.5).

Theorem 8.1. *Fix an integer $\mathfrak{a} \geqslant 1$. For $1 \ll d \ll N^{2/3}$, the following holds uniformly for any $z = 2 + \kappa + \mathrm{i}\eta \in \mathbf{D}_{\mathrm{e}}$.*

- *If $\kappa \geqslant 0$ then*

$$\begin{aligned}
|m(z) - m_d(z)| \prec & \left(\frac{1}{d} + \frac{d}{N} + \frac{1}{(dN\eta)^{1/2}}\right)^{1/3} \frac{1}{N\eta^{2/3}(|\kappa| + \eta)^{1/3}} \\
& + \frac{1}{\sqrt{|\kappa| + \eta}} \left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N} + \frac{d^3}{N^2} + \frac{1}{N^{1/2}d^{3/2}} + \frac{d^{3/2}}{N^{3/2}\eta^{1/2}} + \frac{1}{N\eta^{1/2}} + \frac{1}{d^{3/2}N\eta} + \frac{1}{(N\eta)^2}\right).
\end{aligned} \tag{8.2}$$

- *If $\kappa \leqslant 0$ then*

$$|m(z)-m_d(z)| \prec \frac{1}{(N\eta)^{1/2}}\left(\frac{1}{N\eta}+\frac{d^{5/2}}{N^2}+\frac{1}{d^{3/2}}\right)^{1/2} + \frac{1}{\sqrt{|\kappa|+\eta}}\left(\frac{1}{d^{\mathfrak{a}/2}}+\frac{d}{N}+\frac{d^3}{N^2}+\frac{d^{3/2}}{N\sqrt{N\eta}}\right). \tag{8.3}$$

The analogous statement holds around the left edge.

In this section, we use the notation $X \lesssim Y$ to mean $X = \mathrm{O}(Y)$ and $X \asymp Y$ to mean $X \lesssim Y$ and $Y \lesssim X$.

The estimate for the extremal eigenvalues of the random d-regular graphs, Theorem 1.1, follows as a corollary of Theorem 8.1 and [CGJ18, Theorem 1.1], which states that with high probability, the second-largest eigenvalue and the smallest eigenvalue of a random d-regular graph is of order $\mathrm{O}(1)$.

PROPOSITION 8.2 ([CGJ18, Theorem 1.1]). *Fix $\mathfrak{c} > 0$. For $1 \leqslant d \lesssim N^{2/3}$ and large enough N, there exists a constant $\mathfrak{K}$ depending on $\mathfrak{c}$ such that, with probability at least $1 - N^{-1/\mathfrak{c}}$,*

$$\max\{\lambda_2, -\lambda_N\} \leqslant \mathfrak{K}. \tag{8.4}$$

Proof of Theorem 1.1. We recall the following elementary estimates on $m_d(z)$ which hold for bounded $|z|$. Let $z = 2+\kappa+\mathrm{i}\eta$ or $z = -2-\kappa+\mathrm{i}\eta$ for $0 \leqslant \eta \leqslant \mathfrak{K}, 0 \leqslant 2+\kappa \leqslant \mathfrak{K}$. Then

$$\operatorname{Im}[m_d(z)] \asymp \begin{cases} \sqrt{|\kappa|+\eta} & \text{if } \kappa \leqslant 0, \\ \frac{\eta}{\sqrt{|\kappa|+\eta}} & \text{if } \kappa \geqslant 0. \end{cases} \tag{8.5}$$

This estimate follows from (2.5) and the analogous estimate for m_{sc}; see, for example, [EY17, Lemma 6.2].

We choose $\mathfrak{a} = 6$ in Theorem 8.1, and take $z = 2+\kappa+\mathrm{i}\eta \in \mathbf{D}_{\mathrm{e}}$, where

$$\kappa = N^{\mathfrak{c}}\left(\frac{1}{d^3}+\frac{1}{N^{2/3}}+\frac{d^2}{N^{4/3}}\right), \qquad \eta = \frac{N^{\mathfrak{c}/2}}{N\sqrt{\kappa}}.$$

In particular, $\eta \ll \kappa$. Then under our assumption $1 \ll d \ll N^{2/3}$, (8.2) implies

$$|m(z)-m_d(z)| \ll \frac{1}{N\eta}. \tag{8.6}$$

Since $\operatorname{Im}[m_d(z)] \asymp \eta/\sqrt{\kappa} \ll 1/N\eta$, we get $\operatorname{Im}[m(z)] \ll 1/N\eta$. Since any eigenvalue in $[2+\kappa-\eta, 2+\kappa+\eta]$ would yield a positive contribution of size at least $1/N\eta$ to $\operatorname{Im}[m(z)]$, this implies that there cannot be any eigenvalue on the interval $[2+\kappa-\eta, 2+\kappa+\eta]$. Since we can take any κ in the interval $[N^{\mathfrak{c}}(d^{-3}+N^{-2/3}+d^2N^{-4/3}), \mathfrak{K}-2]$, combining with Proposition 8.2, we conclude that with probability $1-2N^{-1/\mathfrak{c}}$ we have $\lambda_2 \leqslant 2+N^{\mathfrak{c}}(d^{-3}+N^{-2/3}+d^2N^{-4/3})$. Since $\mathfrak{c} > 0$ is arbitrary this implies the same conclusion with probability $1-N^{-1/\mathfrak{c}}$ as in the statement of the theorem. The lower bound for λ_N follows from the same argument. This finishes the proof of Theorem 1.1. □

8.1 Stability of self-consistent equation. Recall from Corollary 6.4 that for $z \in \mathbf{D}_{\mathrm{e}}$,

$$|P_{\mathfrak{a}}'(z, m_d(z))| \asymp \sqrt{|\kappa| + \eta} + \mathrm{O}(d^{-\mathfrak{a}/2}), P_{\mathfrak{a}}''(z, m_d(z)) = 2 + \mathrm{O}(d^{-1/2}),$$
$$P_{\mathfrak{a}}'''(z, m_d(z)) = \mathrm{O}(1). \tag{8.7}$$

The following proposition on the stability of the self-consistent equation, relying on the above square root behaviour at the edge, is essentially [BEKYY14, Lemma 4.5].

PROPOSITION 8.3. *Fix an integer $\mathfrak{a} \geqslant 1$. There exists a constant $\varepsilon > 0$ such that the following holds. Suppose that $\delta : \mathbf{D}_{\mathrm{e}} \to \mathbb{C}$ satisfies $N^{-2} \leqslant \delta(z) \leqslant \varepsilon$ for $z \in \mathbf{D}_{\mathrm{e}}$, and that δ is Lipschitz continuous with Lipschitz constant N. Suppose moreover that for each fixed κ, the function $\eta \mapsto \delta(2 + \kappa + \mathrm{i}\eta)$ is nonincreasing for $\eta > 0$. Suppose that for all $z \in \mathbf{D}_{\mathrm{e}}$ we have*

$$|P_{\mathfrak{a}}(z, m_d(z))| + |P_{\mathfrak{a}}(z, m(z))| \leqslant \delta(z). \tag{8.8}$$

Then we have for $z = 2 + \kappa + \mathrm{i}\eta \in \mathbf{D}_{\mathrm{e}}$,

$$|m(z) - m_d(z)| = \mathrm{O}\left(\frac{\delta(z)}{\sqrt{|\kappa| + \eta + \delta(z)}}\right), \tag{8.9}$$

where the implicit constant is independent of z and N.

Proof. Let $z = 2 + \kappa + \mathrm{i}\eta \in \mathbf{D}_{\mathrm{e}}$. By Taylor expansion, (8.7), and Proposition 2.1, it follows that

$$\begin{aligned} P_{\mathfrak{a}}(z, m(z)) = P_{\mathfrak{a}}(z, m_d(z)) + P_{\mathfrak{a}}'(z, m_d(z))(m(z) - m_d(z)) \\ + (P_{\mathfrak{a}}''(z, m_d(z)) + o(1))(m(z) - m_d(z))^2/2. \end{aligned} \tag{8.10}$$

We abbreviate $\mathcal{R}(z) := P_{\mathfrak{a}}(z, m(z)) - P_{\mathfrak{a}}(z, m_d(z))$. There exists $a(z) \asymp \sqrt{|\kappa| + \eta}$ and $b(z) \asymp 1$, such that

$$\mathcal{R}(z) = a(z)(m(z) - m_d(z)) + b(z)(m(z) - m_d(z))^2. \tag{8.11}$$

With (8.11), Proposition 8.3 follows by a continuity argument that is essentially the same as [BEKYY14, Lemma 4.5]. □

8.2 Estimates on individual Green's function entries. In the following we first prove some estimates on the individual entries of Green's function, which slightly improve the estimates on the diagonal Green's function entries from [BKY17, Theorem 1.1] using the results from the previous sections.

PROPOSITION 8.4. *Fix $1 \ll d \ll N^{2/3}$. Uniformly for $z = 2 + \kappa + \mathrm{i}\eta \in \mathbf{D}_{\mathrm{e}}$, we have (4.1) with Λ_{o} and Λ_{d} given by (2.13). The analogous statement holds for the left edge.*

Proof. From Proposition 2.1, we have (4.1) for all $z \in \mathbf{D}_{\mathrm{e}}$, with Λ_{o} and Λ_{d} given by (2.11). (In all all bounds below, Λ_{o} and Λ_{d} will continue to be given by (2.11).) By the rough bounds $\mathrm{Im}[m] \prec 1$ and $|P'_{\mathfrak{a}}| \prec 1$, it follows from Proposition 7.1 that

$$\begin{aligned}\mathbb{E}[|P_{\mathfrak{a}}|^{2r}] \prec \frac{\Lambda_{\mathrm{d}}^2}{d^{1/2}}\mathbb{E}\left[\frac{|P_{\mathfrak{a}}|^{2r-2}}{N\eta}\right] + \left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\right)\mathbb{E}[|P_{\mathfrak{a}}|^{2r-1}] \\ + \max_{1\leqslant s\leqslant 2r}\mathbb{E}\left[\frac{|P_{\mathfrak{a}}|^{2r-s}}{(N\eta)^s}\right] + \left(\frac{\Lambda_{\mathrm{d}}}{d} + \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}} + \frac{d\Lambda_{\mathrm{o}}}{N}\right)\max_{1\leqslant s\leqslant 2r-1}\mathbb{E}\left[\frac{|P_{\mathfrak{a}}|^{2r-s-1}}{(N\eta)^s}\right].\end{aligned} \tag{8.12}$$

By Jensen's inequality, we get from (8.12) that for any $r \geqslant 1$,

$$\mathbb{E}[|P_{\mathfrak{a}}(z, m(z))|^{2r}] \prec \frac{1}{d^{r\mathfrak{a}}} + \frac{\Lambda_{\mathrm{d}}^{2r}}{d^r} + \frac{1}{(N\eta)^{2r}} + \frac{d^{3r}\Lambda_{\mathrm{o}}^{2r}}{N^{2r}}. \tag{8.13}$$

Therefore

$$|P_{\mathfrak{a}}(z, m(z))| \prec \frac{1}{d^{\mathfrak{a}/2}} + \frac{\Lambda_{\mathrm{d}}^2}{d^{1/2}} + \frac{1}{N\eta} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}, \tag{8.14}$$

uniformly for $z \in \mathbf{D}_{\mathrm{e}}$. Taking

$$\delta(z) = N^{\mathfrak{c}}\left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{\Lambda_{\mathrm{d}}^2}{d^{1/2}} + \frac{1}{N\eta} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\right), \tag{8.15}$$

in Proposition 8.3, where $\mathfrak{c} > 0$, we obtain

$$|m(z) - m_d(z)| \prec \frac{1}{d^{\mathfrak{a}/4}} + \frac{\Lambda_{\mathrm{d}}}{d^{1/4}} + \frac{1}{(N\eta)^{1/2}} + \frac{d^{3/4}\Lambda_{\mathrm{o}}^{1/2}}{N^{1/2}} \prec \frac{1}{d^{1/2}} + \frac{d^{3/2}}{N} + \frac{1}{\sqrt{N\eta}}. \tag{8.16}$$

where the last bound follows from the definitions of Λ_{o} and Λ_{d} in (2.11) and taking $\mathfrak{a} \geqslant 2$. The proof is completed by applying [BKY17, Lemma 5.4]. Indeed, thanks to [BKY17, Lemma 5.4], we have

$$1 + (m(z) + z)G_{ii}(z) = \mathrm{O}_{\prec}\left(\frac{1}{d^{1/2}} + \frac{d^{3/2}}{N} + \frac{1}{\sqrt{N\eta}}\right). \tag{8.17}$$

With (8.16) as input and noticing that $|m_{\mathrm{sc}}(z) - m_d(z)| = \mathrm{O}(1/d)$, we get

$$1 + (m_{\mathrm{sc}}(z) + z)G_{ii}(z) = \mathrm{O}_{\prec}\left(\frac{1}{d^{1/2}} + \frac{d^{3/2}}{N} + \frac{1}{\sqrt{N\eta}}\right). \tag{8.18}$$

Using (2.7) we obtain

$$G_{ii}(z) - m_{\mathrm{sc}}(z) = \mathrm{O}_{\prec}\left(\frac{1}{d^{1/2}} + \frac{d^{3/2}}{N} + \frac{1}{\sqrt{N\eta}}\right), \tag{8.19}$$

from which we deduce that

$$\max_i |G_{ii}(z) - m_d(z)| \prec \frac{1}{d^{1/2}} + \frac{d^{3/2}}{N} + \frac{1}{\sqrt{N\eta}}, \tag{8.20}$$

as claimed. □

8.3 Proof of Theorem 8.1. Throughout this section, we abbreviate $\Lambda(z) = |m(z) - m_d(z)|$, and $\Phi(z) = \mathrm{Im}[m_d(z)]$. Then $\mathrm{Im}[m] = \Phi(z) + \mathrm{O}(\Lambda(z))$, and $P'_{\mathfrak{a}}(z, m(z)) = P'_{\mathfrak{a}}(z, m_d(z)) + \mathrm{O}(\Lambda(z))$ from Corollary 6.4. Moreover, we take Λ_{o} and Λ_{d} to be given by (2.13). By Proposition 8.4, we know that (4.1) holds for any $z \in \mathbf{D}_{\mathrm{e}}$. Notice that Proposition 7.1 implies that, for any $z \in \mathbf{D}_{\mathrm{e}}$,

$$\begin{aligned}\mathbb{E}[|P_{\mathfrak{a}}|^{2r}] &\prec \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}} \max_{1\leqslant s\leqslant 2r-1} \mathbb{E}\left[\left(\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|}{N\eta}\right)^s |P_{\mathfrak{a}}|^{2r-s-1}\right] + \mathbb{E}\left[\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|}{(N\eta)^2}|P_{\mathfrak{a}}|^{2r-2}\right] \\ &+ \frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\mathbb{E}\left[\frac{\mathrm{Im}[m]|P'_{\mathfrak{a}}|^2}{(N\eta)^3}|P_{\mathfrak{a}}|^{2r-3}\right] + \left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\right)\mathbb{E}[|P_{\mathfrak{a}}|^{2r-1}] \\ &+ \max_{1\leqslant s\leqslant 2r} \mathbb{E}\left[\left(\frac{\mathrm{Im}[m]}{N\eta}\right)^s |P_{\mathfrak{a}}|^{2r-s}\right].\end{aligned} \tag{8.21}$$

We notice that for $z = 2+\kappa+\mathrm{i}\eta \in \mathbf{D}_{\mathrm{e}}$, $|P'_{\mathfrak{a}}(z, m(z))| \lesssim \sqrt{|\kappa|+\eta}+\Lambda(z)$. By Jensen's inequality, it follows from (8.21) that

$$\begin{aligned}\mathbb{E}\left[|P(z, m(z))|^{2r}\right] &\prec \mathbb{E}\left[\left(\frac{1}{N\eta} + \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\right)^r \left(\frac{(\Phi+\Lambda)(\sqrt{|\kappa|+\eta}+\Lambda)}{N\eta}\right)^r\right] \\ &+ \mathbb{E}\left[\left(\frac{\Phi+\Lambda}{N\eta}\right)^{2r}\right] + \mathbb{E}\left[\left(\frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\right)^{2r/3} \frac{(\Phi+\Lambda)^{2r/3}(\sqrt{|\kappa|+\eta}+\Lambda)^{4r/3}}{(N\eta)^{2r}}\right] \\ &+ \mathbb{E}\left[\left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}\right)^{2r}\right].\end{aligned} \tag{8.22}$$

Before proving Theorem 8.1, we prove the following weaker estimate which will be used as an input in the Proof of Theorem 8.1.

Proposition 8.5. *Uniformly for any $z = 2+\kappa+\mathrm{i}\eta \in \mathbf{D}_{\mathrm{e}}$, we have*

$$\Lambda(z) = |m(z) - m_d(z)| \prec \sqrt{|\kappa|+\eta}. \tag{8.23}$$

Proof. Thanks to (8.5), we have $\Phi(z) \prec \sqrt{|\kappa|+\eta}$. From our choice of the spectral domain $\mathbf{D}_{\mathrm{e}}$ in (8.1), we have

$$\frac{1}{N\eta} \leqslant \sqrt{|\kappa|+\eta}, \quad \frac{1}{d^{\mathfrak{a}/2}}, \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}, \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\frac{1}{N\eta} \leqslant |\kappa|+\eta. \tag{8.24}$$

Thus (8.22) implies

$$\mathbb{E}\left[|P_{\mathfrak{a}}(z, m(z))|^{2r}\right] \prec (|\kappa|+\eta)^{2r} + (|\kappa|+\eta)^r \mathbb{E}\left[\Lambda(z)^{2r}\right]. \tag{8.25}$$

We have the Taylor expansion (8.10)

$$\begin{aligned} P_{\mathfrak{a}}(z,m(z)) &= P_{\mathfrak{a}}(z,m_d(z)) + P'_{\mathfrak{a}}(z,m_d(z))(m(z) \\ &\quad -m_d(z)) + (1+\mathrm{o}(1))(m(z)-m_d(z))^2, \end{aligned} \tag{8.26}$$

where we used that $P''_{\mathfrak{a}} = 2 + \mathrm{O}(1/d^{1/2} + d/N)$, $P'''_{\mathfrak{a}} = \mathrm{O}(1)$, and $\Lambda(z) \ll 1$. Rearranging the last equation (8.26) and using the definition of Λ, we have arrived at

$$\Lambda(z)^2 \lesssim \Lambda(z)\sqrt{|\kappa|+\eta} + |P_{\mathfrak{a}}(z,m(z))| + |P_{\mathfrak{a}}(z,m_d(z))|, \tag{8.27}$$

and thus

$$\mathbb{E}[\Lambda(z)^{4r}] \lesssim (|\kappa|+\eta)^r \mathbb{E}[\Lambda(z)^{2r}] + \mathbb{E}[|P_{\mathfrak{a}}(z,m(z))|^{2r}] + |P_{\mathfrak{a}}(z,m_d(z))|^{2r}, \tag{8.28}$$

for any fixed integer $r \geqslant 1$. Now we replace $\mathbb{E}[|P_{\mathfrak{a}}(z,m(z))|^{2r}]$ in (8.28) by the right-hand side of (8.25). Moreover, on the domain $\mathbf{D}_{\mathrm{e}}$, from Proposition 6.1, we have $|P_{\mathfrak{a}}(z,m_d(z))| \leqslant |\kappa|+\eta$. With $\mathbb{E}[\Lambda(z)^{2r}] \leqslant \mathbb{E}[\Lambda(z)^{4r}]^{1/2}$ by the Cauchy–Schwarz inequality, we thus obtain

$$\mathbb{E}[\Lambda(z)^{4r}] \prec (|\kappa|+\eta)^{2r}, \tag{8.29}$$

for any $r \geqslant 1$. The claim $\Lambda(z) \prec \sqrt{|\kappa|+\eta}$ follows from Markov's inequality. □

Proof of of Theorem 8.1. We assume that there exists some deterministic control parameter $\Theta(z)$, for any $z = 2+\kappa+\mathrm{i}\eta \in \mathbf{D}_{\mathrm{e}}$, we have the a priori estimate

$$|m(z) - m_d(z)| \prec \Theta(z) \lesssim \sqrt{|\kappa|+\eta}. \tag{8.30}$$

From (8.22) we get

$$\begin{aligned} |P_{\mathfrak{a}}(z,m(z))| &\prec \left(\frac{1}{N\eta} + \frac{\Lambda_{\mathrm{o}}^2}{d^{1/2}}\right)^{1/2} \left(\frac{(\Phi+\Theta)\sqrt{|\kappa|+\eta}}{N\eta}\right)^{1/2} + \frac{\Phi+\Theta}{N\eta} \\ &\quad + \left(\frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\right)^{1/3} \frac{(\Phi+\Theta)^{1/3}(|\kappa|+\eta)^{1/3}}{N\eta} + \frac{1}{d^{\mathfrak{a}/2}} + \frac{d^{3/2}\Lambda_{\mathrm{o}}}{N}. \end{aligned} \tag{8.31}$$

Since $\Phi+\Theta \lesssim \sqrt{|\kappa|+\eta}$, (8.31) simplifies to

$$\begin{aligned} &|P_{\mathfrak{a}}(z,m(z))| \\ &\prec \left(\frac{1}{N\eta} + \frac{d^{5/2}}{N^2} + \frac{1}{d^{3/2}}\right)^{1/2} \left(\frac{(\Phi+\Theta)\sqrt{|\kappa|+\eta}}{N\eta}\right)^{1/2} \\ &\quad + \left(\frac{\Lambda_{\mathrm{o}}}{d^{1/2}}\right)^{1/3} \frac{(\Phi+\Theta)^{1/3}(|\kappa|+\eta)^{1/3}}{N\eta} + \left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N} + \frac{d^3}{N^2} + \frac{d^{3/2}}{N\sqrt{N\eta}}\right). \end{aligned} \tag{8.32}$$

To conclude the proof, we consider the cases $\kappa \leqslant 0$ and $\kappa \geqslant 0$ separately. If $\kappa \leqslant 0$, then we have $\Phi(z) \asymp \sqrt{|\kappa| + \eta}$, and (8.32) simplifies to

$$\begin{aligned} |P_{\mathfrak{a}}(z, m(z))| \prec \left(\frac{1}{N\eta} + \frac{d^{5/2}}{N^2} + \frac{1}{d^{3/2}} \right)^{1/2} \left(\frac{|\kappa| + \eta}{N\eta} \right)^{1/2} \\ + \left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N} + \frac{d^3}{N^2} + \frac{d^{3/2}}{N\sqrt{N\eta}} \right). \end{aligned} \tag{8.33}$$

Thanks to Proposition 8.3, by taking $\delta(z)$ the right-hand side of (8.33) times $N^{\mathfrak{c}}$, we get

$$\begin{aligned} |m(z) - m_d(z)| \prec \frac{1}{(N\eta)^{1/2}} \left(\frac{1}{N\eta} + \frac{d^{5/2}}{N^2} + \frac{1}{d^{3/2}} \right)^{1/2} \\ + \frac{1}{\sqrt{|\kappa| + \eta}} \left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N} + \frac{d^3}{N^2} + \frac{d^{3/2}}{N\sqrt{N\eta}} \right). \end{aligned} \tag{8.34}$$

This finishes the proof of (8.3).

If $\kappa \geqslant 0$, then $\Phi(z) \asymp \eta/\sqrt{|\kappa| + \eta}$, and (8.32) simplifies to

$$\begin{aligned} |P_{\mathfrak{a}}(z, m(z))| \prec \frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N} + \frac{d^3}{N^2} + \frac{1}{N^{1/2} d^{3/4}} + \frac{d^{3/2}}{N^{3/2}\eta^{1/2}} + \frac{1}{N\eta^{1/2}} \\ + \left(\frac{\Lambda_{\mathrm{o}}}{d^{1/2}} \right)^{1/3} \left(\frac{(|\kappa| + \eta)^{1/3}}{N\eta} \Theta^{1/3} + \frac{(|\kappa| + \eta)^{1/6}}{N\eta^{2/3}} \right) \\ + \left(\frac{1}{N\eta} + \frac{d^{5/2}}{N^2} + \frac{1}{d^{3/2}} \right)^{1/2} \left(\frac{\sqrt{|\kappa| + \eta}}{N\eta} \right)^{1/2} \Theta^{1/2}. \end{aligned} \tag{8.35}$$

Thanks to Proposition 8.3, by taking $\delta(z)$ to be the right-hand side of (8.35) times $N^{\mathfrak{c}}$, we get

$$\begin{aligned} |m(z) - m_d(z)| \prec \frac{1}{\sqrt{|\kappa| + \eta}} \left(\frac{1}{d^{\mathfrak{a}/2}} + \frac{d}{N} + \frac{d^3}{N^2} + \frac{1}{N^{1/2} d^{3/4}} + \frac{d^{3/2}}{N^{3/2}\eta^{1/2}} + \frac{1}{N\eta^{1/2}} \right) \\ + \left(\frac{\Lambda_{\mathrm{o}}}{d^{1/2}} \right)^{1/3} \left(\frac{1}{N\eta(|\kappa| + \eta)^{1/6}} \Theta^{1/3} + \frac{1}{N\eta^{2/3}(|\kappa| + \eta)^{1/3}} \right) \\ + \left(\frac{1}{N\eta} + \frac{d^{5/2}}{N^2} + \frac{1}{d^{3/2}} \right)^{1/2} \left(\frac{1}{N\eta\sqrt{|\kappa| + \eta}} \right)^{1/2} \Theta^{1/2}. \end{aligned} \tag{8.36}$$

Since the exponent of Θ on the right-hand side of (8.36) is less than 1, by iterating (8.36) a bounded number of times and recalling the definition of $\prec$, we get

$$|m(z)-m_d(z)| \prec \left(\frac{1}{d}+\frac{d}{N}+\frac{1}{(dN\eta)^{1/2}}\right)^{1/3}\frac{1}{N\eta^{2/3}(|\kappa|+\eta)^{1/3}} + \frac{1}{\sqrt{|\kappa|+\eta}}\left(\frac{1}{d^{a/2}}+\frac{d}{N}+\frac{d^3}{N^2}+\frac{1}{N^{1/2}d^{3/2}}+\frac{d^{3/2}}{N^{3/2}\eta^{1/2}}+\frac{1}{N\eta^{1/2}}+\frac{1}{d^{3/2}N\eta}+\frac{1}{(N\eta)^2}\right). \tag{8.37}$$

This finishes the proof of (8.2). □

9 Edge Universality: Proof of Theorem 1.3

In this section we prove the edge universality of random d-regular graphs in the regime $N^{2/9} \ll d \ll N^{1/3}$ where Theorem 1.1 provides optimal bounds on the extremal eigenvalues. Thus, throughout this section we assume that there exists a universal constant $\mathfrak{e}$ such that

$$N^{2/9+\mathfrak{e}} \leqslant d \leqslant N^{1/3-\mathfrak{e}}. \tag{9.1}$$

Our strategy is based on the usual three-step approach of random matrix theory [EY17]. Starting from the (rescaled) adjacency matrix $H = H_0$ from (1.1), we run a matrix-valued Brownian motion $H(t)$. Using the rigidity estimates from Theorems 1.1 and 8.1 combined with [LY17], we deduce that for $t \gg N^{-1/3}$ the matrix $H(t)$ has GOE edge statistics; see Proposition 9.8 below. The main work in this section is a comparison argument to show that H_0 and $H(t)$, $t \gg N^{-1/3}$, have the same edge statistics; see Proposition 9.11 below. This comparison argument has the same spirit as the one from [BHKY17]. Its underlying principle is that for large d, a Markovian switching dynamics of the graph that leaves the random regular graph invariant is well approximated by Brownian motion, when considering observables that characterize the local spectral statistics. The main difference to [BHKY17] is that, since we are working at the edge, we need to incorporate precise rigidity estimates on the locations of the eigenvalues near the spectral edge, and the necessary cancellations are more delicate.

We recall the constrained GOE W as introduced in [BHKY17, Section 2.1]: W is the centred Gaussian process on the space $\mathcal{M} := \{H \in \mathbb{R}^{N\times N} : H = H^*, H\mathbf{1} = 0\}$ with covariance $\mathbb{E}\langle W, X\rangle\langle W, Y\rangle = \langle X, Y\rangle$, where $\langle X, Y\rangle := \frac{N}{2}\operatorname{Tr}(XY)$ for $X, Y \in \mathcal{M}$. Explicitly, its covariance is given by

$$\mathbb{E}[W_{ij}W_{kl}] = \frac{1}{N}\left(\delta_{ik}-\frac{1}{N}\right)\left(\delta_{jl}-\frac{1}{N}\right)+\frac{1}{N}\left(\delta_{il}-\frac{1}{N}\right)\left(\delta_{jk}-\frac{1}{N}\right). \tag{9.2}$$

It may be viewed as the usual Gaussian Orthogonal Ensemble restricted to matrices with vanishing row and column sums. The following result is a straightforward consequence of (9.2).

LEMMA 9.1. *For the constrained GOE W we have the integration by parts formula*

$$\mathbb{E}[W_{ij}F(W)] = \frac{1}{N^3}\sum_{kl}\mathbb{E}[\partial_{ij}^{kl}F(W)]. \tag{9.3}$$

Next, we define the matrix-valued process

$$H(t) := \mathrm{e}^{-t/2}H + \sqrt{1-\mathrm{e}^{-t}}\,W, \tag{9.4}$$

where H was defined in (1.1) in terms of the adjacency matrix of the d-regular graph. Thus, $H(0) = H$ and $H(\infty) = W$.

We recall the projection matrix $P_\perp = I - \mathbf{1}\mathbf{1}^*/N$ from Section 2. Then the matrix $H(t)$ and $P_\perp$ commute, i.e. $H(t)P_\perp = P_\perp H(t)$. For $z \in \mathbb{C}_+ = \{z \in \mathbb{C} : \mathrm{Im}[z] > 0\}$, we define the *time-dependent Green's function* by

$$G(t;z) := P_\perp(H(t) - z)^{-1}P_\perp, \tag{9.5}$$

so that $G(t;z)$ and $(H(t)-z)^{-1}$ agree on the image of $P_\perp$, i.e. the subspace of $\mathbb{R}^N$ perpendicular to $\mathbf{1}$ which carries the nontrivial spectrum of $H(t)$. The matrix $H(t)$ has a trivial eigenvalue $\lambda_1(t) = \mathrm{e}^{-t/2}d/\sqrt{d-1}$ with eigenvector $\mathbf{1}$. We denote the remaining eigenvalues of $H(t)$ by $\lambda_2(t) \geqslant \lambda_3(t) \geqslant \cdots \lambda_{N-1}(t) \geqslant \lambda_N(t)$, and the Stieltjes transform of the empirical eigenvalue distribution of $H(t)P_\perp$ by $m(t;z)$,

$$m(t;z) = \frac{1}{N}\mathrm{Tr}\,G(t;z) = \frac{1}{N}\sum_{i=2}^{N}\frac{1}{\lambda_i(t)-z}. \tag{9.6}$$

DEFINITION 9.2. *An event holds with* very high probability *if for any $\mathfrak{c} > 0$ it has probability at least $1 - \mathrm{O}_{\mathfrak{c}}(N^{-1/\mathfrak{c}})$.*

Under the assumption (9.1), we have the following corollary of Theorems 1.1 and 8.1 .

COROLLARY 9.3. *Fix a constant $\mathfrak{e} > 0$ and suppose that $N^{2/9+\mathfrak{e}} \leqslant d \leqslant N^{1/3-\mathfrak{e}}$. Then, for any fixed $\mathfrak{c} > 0$ we have with very high probability*

$$|\lambda_2(0) - 2|, |\lambda_N(0) + 2| \leqslant N^{-2/3+\mathfrak{c}}, \tag{9.7}$$

and uniformly for any $z = E + \mathrm{i}\eta$, with $-4 \leqslant E \leqslant 4$ and $\eta \geqslant N^{-2/3}$,

$$|m(0;z) - m_d(z)| \leqslant N^{\mathfrak{c}}\left(\frac{1}{N\eta} + \frac{d}{N\sqrt{\eta}}\right). \tag{9.8}$$

REMARK 9.4. As an easy consequence of Corollary 9.3, we have that for any $z = E + \mathrm{i}\eta$, with $-4 \leqslant E \leqslant 4$ and $\eta \geqslant N^{-2/3+\mathfrak{c}}$, $\mathrm{Im}[m(0;z)] \asymp \mathrm{Im}[m_d(z)]$.

9.1 Free convolution. The asymptotic eigenvalue density of the matrix $H(t)$ is governed by the free additive convolution of the rescaled Kesten–McKay measure with the semicircle law at time $s = 1 - \mathrm{e}^{-t}$. We recall some properties of measures obtained by the free convolution with a semicircle distribution from [Bia97]. The semicircle density $\rho_{\mathrm{sc}}(x)$ is given by (2.6), and the semicircle density of variance s is $s^{-1/2}\rho_{\mathrm{sc}}(s^{-1/2}x)$. Given a probability measure μ on $\mathbb{R}$, we denote its free convolution with a semicircle distribution of variance s by μ_s. The Stieltjes transforms of μ and μ_s are given by $G_\mu(z) = \int \frac{\mathrm{d}\mu(x)}{x-z}$ and $G_{\mu_s}(z) = \int \frac{\mathrm{d}\mu_s(x)}{x-z}$ respectively. Then the following holds [Bia97].

(i) We denote the set $U_s = \{z \in \mathbb{C}_+ : \int \frac{\mathrm{d}\mu(x)}{|z-x|^2} < \frac{1}{s}\}$. Then $z \mapsto z - sG_\mu(z)$ is a homeomorphism from $\bar{U}_s$ to $\mathbb{C}_+ \cup \mathbb{R}$ and conformal from U_s to $\mathbb{C}_+$. We denote its inverse by $F_{\mu_s} : \mathbb{C}_+ \cup \mathbb{R} \mapsto \bar{U}_s$.

(ii) The Stieltjes transform of μ_s is characterized by $G_\mu(z) = G_{\mu_s}(z - sG_\mu(z))$, for any $z \in U_s$.

By the inversion formula for the Stieltjes transform, we deduce from (ii) that the density of μ_s is given by $\mathrm{d}\mu_s(x)/\mathrm{d}x = \mathrm{Im}[G_\mu(F_{\mu_s}(x))]/\pi$.

The asymptotic eigenvalue density of W is the semicircle density $\rho_{\mathrm{sc}}(x)$ and the asymptotic eigenvalue density of $\sqrt{1-\mathrm{e}^{-t}}W$ is the semicircle density at time $s = 1 - \mathrm{e}^{-t}$. The asymptotic eigenvalue density of $H(t)$ is the free convolution of rescaled Kesten–McKay law $\mu(\mathrm{d}x) = \mathrm{e}^{t/2}\rho_d(\mathrm{e}^{t/2}x)\mathrm{d}x$ at time e^{-t} and the semicircle density at time $1 - \mathrm{e}^{-t}$. We denote its density by $\rho_d(t;x)$ and its Stieltjes transform by $m_d(t;z) = \int \frac{\rho_d(t;x)}{x-z}\mathrm{d}x$. Since $G_\mu(\mathrm{e}^{-t/2}z) = \mathrm{e}^{t/2}m_d(z)$, we deduce from (i) and (ii) above that

$$m_d(t;\xi_d(t;z)) = \mathrm{e}^{t/2}m_d(z), \tag{9.9}$$

where

$$\xi_d(t;z) := \mathrm{e}^{-t/2}z - \mathrm{e}^{t/2}(1-\mathrm{e}^{-t})m_d(z) \tag{9.10}$$

is a homeomorphism from the set $\{z \in \mathbb{C}_+ : \int \frac{\rho_d(x)}{|x-z|^2}\mathrm{d}x \leqslant \frac{1}{e^t-1}\}$ to $\mathbb{C}_+ \cup \mathbb{R}$. We denote the functional inverse of $\xi_d(t;z)$ by $F_d(t;z)$ which is a homeomorphism from $\mathbb{C}_+ \cup \mathbb{R}$ to the set $\{z \in \mathbb{C}_+ : \int \frac{\rho_d(x)}{|x-z|^2}\mathrm{d}x \leqslant \frac{1}{e^t-1}\}$. Thus, $\rho_d(t;x) = \mathrm{Im}[m_d(F_d(t;x))]/\pi$. To find the support of the measure $\rho_d(t;x)$, we notice that there exists $z_d^{\pm}(t) \in \mathbb{R}$ such that $\{z \in \mathbb{C}_+ : \int \frac{\rho_d(x)}{|x-z|^2}\mathrm{d}x = \frac{1}{e^t-1}\}$ consists of the intervals $(-\infty, z_d^-(t)] \cup [z_d^+(t), \infty)$ and an arc $\mathcal{C}$ from $z_d^-(t)$ to $z_d^+(t)$. Those two endpoints $z_d^{\pm}(t) \in \mathbb{R}$ are the largest and smallest real solutions to

$$\int \frac{\rho_d(x)}{|x-z|^2}\mathrm{d}x = \frac{1}{e^t-1}, \tag{9.11}$$

and the function $F_d(t;x)$ maps the support of $\rho_d(t;x)$ to the arc $\mathcal{C}$. As a consequence, the right and left edges of the measure $\rho_d(t;x)$ are given by

$$E_d^{\pm}(t) = \xi_d(t; z_d^{\pm}(t)) = \mathrm{e}^{-t/2}z_d^{\pm}(t) - \mathrm{e}^{t/2}(1-\mathrm{e}^{-t})m_d(z_d^{\pm}(t)). \tag{9.12}$$

LEMMA 9.5. *Let $L_t = 2 + t/d$. Then the right and left edges of the measure $\rho_d(t;x)$ are given by*

$$E_d^{\pm}(t) = \pm\left(L_t + \mathrm{O}\left(t^3 + \frac{1}{d^3}\right)\right). \tag{9.13}$$

Proof. From (2.7) and (2.5) we recall that for any $z \in \mathbb{C}$,

$$z = -m_{\mathrm{sc}}(z) - \frac{1}{m_{\mathrm{sc}}(z)}, \quad m_d(z) = \frac{m_{\mathrm{sc}}(z)}{1 - m_{\mathrm{sc}}^2(z)/(d-1)}. \tag{9.14}$$

By taking the derivative we obtain

$$m_d'(z) = \frac{m_{\mathrm{sc}}^2(z)(1 + m_{\mathrm{sc}}^2(z)/(d-1))}{(1 - m_{\mathrm{sc}}^2(z))(1 - m_{\mathrm{sc}}^2(z)/(d-1))^2}. \tag{9.15}$$

We can solve for $z_d^{\pm}(t)$ using (9.11) and (9.15),

$$\begin{aligned} \frac{1}{e^t - 1} &= \int \frac{\rho_d(x)}{|x - z_d^{\pm}(t)|^2}\mathrm{d}x = m_d'(z_d^{\pm}(t)) \\ &= \frac{m_{\mathrm{sc}}^2(z_d^{\pm}(t))(1 + m_{\mathrm{sc}}^2(z_d^{\pm}(t))/(d-1))}{(1 - m_{\mathrm{sc}}^2(z_d^{\pm}(t)))(1 - m_{\mathrm{sc}}^2(z_d^{\pm}(t))/(d-1))^2}. \end{aligned} \tag{9.16}$$

In the regime $t \ll 1$ and $d \gg 1$, we can solve for $m_{\mathrm{sc}}(z_d^{\pm}(t))$ from (9.16) and get

$$m_{\mathrm{sc}}(z_d^{\pm}(t)) = \pm\left(1 - \frac{t}{2} - \frac{3t}{2d} - \frac{t^2}{8} + \mathrm{O}\left(t^3 + \frac{1}{d^3}\right)\right). \tag{9.17}$$

Using (9.14), (9.17) implies that

$$\begin{aligned} z_d^{\pm}(t) &= -m_{\mathrm{sc}}(z_d^{\pm}(t)) - \frac{1}{m_{\mathrm{sc}}(z_d^{\pm}(t))} = \mp\left(2 + \frac{t^2}{4} + \mathrm{O}\left(t^3 + \frac{1}{d^3}\right)\right), \\ m_d(z_d^{\pm}(t)) &= \frac{m_{\mathrm{sc}}(z_d^{\pm}(t))}{1 - m_{\mathrm{sc}}^2(z_d^{\pm}(t))/(d-1)} = \pm\left(1 - \frac{t}{2} + \frac{1}{d} + \mathrm{O}\left(t^2 + \frac{1}{d^2}\right)\right). \end{aligned} \tag{9.18}$$

Lemma 9.5 follows by plugging (9.18) into (9.12) and expanding the exponents to third order. □

9.2 Rigidity and edge universality of $H(t)$. In this section we collect some estimates on the Green's function $G(t;z)$ and the Stieltjes transform $m(t;z)$ of $H(t)$, and state the edge universality result for $H(t)$ when $t \gg N^{-1/3}$. All statements and estimates in this section follow directly from [BHY17, LY17, AH18].

First, using the rigidity estimates of Corollary 9.3 as input, the rigidity estimates on the Stieltjes transform $m(t;z)$ of $H(t)$ follow from [AH18, LY17].

PROPOSITION 9.6. *Fix a constant $\mathfrak{e} > 0$ and $N^{2/9+\mathfrak{e}} \leqslant d \leqslant N^{1/3-\mathfrak{e}}$. For an arbitrarily small constant $\mathfrak{c} > 0$ and any time $0 \leqslant t \ll 1$, with very high probability we have*

$$|\lambda_2(t) - L_t|, |\lambda_N(t) + L_t| \leqslant N^{-2/3+\mathfrak{c}}, \tag{9.19}$$

and uniformly for any $z = E + \mathrm{i}\eta$, with $-4 \leqslant E \leqslant 4$ and $\eta \geqslant N^{-2/3}$,

$$|m(t;z) - m_d(t;z)| \leqslant N^{\mathfrak{c}} \left(\frac{1}{N\eta} + \frac{d}{N\sqrt{\eta}} \right), \tag{9.20}$$

where $L_t = 2 + t/d$ and $m_d(t;z)$ is defined by (9.9) and $m(t;z)$ by (9.6).

Using the rigidity estimates from Corollary 9.3 and the estimates on the Green's function entries of $H(0)$ from Proposition 8.4 as input, the entrywise estimates on Green's function of $H(t)$ with $t > 0$ follow from an argument similar to the proof of [BHY17, Theorem 2.1]. We remark that in the statement of [BHY17, Theorem 2.1], it assumed that $\mathrm{Im}[m_0]$ is bounded from below and that $t \gg \eta_*$. However, in the proof of [BHY17, Theorem 2.3], these assumptions are only used to show that $|m_t(z) - m_{\mathrm{fc},t}(z)|$ is small. With the required estimate of $|m_t(z) - m_{\mathrm{fc},t}(z)|$ already established by (9.20), the remaining part of [BHY17, Theorem 2.1] does not use that $\mathrm{Im}[m_0]$ is bounded from below or that $t \gg \eta_*$. Therefore, with (9.20) given, the remaining proof of [BHY17, Theorem 2.1] applies and gives the following result on the entrywise estimates of Green's function of $H(t)$.

PROPOSITION 9.7. *Fix constant $\mathfrak{e} > 0$ and suppose that $N^{2/9+\mathfrak{e}} \leqslant d \leqslant N^{1/3-\mathfrak{e}}$. For an arbitrarily small constant $\mathfrak{c} > 0$ and any time $0 \leqslant t \ll 1$, with very high probability we have*

$$|G_{ij}(t;z) - \delta_{ij} m_d(t;z)| \leqslant N^{\mathfrak{c}} \left(\frac{1}{d^{1/2}} + \frac{1}{\sqrt{N\eta}} \right), \tag{9.21}$$

uniformly for any $z = E + \mathrm{i}\eta$, with $-4 \leqslant E \leqslant 4$ and $\eta \geqslant N^{-2/3}$.

The edge universality of $H(t)$ for $t \gg N^{-1/3}$ follows from the following result due to [LY17].

PROPOSITION 9.8. *Fix a constant $\mathfrak{e} > 0$ and suppose that $N^{2/9+\mathfrak{e}} \leqslant d \leqslant N^{1/3-\mathfrak{e}}$. Let $\mathfrak{d} > 0$ be a sufficiently small constant and set $t = N^{-1/3+\mathfrak{d}}$. Let $H(t)$ be as in (9.4), which has an eigenvalue $\lambda_1(t) = \mathrm{e}^{-t/2} d/\sqrt{d-1}$, and we denote its remaining eigenvalues by $\lambda_2(t) \geqslant \lambda_3(t) \geqslant \cdots \lambda_{N-1}(t) \geqslant \lambda_N(t)$. Fix $k \geqslant 1$ and $s_1, s_2, \cdots, s_k \in \mathbb{R}$. Then*

$$\begin{aligned} &\mathbb{P}_{H(t)} \left(N^{2/3}(\lambda_{i+1}(t) - L_t) \geqslant s_i, 1 \leqslant i \leqslant k \right) \\ &\quad = \mathbb{P}_{\mathrm{GOE}} \left(N^{2/3}(\mu_i - 2) \geqslant s_i, 1 \leqslant i \leqslant k \right) + \mathrm{o}(1), \end{aligned} \tag{9.22}$$

where $L_t = 2 + t/d$ is as defined in Lemma 9.5, and $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_N$ are the eigenvalues of a the Gaussian Orthogonal Ensemble. The analogous statement holds for the smallest eigenvalues.

REMARK 9.9. By an appropriate modification of the analysis of Dyson Brownian motion from [LY17], Proposition 9.8 also holds for the joint distribution of the k largest and smallest eigenvalues. In particular, this implies that, under the same assumption as in the previous proposition, the asymptotic joint distribution of $N^{2/3}(\lambda_2(t) - L_t, -\lambda_N(t) - L_t)$ is a pair of independent Tracy–Widom$_1$ distributions. Moreover, Proposition 9.11 below can be extended, by merely cosmetic changes, to also cover the joint distribution of k largest and smallest eigenvalues. Thus, we get the following extension of Theorem 1.3.

Theorem 9.10. *Suppose that $N^{2/9} \ll d \ll N^{1/3}$. Fix $k \geqslant 1$ and $s_1^{\pm}, s_2^{\pm}, \cdots, s_k^{\pm} \in \mathbb{R}$. Then*

$$\begin{aligned}&\mathbb{P}_H\Big(N^{2/3}(\lambda_{i+1} - 2) \\ &\quad \geqslant s_i^+, N^{2/3}(-\lambda_{N+1-i} - 2) \geqslant s_i^-, 1 \leqslant i \leqslant k\Big) \\ &\quad = \mathbb{P}_{\mathrm{GOE}}\Big(N^{2/3}(\mu_i - 2) \geqslant s_i^+, N^{2/3}(-\mu_{N+1-i} - 2) \geqslant s_i^-, 1 \leqslant i \leqslant k\Big) + \mathrm{o}(1).\end{aligned}$$

9.3 Green's function comparison. In this section we prove the following short-time comparison result for the edge eigenvalue statistics of $H(t)$.

PROPOSITION 9.11. *Fix a constant $\mathfrak{e} > 0$ and suppose that $N^{2/9+\mathfrak{e}} \leqslant d \leqslant N^{1/3-\mathfrak{e}}$. Let $\mathfrak{d} > 0$ be a sufficiently small constant and set $t = N^{-1/3+\mathfrak{d}}$. Let $H(t)$ be as in (9.4), which has an eigenvalue $\lambda_1(t) = \mathrm{e}^{-t/2} d/\sqrt{d-1}$, and we denote its remaining eigenvalues by $\lambda_2(t) \geqslant \lambda_3(t) \geqslant \cdots \lambda_{N-1}(t) \geqslant \lambda_N(t)$. Fix $k \geqslant 1$ and $s_1, s_2, \cdots, s_k \in \mathbb{R}$. Then*

$$\begin{aligned}&\mathbb{P}_H\Big(N^{2/3}(\lambda_{i+1}(0) - 2) \geqslant s_i, 1 \leqslant i \leqslant k\Big) \\ &\quad = \mathbb{P}_{H(t)}\Big(N^{2/3}(\lambda_{i+1}(t) - L_t) \geqslant s_i, 1 \leqslant i \leqslant k\Big) + \mathrm{o}(1),\end{aligned} \tag{9.23}$$

where $L_t = 2 + t/d$ is as defined in Lemma 9.5. The analogous statement holds for the smallest eigenvalues.

Before proving Proposition 9.11 we use it to conclude the proof of edge universality of random d-regular graphs, Theorem 1.3.

Proof of Theorem 1.3. Combine Propositions 9.8 and 9.11 . □

The rest of this section is devoted to the proof of Proposition 9.11. For any E, we define

$$\mathcal{N}_t(E) := |\{i \geqslant 2 : \lambda_i(t) \geqslant L_t + E\}|, \tag{9.24}$$

and we write $\mathcal{N}_0(E)$ as $\mathcal{N}(E)$. We take $\ell = N^{-2/3-\mathfrak{d}/9}$ and $\eta = N^{-2/3-\mathfrak{d}}$. Then with very high probability, from Propositions 9.6 and 9.7 , with (8.5) and (9.9), and with

the fact that $\eta \operatorname{Im}[m(t; L_t + \kappa + \mathrm{i}\eta)]$ is a monotone decreasing function of η (which is immediate from the spectral representation), for any $|\kappa| \leqslant N^{-2/3+\mathfrak{d}}$, we have

$$\operatorname{Im}[m(t; L_t + \kappa + \mathrm{i}\eta)] \leqslant N^{-1/3+C\mathfrak{d}}, \tag{9.25}$$

and similarly, since $\max_{ij} \eta |G_{ij}(t; L_t + \kappa + \mathrm{i}\eta)|$ is decreasing in η (see [BKY17, Lemma 2.1]),

$$\max_{ij} |G_{ij}(t; L_t + \kappa + \mathrm{i}\eta)| \leqslant N^{C\mathfrak{d}}, \quad \sum_{ij} |G_{ij}(t; L_t + \kappa + \mathrm{i}\eta)|^2 \leqslant N^{4/3+C\mathfrak{d}}. \tag{9.26}$$

Moreover, we have $|\partial_z m(t; z)| \leqslant \operatorname{Im}[m(t; z)]/\operatorname{Im}[z] \leqslant N^{1/3+C\mathfrak{d}}$ for any z with $|\operatorname{Re}[z] - L_t| \leqslant N^{-2/3+\mathfrak{d}}$ and $N^{-2/3-\mathfrak{d}} \leqslant \operatorname{Im}[z] \leqslant N^{-2/3}$. Therefore, by integrating from $z = 2 + \kappa + \mathrm{i}\eta$ to $z = 2 + \kappa + \mathrm{i}N^{-2/3}$, we get with very high probability

$$\begin{aligned} m(t; L_t + \kappa + \mathrm{i}\eta) &= m(t; L_t + \kappa + \mathrm{i}N^{-2/3}) + \mathrm{O}(N^{-1/3+C\mathfrak{d}}) \\ &= m_d(t; L_t + \kappa + \mathrm{i}N^{-2/3}) + \mathrm{O}(N^{-1/3+C\mathfrak{d}}) \\ &= m_d(t; E_d^+(t)) + \mathrm{O}(N^{-1/3+C\mathfrak{d}}) = -1 + \mathrm{O}(1/d) \end{aligned} \tag{9.27}$$

where we used Lemma 9.5, Proposition 9.6 and the square root behavior of the density $\rho_d(t; x)$:

$$|m_d(t; L_t + \kappa + \mathrm{i}N^{-2/3}) - m_d(t; E_d^+(t))| \asymp \sqrt{|L_t + \kappa + \mathrm{i}N^{-2/3} - E_d^+(t)|} \lesssim N^{-1/3+C\mathfrak{d}}, \tag{9.28}$$

which follows from [AH18, Proposition A.1].

Next, we define

$$\chi_E(x) = \mathbf{1}_{[E, N^{-2/3+\mathfrak{d}}]}(x - L_t), \quad \theta_\eta(x) := \frac{\eta}{\pi(x^2 + \eta^2)} = \frac{1}{\pi} \operatorname{Im} \frac{1}{x + \mathrm{i}\eta}. \tag{9.29}$$

We have that $\lambda_1(t) = \mathrm{e}^{-t/2} d \gg L_t$ and with very high probability it holds $\lambda_2(t) \leqslant L_t + N^{-2/3+\mathfrak{d}}$ by Proposition 9.6. By the same argument as in [KY13, Lemma 2.7], we get that

$$\operatorname{Tr}(\chi_{E+\ell} * \theta_\eta)(H(t)) - N^{-\mathfrak{d}/9} \leqslant \mathcal{N}_t(E) \leqslant \operatorname{Tr}(\chi_{E-\ell} * \theta_\eta)(H(t)) + N^{-\mathfrak{d}/9} \tag{9.30}$$

with very high probability. Let $K_i : \mathbb{R} \mapsto [0, 1]$ be a monotonic smooth function satisfying,

$$K_i(x) = \begin{cases} 0 \text{ if } x \leqslant i - 2/3, \\ 1 \text{ if } x \geqslant i - 1/3. \end{cases} \tag{9.31}$$

Since $\mathbf{1}_{\mathcal{N}_t(E) \geqslant i} = K_i(\mathcal{N}_t(E))$ and K_i is monotonically increasing, we have

$$\begin{aligned} K_i\left(\operatorname{Tr}(\chi_{E+\ell} * \theta_\eta)(H(t))\right) + \mathrm{O}(N^{-\mathfrak{d}/9}) &\leqslant \mathbf{1}_{\mathcal{N}_t(E) \geqslant i} \\ &\leqslant K_i\left(\operatorname{Tr}(\chi_{E-\ell} * \theta_\eta)(H(t))\right) + \mathrm{O}(N^{-\mathfrak{d}/9}). \end{aligned} \tag{9.32}$$

In this way we can express the locations of eigenvalues in terms of integrals of the Stieltjes transform of the empirical eigenvalue densities:

$$\begin{aligned}
&\mathbb{E}_{H(t)}\left[\prod_{i=1}^{k} K_i\left(\operatorname{Im}\left[\frac{N}{\pi}\int_{s_i N^{-2/3}+\ell}^{N^{-2/3+\mathfrak{d}}} m(t; L_t + y + \mathrm{i}\eta)\,\mathrm{d}y\right]\right)\right] + \mathrm{O}\left(N^{-\mathfrak{d}/9}\right) \\
&\quad \leqslant \mathbb{P}_{H(t)}\left(N^{2/3}(\lambda_{i+1}(t) - L_t) \geqslant s_i, 1 \leqslant i \leqslant k\right) = \mathbb{E}\left[\prod_{i=1}^{k} \mathbf{1}_{\mathcal{N}_t(s_i N^{-2/3}) \geqslant i}\right] \\
&\quad \leqslant \mathbb{E}_{H(t)}\left[\prod_{i=1}^{k} K_i\left(\operatorname{Im}\left[\frac{N}{\pi}\int_{s_i N^{-2/3}-\ell}^{N^{-2/3+\mathfrak{d}}} m(t; L_t + y + \mathrm{i}\eta)\,\mathrm{d}y\right]\right)\right] + \mathrm{O}\left(N^{-\mathfrak{d}/9}\right).
\end{aligned} \tag{9.33}$$

For the product of the functions of Stieltjes transform, we have the following comparison theorem.

PROPOSITION 9.12. *Fix constant $\mathfrak{e} > 0$ and $N^{2/9+\mathfrak{e}} \leqslant d \leqslant N^{1/3-\mathfrak{e}}$. Let $\mathfrak{d} > 0$ be sufficiently small and set $t = N^{-1/3+\mathfrak{d}}$ and $\eta = N^{-2/3-\mathfrak{d}}$. Let $F : \mathbb{R}^k \mapsto \mathbb{R}$ be a fixed smooth test function. Then for $E_1, E_2, \cdots, E_k = \mathrm{O}(N^{-2/3})$ we have*

$$\begin{aligned}
&\mathbb{E}_H\left[F\left(\left\{\operatorname{Im}\left[N\int_{E_i}^{N^{-2/3+\mathfrak{d}}} m(0; L_0 + y + \mathrm{i}\eta)\,\mathrm{d}y\right]\right\}_{i=1}^{k}\right)\right] \\
&\quad = \mathbb{E}_{H(t)}\left[\left\{F\left(\operatorname{Im}\left[N\int_{E_i}^{N^{-2/3+\mathfrak{d}}} m(t; L_t + y + \mathrm{i}\eta)\,\mathrm{d}y\right]\right\}_{i=1}^{k}\right)\right] \\
&\qquad + \mathrm{O}\left(\frac{N^{2/3+C\mathfrak{d}}t}{d^{3/2}} + \frac{N^{1/3+C\mathfrak{d}}t}{d^{1/2}}\right).
\end{aligned} \tag{9.34}$$

Proof of Proposition 9.11. Since $d \geqslant N^{2/9+\mathfrak{e}}$ and $t = N^{-1/3+\mathfrak{d}}$, for $\mathfrak{d}$ small, the error terms in (9.34) are of order $\mathrm{O}(N^{-\mathfrak{d}})$. By combining (9.33) and (9.34), we thus get

$$\begin{aligned}
&\mathbb{P}_{H(t)}\left(N^{2/3}(\lambda_{i+1}(t) - L_t) \geqslant s_i + 2N^{2/3}\ell, 1 \leqslant i \leqslant k\right) \\
&\quad \leqslant \mathbb{E}_{H(t)}\left[\prod_{i=1}^{k} K_i\left(\operatorname{Im}\left[\frac{N}{\pi}\int_{s_i N^{-2/3}+\ell}^{N^{-2/3+\mathfrak{d}}} m(t; L_t + y + \mathrm{i}\eta)\,\mathrm{d}y\right]\right)\right] + \mathrm{O}\left(N^{-\mathfrak{d}/9}\right) \\
&\quad \leqslant \mathbb{P}_H\left(N^{2/3}(\lambda_{i+1}(0) - L_0) \geqslant s_i, 1 \leqslant i \leqslant k\right) + \mathrm{O}\left(N^{-\mathfrak{d}/9}\right) \\
&\quad \leqslant \mathbb{E}_{H(t)}\left[\prod_{i=1}^{k} K_i\left(\operatorname{Im}\left[\frac{N}{\pi}\int_{s_i N^{-2/3}-\ell}^{N^{-2/3+\mathfrak{d}}} m(t; L_t + y + \mathrm{i}\eta)\,\mathrm{d}y\right]\right)\right] + \mathrm{O}\left(N^{-\mathfrak{d}/9}\right) \\
&\quad \leqslant \mathbb{P}_{H(t)}\left(N^{2/3}(\lambda_{i+1}(t) - L_t) \geqslant s_i - 2N^{2/3}\ell, 1 \leqslant i \leqslant k\right) + \mathrm{O}(N^{-\mathfrak{d}/9}).
\end{aligned} \tag{9.35}$$

Since $N^{2/3}\ell = N^{-\mathfrak{d}/9} \ll 1$, (9.23) follows. □

Proof of Proposition 9.12. For simplicity of notation, we only prove the case $k = 1$; the general case can be proved in the same way. Let

$$X_t = \operatorname{Im}\left[N \int_E^{N^{-2/3+\mathfrak{d}}} m(t; L_t + y + \mathrm{i}\eta)\,\mathrm{d}y\right]. \tag{9.36}$$

We shall prove that

$$|\mathbb{E}[F(X_t)] - \mathbb{E}[F(X_0)]| \leqslant N^{C\mathfrak{d}}\left(\frac{N^{2/3}t}{d^{3/2}} + \frac{N^{1/3}t}{d^{1/2}}\right). \tag{9.37}$$

The derivative of $\mathbb{E}[F(X_t)]$ with respect to the time t is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[F(X_t)] &= \mathbb{E}\left[F'(X_t)\frac{\mathrm{d}X_t}{\mathrm{d}t}\right] \\ &= \mathbb{E}\left[F'(X_t)\operatorname{Im}\int_E^{N^{-2/3+\mathfrak{d}}}\left(N\sum_{ij}\dot{H}_{ij}(t)\frac{\partial m}{\partial H_{ij}} + \dot{L}_t\sum_{ij}G_{ij}^2\right)\mathrm{d}y\right] \\ &= \mathbb{E}\left[F'(X_t)\operatorname{Im}\int_E^{N^{-2/3+\mathfrak{d}}}\left(-\sum_{ija}\dot{H}_{ij}(t)G_{ai}G_{aj} + \dot{L}_t\sum_{ij}G_{ij}^2\right)\mathrm{d}y\right]\end{aligned} \tag{9.38}$$

where we abbreviate $G(t; L_t + y + \mathrm{i}\eta)$ and $m(t; L_t + y + \mathrm{i}\eta)$ by G and m respectively. In the following, we estimate the right-hand side. By definition,

$$\dot{H}_{ij}(t) = -\frac{1}{2}\mathrm{e}^{-t/2}H_{ij} + \frac{\mathrm{e}^{-t}}{2\sqrt{1-\mathrm{e}^{-t}}}W_{ij}. \tag{9.39}$$

In the following, we use the notation ∂_{ij}^{kl} applied to a function of $H(t)$, such as G or m, to denote the directional derivative (4.5) with respect to $H(t)$. For any directional derivative ∂ we therefore have

$$\frac{\partial}{\partial H}F(H(t)) = \mathrm{e}^{-t/2}\,\partial F(H(t)), \qquad \frac{\partial}{\partial W}F(H(t)) = \sqrt{1-\mathrm{e}^{-t}}\,\partial F(H(t)).$$

Plugging (9.39) into (9.38), and using the Gaussian integration by parts (9.3), we therefore obtain

$$\begin{aligned}-\sum_{ija}\mathbb{E}\left[\dot{H}_{ij}(t)F'(X_t)G_{ai}G_{ja}\right] &= \frac{1}{2}\sum_{ija}\mathbb{E}\left[\left(\mathrm{e}^{-t/2}H_{ij} - \frac{\mathrm{e}^{-t}}{\sqrt{1-\mathrm{e}^{-t}}}W_{ij}\right)F'(X_t)G_{ai}G_{ja}\right] \\ &= \frac{\mathrm{e}^{-t/2}}{2\sqrt{d-1}}\sum_{ija}\mathbb{E}\left[A_{ij}(F'(X_t)G_{ai}G_{ja})\right] - \frac{\mathrm{e}^{-t}}{2N^3}\sum_{ijkla}\mathbb{E}[\partial_{ij}^{kl}(F'(X_t)G_{ai}G_{ja})].\end{aligned} \tag{9.40}$$

To estimate the first term, we apply Corollary 3.2 with the random variable $F_{ij} = F'(X_t)G_{ai}G_{ja}$ in (3.7). Since $\mathcal{C}_{ij}(F_{ij}, A) \prec |G_{ai}G_{ja}|$, using the Ward identity (4.2) and $\sum_{ij} A_{ij} = Nd$, the resulting error term is bounded by

$$\frac{1}{\sqrt{d-1}} \sum_{ija} \frac{d\mathbb{E}[A_{ij}\mathcal{C}_{ij}(F_{ij}, A)]}{N} \prec \frac{d^{1/2}}{N} \sum_{ija} \mathbb{E}[A_{ij}|G_{ai}G_{ja}|] \prec d^{3/2} \frac{\mathbb{E}[\operatorname{Im} m]}{\eta} \leqslant Cd^{3/2}N^{1/3+C\mathfrak{d}}; \tag{9.41}$$

in the last inequality, we used (9.25) and that $\eta = N^{-2/3-\mathfrak{d}}$ and $|\kappa| \leqslant N^{-2/3+\mathfrak{d}}$. In summary, Taylor expanding the discrete derivative in (3.7) and noting that the first term on the right-hand side of (3.7) vanishes by $\sum_i G_{ai} = 0$, we find that (9.40) is bounded by

$$\begin{aligned} &\frac{\mathrm{e}^{-t}}{2Nd(d-1)} \sum_{ijkla} \mathbb{E}\left[A_{ik}A_{jl}\partial_{ij}^{kl}(F'(X_t)G_{ai}G_{ja})\right] - \frac{\mathrm{e}^{-t}}{2N^3} \sum_{ijkla} \mathbb{E}[\partial_{ij}^{kl}(F'(X_t)G_{ai}G_{ja})] \\ &+ \sum_{n=2}^{\mathfrak{b}} \frac{\mathrm{e}^{-(n+1)t/2}}{2Nd(d-1)^{(n+1)/2}n!} \sum_{ijkla} \mathbb{E}\left[A_{ik}A_{jl}(\partial_{ij}^{kl})^n(F'(X_t)G_{ai}G_{ja})\right] \\ &+ \mathrm{O}\left(d^{3/2}N^{1/3+C\mathfrak{d}}\right) \end{aligned} \tag{9.42}$$

for some constant $\mathfrak{b}$. The terms in (9.42) are estimated in the following two claims; we postpone their proofs to the end of this section.

Claim 9.13. *For the first two terms in* (9.42) *we have,*

$$\begin{aligned} &\frac{\mathrm{e}^{-t}}{2Nd(d-1)} \sum_{ijkla} \mathbb{E}\left[A_{ik}A_{jl}\partial_{ij}^{kl}(F'(X_t)G_{ai}G_{ja})\right] - \frac{\mathrm{e}^{-t}}{2N^3} \sum_{ijkla} \mathbb{E}[\partial_{ij}^{kl}(F'(X_t)G_{ai}G_{ja})] \\ &= \frac{3}{d} \sum_{ij} \mathbb{E}[F'(X_t)G_{ij}^2] + \mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right). \end{aligned} \tag{9.43}$$

Claim 9.14. *For any $n \geqslant 2$, let,*

$$J_n := \frac{\mathrm{e}^{-(n+1)t/2}}{2Nd(d-1)^{(n+1)/2}n!} \sum_{ijkla} \mathbb{E}\left[A_{ik}A_{jl}(\partial_{ij}^{kl})^n(F'(X_t)G_{ai}G_{ja})\right]. \tag{9.44}$$

Then, we have the estimates

$$J_2 = -\frac{12}{d} \sum_{ij} \mathbb{E}[F'(X_t)G_{ij}^2] + \mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}} + \frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right), \tag{9.45}$$

$$J_3 = \frac{8}{d}\sum_{ij}\mathbb{E}[F'(X_t)G_{ij}^2] + \mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}} + \frac{N^{1+C\mathfrak{d}}}{d}\right), \tag{9.46}$$

and for any $n \geqslant 4$, $J_n = \mathrm{O}(N^{4/3+C\mathfrak{d}}/d^{(n-1)/2})$.

It follows from Claims 9.13 and 9.14 and (9.42) that (9.40) can be estimated by

$$\begin{aligned}-\sum_{ija}\mathbb{E}\left[\dot{H}_{ij}(t)F'(X_t)G_{ai}G_{ja}\right] &= \frac{3}{d}\sum_{ij}\mathbb{E}[F'(X_t)G_{ij}^2] - \frac{12}{d}\sum_{ij}\mathbb{E}[F'(X_t)G_{ij}^2] \\ &\quad + \frac{8}{d}\sum_{ij}\mathbb{E}[F'(X_t)G_{ij}^2] + \mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}} + \frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right) \\ &= -\frac{1}{d}\sum_{ij}\mathbb{E}[F'(X_t)G_{ij}^2] + \mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}} + \frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right).\end{aligned} \tag{9.47}$$

From Lemma 9.5, we recall that $L_t = 2 + t/d$, i.e., $\dot{L}_t = 1/d$. Therefore plugging (9.47) into (9.38), the two terms in (9.38) cancel up to an error, and we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[F(X_t)] &= \mathbb{E}\left[\int_E^{N^{-2/3+\mathfrak{d}}} \mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}} + \frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right)\mathrm{d}y\right] \\ &= \mathrm{O}\left(\frac{N^{2/3+C\mathfrak{d}}}{d^{3/2}} + \frac{N^{1/3+C\mathfrak{d}}}{d^{1/2}}\right),\end{aligned} \tag{9.48}$$

and (9.37) follows. This finishes the proof of Proposition 9.12. □

Proof of Claim 9.13. We have

$$\partial_{ij}^{kl}(F'(X_t)G_{ai}G_{ja}) = \partial_{ij}^{kl}(F'(X_t))G_{ai}G_{ja} + F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja}). \tag{9.49}$$

For the first term in (9.49), with very high probability we have

$$|\partial_{ij}^{kl}F'(X_t)| = \left|F''(X_t)\,\mathrm{Im}\left[N\int_E^{N^{-2/3+\mathfrak{d}}}\partial_{ij}^{kl}m(t; L_t + y + \mathrm{i}\eta)\,\mathrm{d}y\right]\right| \leqslant N^{-1/3+C\mathfrak{d}}, \tag{9.50}$$

where we used (4.14) and (9.25). Therefore, the sum arising from the first term in (9.49) can be estimated as

$$\begin{aligned}&\frac{\mathrm{e}^{-t}}{2Nd(d-1)}\sum_{ijkla}\mathbb{E}\left[A_{ik}A_{jl}\partial_{ij}^{kl}(F'(X_t))G_{ai}G_{ja}\right] \\ &= \frac{\mathrm{e}^{-t}}{2Nd(d-1)}\sum_{ijkla}\frac{d^2}{N^2}\mathbb{E}\left[\partial_{ij}^{kl}(F'(X_t))G_{ai}G_{ja}\right] + \mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right) \\ &= \frac{\mathrm{e}^{-t}}{2N^3}\sum_{ijkla}\mathbb{E}\left[\partial_{ij}^{kl}(F'(X_t))G_{ai}G_{ja}\right] + \mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right).\end{aligned} \tag{9.51}$$

where we used Corollary 3.4, (9.50) and that with very high probability $\sum_{ij} |(G^2)_{ij}| = N \operatorname{Im}[m(t; L_t + y + \mathrm{i}\eta)]/\eta \leqslant N^{4/3+C\mathfrak{d}}$ from (9.25). For the sum corresponding to the second term in (9.49), we use the notation $A(t) = \sqrt{d-1}H(t)$ and write

$$\begin{aligned}
&\frac{\mathrm{e}^{-t}}{2Nd(d-1)} \sum_{ijkla} \mathbb{E}\left[A_{ik}A_{jl}F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja})\right] \\
&\quad = \frac{1}{2Nd(d-1)} \sum_{ijkla} \mathbb{E}\left[A_{ik}(t)A_{jl}(t)F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja})\right] \\
&\qquad - \frac{\mathrm{e}^{-t/2}\sqrt{1-\mathrm{e}^{-t}}}{2Nd(d-1)^{1/2}} \sum_{ijkla} \mathbb{E}\left[(A_{ik}W_{jl} + W_{ik}A_{jl})F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja})\right] \\
&\qquad - \frac{(1-\mathrm{e}^{-t})}{2Nd} \sum_{ijkla} \mathbb{E}\left[W_{ik}W_{jl}F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja})\right].
\end{aligned} \tag{9.52}$$

We can estimate the second and third term on the right-hand side of (9.52) using Lemma 9.1, e.g.

$$\begin{aligned}
&\frac{\mathrm{e}^{-t/2}\sqrt{1-\mathrm{e}^{-t}}}{2Nd(d-1)^{1/2}} \sum_{ijkla} \mathbb{E}\left[W_{ik}A_{jl}F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja})\right] \\
&\quad = \frac{\mathrm{e}^{-t/2}(1-\mathrm{e}^{-t})}{2N^4d(d-1)^{1/2}} \sum_{ijkli'k'a} \mathbb{E}\left[A_{jl}\partial_{ik}^{i'k'}(F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja}))\right] \\
&\quad = \mathrm{O}\left(\frac{(1-\mathrm{e}^{-t})}{2N^4d(d-1)^{1/2}} \sum_{ijkli'k'} \mathbb{E}\left[A_{jl}\operatorname{Im}[m]/\eta\right]\right) = \mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right),
\end{aligned} \tag{9.53}$$

where in the third line we used (4.12) and (9.25). In the same way, we in fact have that the second and third term on the right-hand side of (9.52) are all bounded by $\mathrm{O}(N^{1+C\mathfrak{d}}/d^{1/2})$.

In the following we estimate the first term on the right-hand side of (9.52). These are terms in the form $\sum_{ijkla} A_{ik}(t)A_{jl}(t)F'(X_t)\times\{$monomial of Green's function entries

, where, we recall, the Green's function is that of $H(t) = (d-1)^{-1/2}A(t)$. For them we can

(i) sum over indices which appear only once and use the relations $\sum_i A_{ij}(t) = \sum_j A_{ij}(t) = \mathrm{e}^{-t/2}d$ and $\sum_i G_{ij} = \sum_j G_{ij} = 0$ to get expressions involving $\operatorname{Tr} G^2$, $\operatorname{Tr} G^3$, or $A(t)G$;
(ii) simplify the expressions using the identity $A(t)G = GA(t) = (d-1)^{1/2}(H(t) - z + z)G = (d-1)^{1/2}(zG + P_\perp)$;
(iii) estimate the final expressions using (4.11), (4.12), (9.27), $z = 2 + \mathrm{O}(N^{-1/3+\mathfrak{d}})$, $\max_{i\neq j}|G_{ij}| \prec 1/\sqrt{d}$ and $\max_i |G_{ii}+1| \prec 1/\sqrt{d}$ from Proposition 9.7.

Using the above procedure, we get that

$$\begin{aligned}&\sum_{ijkla} A_{ik}(t)A_{jl}(t)\partial_{ij}^{kl}(G_{ai}G_{ja})\\&\quad=-2\mathrm{e}^{-t}d^2(\operatorname{Tr}G^3+Nm\operatorname{Tr}G^2)+4(d-1)N\operatorname{Tr}G^2+\mathrm{O}(N^{7/3+C\mathfrak{d}}),\end{aligned}\tag{9.54}$$

and

$$\sum_{ijkla}\partial_{ij}^{kl}(G_{ai}G_{ja})=-2N^2(\operatorname{Tr}G^3+Nm\operatorname{Tr}G^2).\tag{9.55}$$

It follows by combining (9.51), (9.54) and (9.55) that

$$\begin{aligned}&\frac{\mathrm{e}^{-t}}{2Nd(d-1)}\sum_{ijkla}\mathbb{E}\left[A_{ik}A_{jl}\partial_{ij}^{kl}(F'(X_t)G_{ai}G_{ja})\right]-\frac{\mathrm{e}^{-t}}{2N^3}\sum_{ijkla}\mathbb{E}[\partial_{ij}^{kl}(F'(X_t)G_{ai}G_{ja})]\\&\quad=\frac{1}{2Nd(d-1)}\sum_{ijkla}\mathbb{E}\left[A_{ik}(t)A_{jl}(t)F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja})\right]\\&\qquad-\frac{\mathrm{e}^{-t}}{2N^3}\sum_{ijkla}\mathbb{E}[F'(X_t)\partial_{ij}^{kl}(G_{ai}G_{ja})]+\mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right)\\&\quad=-\frac{\mathrm{e}^{-t}}{(d-1)N}\mathbb{E}\left[F'(X_t)(\operatorname{Tr}G^3+Nm\operatorname{Tr}G^2)\right]+\frac{2}{d}\mathbb{E}[F'(X_t)\operatorname{Tr}G^2]+\mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right)\\&\quad=\frac{3}{d}\mathbb{E}[F'(X_t)\operatorname{Tr}G^2]+\mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right),\end{aligned}$$

where in the last line we used (4.11) and (9.27). □

Proof of Claim 9.14 For (9.45), similarly to (9.50), we have $|(\partial_{ij}^{kl})^2F'(X_t)|\leqslant N^{-1/3+C\mathfrak{d}}$ with very high probability and

$$J_2=\frac{\mathrm{e}^{-3t/2}}{4Nd(d-1)^{3/2}}\sum_{ijkla}\mathbb{E}\left[A_{ik}A_{jl}F'(X_t)(\partial_{ij}^{kl})^2(G_{ai}G_{ja})\right]+\mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right).\tag{9.56}$$

Thanks to Proposition 9.7, we have $\max_{i\neq j}|G_{ij}|\prec 1/\sqrt{d}$. Those terms from $(\partial_{ij}^{kl})^2(G_{ai}G_{ja})$ that contain four off-diagonal terms yield a contribution of the form

$$\begin{aligned}&\frac{\mathrm{e}^{-3t/2}}{4Nd(d-1)^{3/2}}\sum_{ijkla}\mathbb{E}\left[A_{ik}A_{jl}F'(X_t)\{\text{terms with 4 off-diagonal terms}\}\right]\\&\quad=\mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}}\right).\end{aligned}\tag{9.57}$$

The leading contributions are from those terms from $(\partial_{ij}^{kl})^2(G_{ai}G_{ja})$ which contain two or three off-diagonal Green's function entries. By the same estimate as in (9.52), we have

$$\begin{aligned}&\frac{\mathrm{e}^{-3t/2}}{4Nd(d-1)^{3/2}}\sum_{ijkla}\mathbb{E}\left[A_{ik}A_{jl}F'(X_t)\{\text{terms with } \leqslant 3 \text{ off-diagonal terms}\}\right]\\&\quad=\frac{\mathrm{e}^{-t/2}}{4Nd(d-1)^{3/2}}\sum_{ijkla}\mathbb{E}\left[A_{ik}(t)A_{jl}(t)F'(X_t)\{\text{terms with } \leqslant 3 \text{ off-diagonal terms}\}\right]\\&\qquad+\mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d}\right).\end{aligned}\tag{9.58}$$

Those terms in (9.58) can be treated by the same procedure as described in the proof of Claim 9.13, and we get

$$(9.58)=-\frac{12}{d}\sum_{ij}\mathbb{E}[F'(X_t)G_{ij}^2]+\mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}}+\frac{N^{1+C\mathfrak{d}}}{d^{1/2}}\right).\tag{9.59}$$

The claim (9.45) follows from combining (9.56), (9.57), (9.58) and (9.59).

In the following we prove (9.46). Similarly to (9.50), we have $|(\partial_{ij}^{kl})^3F'(X_t)|\leqslant N^{-1/3+C\mathfrak{d}}$ with very high probability and

$$J_3=\frac{\mathrm{e}^{-2t}}{12Nd(d-1)^2}\sum_{ijkla}\mathbb{E}\left[A_{ik}A_{jl}F'(X_t)(\partial_{ij}^{kl})^3(G_{ai}G_{ja})\right]+\mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d}\right).\tag{9.60}$$

Thanks to Proposition 9.7, we have $\max_{i\neq j}|G_{ij}|\prec 1/\sqrt{d}$. Those terms from $(\partial_{ij}^{kl})^3(G_{ai}G_{ja})$ which contain at least three off-diagonal terms yield a contribution of the form

$$\begin{aligned}&\frac{\mathrm{e}^{-2t}}{12Nd(d-1)^2}\sum_{ijkla}\mathbb{E}\left[A_{ik}A_{jl}F'(X_t)\{\text{terms with}\right.\\&\qquad\left.\geqslant 3 \text{ off-diagonal terms}\}\right]=\mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}}\right).\end{aligned}\tag{9.61}$$

The leading contribution is from those terms that contain exactly two off-diagonal terms. By the same estimate as in (9.52), we have

$$\begin{aligned}&\frac{\mathrm{e}^{-2t}}{12Nd(d-1)^2}\sum_{ijkla}\mathbb{E}\left[A_{ik}A_{jl}F'(X_t)\{\text{terms with 2 off-diagonal terms}\}\right]\\&\quad=\frac{\mathrm{e}^{-t}}{12Nd(d-1)^2}\sum_{ijkla}\mathbb{E}\left[A_{ik}(t)A_{jl}(t)F'(X_t)\{\text{terms with 2 off-diagonal terms}\}\right]\\&\qquad+\mathrm{O}\left(\frac{N^{1+C\mathfrak{d}}}{d^{3/2}}\right).\end{aligned}\tag{9.62}$$

We can estimate above terms using the procedure as described in the proof of Claim 9.13, and get

$$\begin{aligned}&\frac{\mathrm{e}^{-t}}{12Nd(d-1)^2}\sum_{ijkla}\mathbb{E}\left[A_{ik}(t)A_{jl}(t)F'(X_t)\{\text{terms with 2 off-diagonal terms}\}\right]\\&=\frac{8}{d}\mathbb{E}\left[F'(X_t)\operatorname{Tr}G^2\right]+\mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{3/2}}\right).\end{aligned}\tag{9.63}$$

The claim (9.46) follows from combining (9.60), (9.61), (9.62) and (9.63).

For fixed $n\geqslant 4$, we have the trivial bound

$$\begin{aligned}|J_n|&\lesssim\frac{1}{Nd^{(n+3)/2}}\sum_{ijkla}\mathbb{E}[A_{ik}A_{jl}(\partial_{ij}^{kl})^n(F'(X_t)G_{ai}G_{ja})]\\&\lesssim\frac{1}{Nd^{(n+3)/2}}\sum_{ijkl}\mathbb{E}[A_{ik}A_{jl}\operatorname{Im}[m]/\eta]=\mathrm{O}\left(\frac{N^{4/3+C\mathfrak{d}}}{d^{(n-1)/2}}\right).\end{aligned}$$

□

Acknowledgments

The work of J.H. is supported by the Institute for Advanced Study. A.K. gratefully acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 715539_RandMat) and from the Swiss National Science Foundation through the SwissMAP grant. The work of H.-T.Y. is partially supported by NSF Grants DMS-1606305 and DMS-1855509, and a Simons Investigator award.

References

[AH18] A. Adhikari and J. Huang. Dyson Brownian motion for general β and potential at the edge. Preprint arXiv:1810.08308 (2018)

[AC15] B. Adlam and Z. Che. Spectral statistics of sparse random graphs with a general degree distribution. Preprint 1509.03368 (2015)

[AW15] M. Aizenman and S. Warzel. Random operators. In: *Disorder Effects on Quantum Spectra and Dynamics. Graduate Studies in Mathematics*, Vol. 168. American Mathematical Society, Providence, RI (2015)

[ADK19] J. Alt, R. Ducatez, and A. Knowles. Extremal eigenvalues of critical Erdős–Rényi graphs. Preprint 1905.03243 (2019)

[BS19] Á. Backhausz and B. Szegedy. On the almost eigenvectors of random regular graphs. *Ann. Probab.*, (3)47 (2019), 1677–1725.

[BCZ18] A. Basak, N. Cook, and O. Zeitouni. Circular law for the sum of random permutation matrices. *Electron. J. Probab.*, 23:Paper No. 33 (2018), 51

[BHKY17] R. Bauerschmidt, J. Huang, A. Knowles, and H.-T. Yau. Bulk eigenvalue statistics for random regular graphs. *Ann. Probab.*, (6A)45 (2017), 3626–3663.

[BHY19] R. Bauerschmidt, J. Huang, and H.-T. Yau. Local Kesten-McKay law for random regular graphs. *Comm. Math. Phys.*, (2)369 (2019), 523–636

[BKY17] R. Bauerschmidt, A. Knowles, and H.-T. Yau. Local semicircle law for random regular graphs. *Comm. Pure Appl. Math.*, (10)70 (2017), 1898–1960

[BBK17] F. Benaych-Georges, C. Bordenave, and A. Knowles. Spectral radii of sparse random matrices. Preprint 1704.02945 (2017)

[Bia97] P. Biane. On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.*, (3)46 (1997), 705–718

[BEKYY14] 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.*, (33)19 (2014), 53.

[Bor15] C. Bordenave. A new proof of Friedman's second eigenvalue Theorem and its extension to random lifts. *To appear in Ann. Sci. Éc. Norm. Supér.* (2015)

[BC12] C. Bordenave and D. Chafa¨i. Around the circular law. *Probability Surveys* 9 (2012), 1–89

[BEYY16] P. Bourgade, L. Erdős, H.-T. Yau, and J. Yin. Fixed energy universality for generalized Wigner matrices. *Comm. Pure Appl. Math.*, (10)69 (2016), 1815–1881

[BHY17] P. Bourgade, J. Huang, and H.-T. Yau. Eigenvector statistics of sparse random matrices. *Electron. J. Probab.*, 22:Paper No. 64 (2017), 38

[BY16] P. Bourgade and H.-T. Yau. The eigenvector moment flow and local quantum unique ergodicity. *Comm. Math. Phys.*, (1)350 (2016) 231–278.

[Coh16] M.B. Cohen. Ramanujan graphs in polynomial time. In: *57th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2016*. IEEE Computer Society , Los Alamitos, CA (2016), pp. 276–281

[Coo15] N. Cook. On the singularity of adjacency matrices for random regular digraphs. *Probab. Theory Related Fields*, 167 (2015), 143–200

[CGJ18] N. Cook, L. Goldstein, and T. Johnson. Size biased couplings and the spectral gap for random regular graphs. *Ann. Probab.*, (1)46 (2018), 72–125

[Cok15] N.A. Cook. The circular law for signed random regular digraphs. Preprint 1508.00208 (2015)

[Coo17] N.A. Cook. The circular law for random regular digraphs. Preprint 1703.05839 (2017)

[DJPP13] I. Dumitriu, T. Johnson, S. Pal, and E. Paquette. Functional limit theorems for random regular graphs. *Probab. Theory Related Fields*, (3-4)156 (2013), 921–975

[Ede97] A. Edelman. The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law. *J. Multivariate Anal.*, (2)60 (1997), 203–232

[Elo09] Y. Elon. Gaussian waves on the regular tree. Preprint 0907.5065 (2009)

[EY17] L. Erdős and H.-T. Yau. *A Dynamical Approach to Random Matrix Theory. Courant Lecture Notes in Mathematics*, Vol. 28. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI (2017)

[EKYY12] 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.*, (3)314 (2012), 587–640

[EKYY13] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. Spectral statistics of Erdős-Rényi graphs I: local semicircle law. *Ann. Probab.*, (3B)41 (2013), 2279–2375

[EPRSY10] L. Erdős, S. Péché, J.A. Ram´irez, B. Schlein, and H.-T. Yau. Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.*, (7)63 (2010), 895–925

[ERSY10] L. Erdős, J.A. Ramirez, B. Schlein, and H.-T. Yau. Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation. *Electron. J. Probab.*, (18)15 (2010), 526–603

[ESY11] L. Erdős, B. Schlein, and H.-T. Yau. Universality of random matrices and local relaxation flow. *Invent. Math.*, (1)185 (2011), 75–119

[EY15] L. Erdős and H.-T. Yau. Gap universality of generalized Wigner and β-ensembles. *J. Eur. Math. Soc. (JEMS)*, (8)17 (2015), 1927–2036

[EYY12] L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields*, (1-2)154 (2012), 341–407

[EYY12] L. Erdős, H.-T. Yau, and J. Yin. Rigidity of eigenvalues of generalized Wigner matrices. *Adv. Math.*, (3)229 (2012), 1435–1515

[FS10] O.N. Feldheim and S. Sodin. A universality result for the smallest eigenvalues of certain sample covariance matrices. *Geom. Funct. Anal.*, (1)20 (2010), 88–123

[Fri08] J. Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.*, (910)195 (2008), viii+100

[FKS89] J. Friedman, J. Kahn, and E. Szemerédi. On the second eigenvalue of random regular graphs. In: *Proceedings of the Twenty-first Annual ACM Symposium on Theory of Computing, STOC '89*. ACM, New York, NY, USA (1989), pp. 587–598

[GT10] F. Götze and A. Tikhomirov. The circular law for random matrices. *Ann. Probab.*, (4)38 (2010), 1444–1491

[HK20] Y. He and A. Knowles. Fluctuations of extreme eigenvalues of sparse Erdős–Rényi graphs. Preprint arXiv:2005.02254 (2020)

[HKM19] Y. He, A. Knowles, and M. Marcozzi. Local law and complete eigenvector delocalization for supercritical Erdős-Rényi graphs. *Ann. Prob.*, (5)47 (2019), 3278–3302

[HL15] J. Huang and B. Landon. Spectral statistics of sparse Erdős-Rényi graph Laplacians. Ann. Inst. H. Poincar'e Probab. Statist., (1)56 (2020), 120–154

[HLY15] J. Huang, B. Landon, and H.-T. Yau. Bulk universality of sparse random matrices. *J. Math. Phys.*, (12)56 (2015), 123301

[HLY17] J. Huang, B. Landon, and H.-T. Yau. Transition from Tracy-Widom to gaussian fluctuations of extremal eigenvalues of sparse Erdős-Rényi graphs. Ann. Probab., (2)48 (2020), 916–962

[Joh01] K. Johansson. Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.* 215(3) (2001), 683–705

[Joh15] T. Johnson. Exchangeable pairs, switchings, and random regular graphs. *Electron. J. Combin.*, (1)22:Paper 1.33 (2015), 28

[Kes59] H. Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, 92 (1959), 336–354

[KY13] A. Knowles and J. Yin. Eigenvector distribution of Wigner matrices. *Probab. Theory Related Fields*, (3-4)155 (2013), 543–582

[KSVW01] M. Krivelevich, B. Sudakov, V.H. Vu, and N.C. Wormald. Random regular graphs of high degree. *Random Structures Algorithms*, (4)18 (2001), 346–363

[LY17] B. Landon and H.-T. Yau. Edge statistics of dyson brownian motion. Preprint arXiv:1712.03881 (2017)

[LS18] J.O. Lee and K. Schnelli. Local law and Tracy-Widom limit for sparse random matrices. *Probab. Theory Related Fields*, (1-2)171 (2018), 543–616

[LLTTY18] A.E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef. Circular law for sparse random regular digraphs. Preprint arXiv:1801.05576 (2018)

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, (3)8 (1988), 261–277

[MSS15] A.W. Marcus, D.A. Spielman, and N. Srivastava. Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. of Math. (2)*, (1)182 (2015), 307–325

[MSS18] A.W. Marcus, D.A. Spielman, and N. Srivastava. Interlacing families IV: Bipartite Ramanujan graphs of all sizes. *SIAM J. Comput.*, (6)47 (2018), 2488–2509

[Mar88] G.A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, (1)24 (1988), 51–60

[MPL14] F.L. Metz, G. Parisi, and L. Leuzzi. Finite-size corrections to the spectrum of regular random graphs: an analytical solution. *Phys. Rev. E*, 90 (2014), 052109

[MN08] S.J. Miller and T. Novikoff. The distribution of the largest nontrivial eigenvalues in families of random regular graphs. *Experiment. Math.*, (2)17 (2008), 231–244

[PZ10] G. Pan and W. Zhou. Circular law, extreme singular values and potential theory. *J. Multivariate Anal.*, (3)101 (2010), 645–656

[Pec09] S. Péché. Universality results for the largest eigenvalues of some sample covariance matrix ensembles. *Probab. Theory Related Fields*, (3-4)143 (2009), 481–516

[Pud15] D. Puder. Expansion of random graphs: new proofs, new results. *Invent. Math.*, (3)201 (2015), 845–908

[Sar04] P. Sarnak. What is an expander? *Notices Amer. Math. Soc.*, (7)51 (2004), 762–763

[Sod10] S. Sodin. The spectral edge of some random band matrices. *Ann. of Math. (2)*, (3)172 (2010), 2223–2251

[Sos99] A. Soshnikov Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.*, (3)207 (1999), 697–733

[TV10] T. Tao and V. Vu. Random matrices: universality of ESDs and the circular law. *Ann. Probab.*, (5)38 (2010), 2023–2065. With an appendix by Manjunath Krishnapur

[TV10] T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.*, (2)298 (2010), 549–572

[TV11] T. Tao and V. Vu. Random matrices: universality of local eigenvalue statistics. *Acta Math.*, (1)206 (2011), 127–204

[TV12] T. Tao and V. Vu. Random matrices: universal properties of eigenvectors. *Random Matrices Theory Appl.*, (1)1 (2012), 1150001

[TY19] K. Tikhomirov and P. Youssef. The spectral gap of dense random regular graphs. *Ann. Probab.*, (1)47 (2019), 362–419

[V08] V. Vu. Random discrete matrices. In *Horizons of Combinatorics. Bolyai Society Mathematical Studies*, Vol. 17. Springer, Berlin, (2008), pp. 257–280

[Vu14] V.H. Vu. Combinatorial problems in random matrix theory. In: *Proceedings of the International Congress of Mathematicians—Seoul 2014*, Vol. IV, Kyung Moon Sa, Seoul (2014), pp. 489–508

Roland Bauerschmidt
University of Cambridge, Cambridge, UK. rb812@cam.ac.uk

Jiaoyang Huang
IAS, Princeton, USA. jiaoyang@ias.edu

Antti Knowles
University of Geneva, Geneva, Switzerland. antti.knowles@unige.ch

Horng-Tzer Yau
Harvard University, Cambridge, USA. htyau@math.harvard.edu

Received: October 22, 2019
Revised: April 8, 2020
Accepted: June 6, 2020