SPECTRAL MEASURE FOR UNIFORM d-REGULAR DIGRAPHS

ARKA ADHIKARI AND AMIR DEMBO

ABSTRACT. Consider the matrix $\mathsf{A}_\mathcal{G}$ chosen uniformly at random from the finite set of all N-dimensional matrices of zero main-diagonal and binary entries, having each row and column of $\mathsf{A}_\mathcal{G}$ sum to d. That is, the adjacency matrix for the uniformly random d-regular simple digraph $\mathcal{G}$. Fixing $d \geq 3$, it has long been conjectured that as $N \to \infty$ the corresponding empirical eigenvalue distributions converge weakly, in probability, to an explicit non-random limit, given by the Brown measure of the free sum of d Haar unitary operators. We reduce this conjecture to bounding the decay in N of the probability that the minimal singular value of the shifted matrix $\mathsf{A}(w) = \mathsf{A}_\mathcal{G} - w\mathsf{I}$ is very small. While the latter remains a challenging task, the required bound is comparable to the recently established control on the singularity of $\mathsf{A}_\mathcal{G}$. The reduction is achieved here by sharp estimates on the behavior at large N, near the real line, of the Green's function (aka resolvent) of the Hermitization of $\mathsf{A}(w)$, which is of independent interest.

1. INTRODUCTION

The method of moments and the Stieltjes transform approach provide rather precise information on asymptotics of the Empirical Spectral Distribution (in short ESD), for many *Hermitian* random matrix models. In contrast, both methods fail for *non-Hermitian* matrix models, and the only available general scheme for finding the limiting spectral distribution in such cases is the one proposed by Girko (in [13]). It is extremely challenging to rigorously justify this scheme, even for the matrix model consisting of i.i.d. entries (of zero mean and finite variance). Indeed, the *circular law* conjecture, for the i.i.d. case, was established in full generality only in 2010, after rather long series of partial results (see [30] and historical references in [8]). Beyond this simple model, only a few results have been rigorously proved in the non-Hermitian regime. Our focus here is on the long standing conjecture about the limit, as $N \to \infty$, of the ESD of the adjacency matrix $\mathsf{A}_\mathcal{G}$ in case $\mathcal{G}$ is a uniformly random d-regular *directed graph* (aka digraph) of N vertices. Specifically, in this context it is conjectured (see [8]), that, fixing $d \geq 3$, such ESD converge weakly, in probability, to the oriented Kesten-McKay law. Namely, the measure μ_d on the complex plane, whose density with respect to Lebesgue measure on $\mathbb{C}$ is

$$(1.1) \qquad h_d(w) := \frac{1}{\pi} \frac{d^2(d-1)}{(d^2 - |w|^2)^2} \mathbb{I}_{\{|w| \leq \sqrt{d}\}}$$

(c.f. [9, 24] for more insights on this conjecture). Related to this conjecture, [10] applies the high trace method to study the spectral gap of such matrices, proving in particular that the limiting ESD support must be the disk of radius $\sqrt{d}$ with no outliers apart from the trivial eigenvalue. This is also one of the consequences of [11] work on the asymptotic of the characteristic polynomial of such $\mathsf{A}_\mathcal{G}$ outside the disk of radius $\sqrt{d}$ (for more on this direction, see [11] and the references therein). Note that as $d \to \infty$, the oriented Kesten-McKay law μ_d, rescaled by $\sqrt{d}$, converges to the circular law. This suggests that the circular law should hold for the limiting ESD of $\frac{1}{\sqrt{d}}\mathsf{A}_\mathcal{G}$ whenever $d \wedge (N-d) \to \infty$ as $N \to \infty$, as indeed shown by [9, 21] (see also [2] for this result in the related permutation model, with at least poly-log growing degree).

From a graph-theoretic point of view, it is most natural to consider the *simple digraph* (SD) version of this problem, whereby we restrict $\mathcal{G}$ to be a simple digraph, or alternatively choose $\mathsf{A}_\mathcal{G}$ uniformly from the ensemble of all matrices of $\{0, 1\}$-valued entries with zero main diagonal and having each of its row and column sum to d. Taking advantage of the contiguity results of [19], one may alternatively aim at proving

this conjecture for the *configuration model* (CM), where $A_{\mathcal{G}}$ is uniformly chosen from among all matrices of non-negative integer entries with all rows and columns sum to d (that is, $\mathcal{G}$ is still uniform, directed and d-regular, but now allowing also for self and multiple edges). The preceding conjecture is best appreciated in the context of yet another version, the so called *permutation model* (PM), where $A_{\mathcal{G}}$ is taken as the sum of d uniformly and independently chosen N-dimensional permutation matrices (that is, with $\mathcal{G}$ a d-regular multi-graph, now of a slightly non-uniform law).

Indeed, permutation matrices form a subgroup of the group of orthogonal matrices, and it was shown in [3] that as $N \to \infty$, the ESD for the sum of d independently chosen, *Haar distributed* N-dimensional, orthogonal matrices, converge weakly, in probability, to the measure μ_d whose density is $h_d(\cdot)$ of (1.1). This result of [3] brings us back to the rationale behind the conjecture (which should hold for all three versions of uniform d-regular digraph we have mentioned). Namely, the observation that μ_d is the *Brown measure* of the *free sum* of $d \geq 2$ *Haar unitary operators* (see [16, Example 5.5]). To understand the plausible route and challenges in affirming this conjecture, recall the definition of Brown measure for a bounded operator (see [16, Page 333]), where $\langle \mathrm{Log}, \mu \rangle := \int_{\mathbb{R}} \log |x| d\mu(x)$ for any probability measure μ on $\mathbb{R}$ (for which such integral is well defined).

Definition 1.1. *Let $(\mathcal{A}, \tau)$ be a non-commutative W^*-probability space, i.e. a von Neumann algebra $\mathcal{A}$ with a normal faithful tracial state τ (see [1, Defn. 5.2.26]). For h a positive element in $\mathcal{A}$, let μ_h denote the unique probability measure on $\mathbb{R}^+$ such that $\tau(h^n) = \int t^n d\mu_h(t)$ for all $n \in \mathbb{Z}^+$. The Brown measure μ_a associated with each bounded $a \in \mathcal{A}$, is the Riesz measure corresponding to the $[-\infty, \infty)$-valued sub-harmonic function $w \mapsto \langle \mathrm{Log}, \mu_{|a-w|} \rangle$ on $\mathbb{C}$. That is, μ_a is the unique Borel probability measure on $\mathbb{C}$ such that*

$$(1.2) \qquad d\mu_a(w) = \frac{1}{2\pi} \Delta_w \langle \mathrm{Log}, \mu_{|a-w|} \rangle \, dw,$$

where Δ_w denotes the two-dimensional Laplacian operator (with respect to $w \in \mathbb{C}$), and the identity (1.2) holds in distribution sense (i.e. when integrated against any test function $\psi \in C_c^\infty(\mathbb{C})$).

Recall that as $N \to \infty$, independent Haar distributed N-dimensional orthogonal matrices converge in $\star$-moments (see [29] for a definition), to the collection $\{u_i\}_{i=1}^d$ of $\star$-*free* Haar unitary operators (see [1, Thm. 5.4.10]). The challenge, even in the context of [3], stems from the fact that the convergence of $\star$-moments (or even the stronger convergence in distribution of traffics, see [23]), does not imply convergence of the corresponding Brown measures[1] (see [29, §2.6]). Moreover, while [29, Thm. 6] shows that if the original matrices are perturbed by adding small Gaussian (of *unknown variance*), then the Brown measures do converge, removing the Gaussian, or merely identifying the variance needed, are often hard tasks (e.g. [15, Prop. 7 & Cor. 8] provide an example of an ensemble where no Gaussian matrix of polynomially vanishing variance can regularize the Brown measures, in this sense). In contrast, for Toeplitz matrices, many spectral features of regularization by adding a matrix of polynomially small norm, are by now fairly well understood (c.f. [28] for Gaussian perturbations, or [5] for non-Gaussian ones).

Thus, instead of relying on the preceding free probability intuition, we shall utilize Girko's approach, which we now summarize. First, one associates to any N-dimensional (non-Hermitian) matrix A_N and every $w \in \mathbb{C}$ the $2N$-dimensional *Hermitian* matrix

$$(1.3) \qquad H^{A_N}(0, w) := \begin{bmatrix} 0 & A_N - wI_N \\ A_N^* - \overline{w}I_N & 0 \end{bmatrix}.$$

Applying Green's formula to the characteristic polynomial of a matrix A_N (whose ESD we denote hereafter by L_{A_N}), results with

$$\int_{\mathbb{C}} \psi(w) dL_{A_N}(w) = \frac{1}{2\pi N} \int_{\mathbb{C}} \Delta\psi(w) \log |\det(wI_N - A_N)| dw, \qquad \forall \psi \in C_c^2(\mathbb{C}).$$

---

[1]The Brown measure of a matrix is its ESD (see [29, Prop. 1])

The eigenvalues of $\mathsf{H}^{\mathsf{A}_N}(0,w)$ are merely ± 1 times the singular values of $w\mathsf{I}_N - \mathsf{A}_N$. Hence, denoting the ESD of $\mathsf{H}^{\mathsf{A}_N}(0,w)$ by ν_N^w, we have that

$$\frac{1}{N}\log|\det(w\mathsf{I}_N - \mathsf{A}_N)| = \frac{1}{2N}\log|\det\mathsf{H}^{\mathsf{A}_N}(0,w)| = \langle\mathrm{Log}, \nu_N^w\rangle\,,$$

out of which we deduce Girko's formula,

$$(1.4) \qquad \int_{\mathbb{C}}\psi(w)dL_{\mathsf{A}_N}(w) = \frac{1}{2\pi}\int_{\mathbb{C}}\Delta\psi(w)\langle\mathrm{Log}, \nu_N^w\rangle dw\,, \qquad \forall\psi\in C_c^2(\mathbb{C})\,.$$

The preceding identity suggests the following recipe for proving convergence of L_{A_N} per given family of non-Hermitian random matrices $\{\mathsf{A}_N\}$ (to which we referred already as Girko's method).

Step 1: Show that for (Lebesgue almost) every $w\in\mathbb{C}$, as $N\to\infty$ the measures ν_N^w converge weakly, in probability, to some measure ν^w.

Step 2: Justify that $\langle\mathrm{Log}, \nu_N^w\rangle \to \langle\mathrm{Log}, \nu^w\rangle$ in probability, for almost every $w\in\mathbb{C}$ (which is the main technical challenge of this approach).

Step 3: A uniform integrability argument allows one to convert the w-a.e. convergence of $\langle\mathrm{Log}, \nu_N^w\rangle$ to the corresponding convergence for a suitable collection $\mathcal{S}\subseteq C_c^2(\mathbb{C})$ of (smooth) test functions. It then follows from (1.4) that for each fixed $\psi\in\mathcal{S}$, we have the convergence in probability

$$\int_{\mathbb{C}}\psi(w)dL_{\mathsf{A}_N}(w) \to \frac{1}{2\pi}\int_{\mathbb{C}}\Delta\psi(w)\langle\mathrm{Log}, \nu^w\rangle dw$$
$$(1.5) \qquad\qquad = \int_{\mathbb{C}}\psi(w)h(w)dw\,, \quad\text{for}\quad h(w) := \frac{1}{2\pi}\Delta\langle\mathrm{Log}, \nu^w\rangle\,,$$

provided $w\mapsto\langle\mathrm{Log}, \nu^w\rangle$ is smooth enough to justify the integration by parts. For $\mathcal{S}$ large enough, this implies the convergence in probability of the ESD-s L_{A_N} to a limit whose density is $h(w)$.

For example, [3] follows this approach, inductively over $d\geq 2$, and for the core induction step they adapt the arguments of [14] (who for Haar distributed orthogonal O_N, independent of the uniformly bounded, non-negative definite, diagonal D_N, show that the weak convergence of L_{D_N}, in probability, implies the same for $L_{\mathsf{O}_N\mathsf{D}_N}$, whose limit is the Brown measure of a certain operator). Utilizing this, it suffices in our case to only consider **Step 2**. Indeed, for any adjacency matrix A_N of a d-regular digraph we have that $d^{-2}\mathsf{A}_N^*\mathsf{A}_N$ is a (Markov) probability transition matrix, hence of spectral norm one. Any such A_N must then have singular values bounded by d and a Hilbert-Schmidt norm $\|\mathsf{A}_N\|_2 \leq d\sqrt{N}$. The same applies for any sum of d unitary matrices, and in particular to the sum B_N of d independent Haar distributed unitary matrices. Recall, from [3, Prop. 1.4], that L_{B_N} converges weakly, in probability, to the measure μ_d (of density $h_d(w)$ as in (1.1)), and from [3, Lemma 3.2], that for almost every $w\in\mathbb{C}$, as $N\to\infty$,

$$\frac{1}{N}\log|\det(w\mathsf{I}_N - \mathsf{B}_N)| \to \langle\mathrm{Log}, \nu^w\rangle\,.$$

It then follows from applying the *replacement principle* of [30, Thm. 2.1] for ensembles (A_N) and (B_N), that **Step 2** already yields the conjectured convergence for any such L_{A_N}.

The following theorem, which is our main result, addresses the latter convergence, while also providing the corresponding local limit law.

Theorem 1.2. *Let $\mathbb{P}_{N,d}$ denote the uniform distribution over all d-regular, simple digraphs on N vertices, with $\mathsf{A}_N = A_{\mathcal{G}}$ the adjacency matrix of such a digraph, and $s_1(w)$ the minimal singular value of $\mathsf{A}_N - w\mathsf{I}_N$. Suppose for $\mathfrak{a} = 0$, any fixed $\kappa > 0$ and a.e $w\in\mathbb{C}$,*

$$(1.6) \qquad \lim_{N\to\infty}N^{\mathfrak{a}}\,\mathbb{P}_{N,d}(-\log s_1(w) \geq N^{\kappa}) = 0\,.$$

(a). Girko's method **Step 2** *holds for our ν_N^w (i.e. $\langle\mathrm{Log}, \nu_N^w\rangle \to \langle\mathrm{Log}, \nu^w\rangle$, in probability), and the ESD $\{L_{\mathsf{A}_N}\}$ converge weakly, in probability to μ_d, whose density is $h_d(w)$ of (1.1).*

(b). Further, if (1.6) holds with $\mathfrak{a}(\kappa) > 0$, then for some $\mathfrak{a}_o > 0$ any $w_o \in \mathbb{C}$ and $\psi \in C_c^2(\mathbb{C})$, setting $\psi_r(\cdot) := r^2 \psi(r \cdot)$, we have that

$$(1.7) \qquad \lim_{N \to \infty} \mathbb{P}_{N,d} \Big(\big| \langle \psi_{N^a}(\cdot - w_o), L_{\mathfrak{A}_N} - \mu_d \rangle \big| > \eta \Big) = 0 , \qquad \forall \eta > 0, a \in (0, \mathfrak{a}_o) .$$

There is a long history regarding bounds on the smallest singular value of random matrix ensembles which are of the type of (1.6) (see, e.g. [30, 26, 2, 20, 22, 27]). Of particular note are the paper [17], which shows that $\mathbb{P}_{N,d}(s_1(0) = 0) \to 0$ as $N \to \infty$ (c.f. [25] for better decay rate), and [27], which establishes (1.6) at $\kappa = 1$ for the adjacency matrices of Erdos-Renyi(d/N) digraphs. Unfortunately, [17, 25] rely on considering digraphs over finite fields, which does not seem applicable for an estimate such as (1.6), whereas independence of the entries of A_N seems to be key to the success of the approach taken in [27].

As we shortly show, Theorem 1.2 is an immediate consequence of the following local law for the Green's function $\mathsf{G}_N(z, w) := [\mathsf{H}^{\mathsf{A}_N}(0, w) - z \mathsf{I}_{2N}]^{-1}$ of our Hermitized matrices.

Theorem 1.3. *Let $m_\star(z, w) := \int_{\mathbb{R}} (\lambda - z)^{-1} d\nu^w(\lambda)$ (which is well defined for $\Im(z) > 0$). For any $0 < \mathfrak{c} < 1$ and $d \geq 3$, there exist $\varepsilon = \varepsilon(\mathfrak{c}, d) > 0$ and $\delta = \delta(\mathfrak{c}, d) > 0$ so that $\mathbb{P}_{N,d}(\Omega'_N) \geq 1 - O(N^{-\mathfrak{c}})$ for some sets Ω'_N of d-regular, simple digraphs and for any compact $\mathbb{D} \subset \mathbb{C}$, as $N \to \infty$,*

$$(1.8) \qquad \sup_{\mathcal{G} \in \Omega'_N} \sup_{w \in \mathbb{D}, \Im(z) \geq N^{-\varepsilon}} \Big| \frac{1}{2N} \mathrm{trace}(\mathsf{G}_N(z, w)) - m_\star(z, w) \Big| \leq N^{-\delta} .$$

Remark 1.4. The values of $\delta(\mathfrak{c}, d)$ and $\varepsilon(\mathfrak{c}, d)$ in our proof are sub-optimal. While plausible to expect our method to yield $\varepsilon(\mathfrak{c}, d) \to 1$ as $d \to \infty$ and $c \to 0$, this would entail substantially more effort. We further note that near the edge of the support of μ_d (namely, at $|w| = \sqrt{d}$), the singularity of the self-consistency equations causes weaker estimates on the Green's function, but having small ε and δ allows us to get the uniform bound in (1.8), avoiding effects due to such singularity near the edge.

Remark 1.5. We have stated Theorems 1.2 and 1.3 in terms of uniformly random, *simple d-regular digraphs*, namely, taken from the SD model. However, our work yields precisely the same results when $\mathcal{G}$ is taken from the configuration model.

Remark 1.6. Combining Theorem 1.3 with [4, Theorem 2.1] (at $\alpha = 0$), one has, for the standard complex Ginibre matrix G_N, that under $\mathbb{P}_{N,d}$ the ESD of $\mathsf{A}_N + N^{-\beta - \frac{1}{2}} \mathsf{G}_N$ converges weakly, in probability, to the oriented Kesten-McKay law μ_d. Utilizing [5, Theorem 1.8], the same applies for $\mathsf{A}_N + N^{-\beta} \mathsf{O}_N$, without requiring (1.6) to hold.

Many recent papers deal with establishing local laws for graph models. For example, [6, 7, 18] derive local laws via switching estimates that generate useful self-consistent equations. The most relevant here are [6, 18], which introduce the strategy of preserving local neighborhoods and performing switches along the boundary of the local neighborhood. In the graphs considered in those papers the local neighborhoods of most vertices match those of the root of an infinite d-regular tree, so one applies perturbation theory and averaging estimates to compare the Green's functions of the graph with the one at the root of the infinite d-regular tree. The key for doing so are estimates on the Green's function on the tree, which thanks to its homogeneous structure depend only on the distance between the two vertices. This homogeneity of the tree Green's function helps one establish the contractive property of the linearized self-consistency equations for the perturbation, which is essential to the success of this approach.

While we follow the same philosophy in our proof of Theorem 1.3, a major new challenge arises when considering digraphs, whereby the limit object with respect to which we now perturb is the infinite d-regular directed tree (see Definition 1.10). Due to having separate in-edges and out-edges in the directed tree, homogeneity is lost, with the Green's function estimates now depending on the precise path taken between the two vertices on the tree. This adds complexity to the perturbation estimates we need (see Section 3), with much extra care now required for comparing the Green's functions of the original and the switched graphs (specifically, for controlling the effect that applying Schur complement formula for the removal of large neighborhoods has on the Green's functions of the switched graphs). Lastly and most importantly, as further detailed at the start of Section 2, in the digraph case the limiting (non-perturbative) Green's function

evolution along some edges is expansive (instead of contractive), and a very precise averaging is crucial to get the stability of the self-consistency equations.

Deferring for awhile the construction and the proof of all these Green's function estimates, we proceed with a standard, short derivation of Theorem 1.2 out of Theorem 1.3. The remainder of this paper is devoted to the proof of Theorem 1.3, and other results which we get along the way (e.g. about the asymptotic behavior of the off-diagonal entries of G_N).

1.1. **Proof of Theorem 1.2.** Fixing $d \geq 3$, $0 < \mathfrak{c} < 1$ and $\kappa < \varepsilon(\mathfrak{c}, d)$ we restrict ourselves hereafter WLOG to digraphs in Ω'_N from Theorem 1.3, where the estimate (1.8) holds and denote by $\Omega_N(w)$ those digraphs for which $s_1(w)$, the minimal singular value of $\mathsf{A}_{\mathcal{G}} - w\mathsf{I}_N$ exceeds $\delta_N := e^{-N^\kappa}$.
(a). Fixing $w \in \mathbb{C}$ where (1.6) holds (namely, $\mathbb{P}_{N,d}(\Omega_N(w)) \to 1$), recall that for any $z \in \mathbb{C}_+$ and N,

$$(1.9) \qquad \frac{1}{2N}\mathrm{trace}(\mathsf{G}_N(z, w)) = \int_{\mathbb{R}} (\lambda - z)^{-1} d\nu_N^w(\lambda) \,.$$

In particular, (1.8) implies the convergence as $N \to \infty$ of the Stieltjes transform of ν_N^w to that of ν^w, at any $z \in \mathbb{C}_+$, uniformly over Ω'_N. This of course implies that $|\langle f, \nu_N^w \rangle - \langle f, \nu^w \rangle| \to 0$ for any $f \in C_b(\mathbb{C})$ (which is **Step 1** of Girko's method), also uniformly over digraphs in Ω'_N. Further, with $\langle \mathrm{Log}, \mu \rangle_a^b := \int \mathbf{1}_{|x| \in [a,b]} \log|x| d\mu(x)$ for $a < b$, and recalling that the spectral norm of $\mathsf{A}_{\mathcal{G}}(w)$ is uniformly bounded (by $d + |w|$), we have for $\eta > 0$ fixed, uniformly over Ω'_N, that as $N \to \infty$,

$$|\langle \mathrm{Log}, \nu_N^w \rangle_\eta^\infty - \langle \mathrm{Log}, \nu^w \rangle_\eta^\infty| \to 0 \,.$$

By definition, on $\Omega_N(w)$ we have that $\nu_N^w(-(\delta_N, \delta_N)) = 0$. Further, from [3, (4.18)] we have for a.e. w,

$$(1.10) \qquad \left| \langle \mathrm{Log}, \nu^w \rangle_0^\eta \right| \leq C\eta |\log \eta| \,, \qquad \forall 0 \leq \eta \leq 1/2 \,,$$

Thus, it remains only to show that $\langle \mathrm{Log}, \nu_N^w \rangle_{\delta_N}^\eta \to 0$ uniformly on Ω_N, when $N \to \infty$ followed by $\eta \to 0$. To this end, note that (1.8) also implies that for some $C < \infty$, any N, all digraphs in Ω'_N and $y > 0$,

$$\nu_N^w((-y, y)) \leq C(y \vee N^{-\varepsilon})$$

(see [3, (4.15)] or [14, Lemma 15]). Combining this bound with integration by parts we find that

$$(1.11) \qquad \langle \mathrm{Log}, \nu_N^w \rangle_{\delta_N}^\eta \leq C \int_{\delta_N}^\eta \frac{1}{y}(y \vee N^{-\varepsilon}) dy \leq C(N^{-\varepsilon}|\log \delta_N| + \eta) = C(N^{\kappa - \varepsilon} + \eta) \,,$$

which goes to zero when $N \to \infty$ followed by $\eta \to 0$, thereby establishing **Step 2** of Girko's method.
(b). Next, fixing $\eta > 0$ and $a < \mathfrak{a}_o$, since the Lipschitz norm of $\psi_r(\cdot)$ is at most $r^3\|\nabla\psi\|_\infty$, we can and shall replace $\langle \psi_{N^a}(\cdot - w_o), L_{\mathsf{A}_N} \rangle$ in (1.7), by $\langle \widetilde{\psi}_{N^a}(\cdot - w_o), L_{\mathsf{A}_N} \rangle$, where $\widetilde{\psi}_r(\cdot) := \mathbb{E}_{U_N}[\psi_r(\cdot - U_N)]$ for U_N uniformly distributed on $\mathbb{B}(0, N^{-3\mathfrak{a}_o})$ independently of A_N. Hence, we let

$$\Gamma_N(w) := \mathbb{E}_{U_N}\left[\langle \mathrm{Log}, \nu_N^{w+U_N} \rangle \right] - \langle \mathrm{Log}, \nu^w \rangle \,,$$

and combine Girko's formula (1.4) for $\widetilde{\psi}_{N^a}(\cdot)$ and $\psi_{N^a}(\cdot)$, with Fubini's theorem, to get (1.7) out of

$$\lim_{N \to \infty} \mathbb{P}_{N,d}\left(\left| \int_{\mathbb{C}} \Delta\psi_{N^a}(w)\Gamma_N(w_o + w) dw \right| > \eta \right) = 0 \,.$$

We will now obtain estimates on the differences of the empirical densities of ν_N^w and ν^w via the Helffer-Sjostrand formula. First, we remark that the Green's function $m_{\mathcal{T}}^d(z, w)$ is bounded by a constant via the computations of Lemma A.1. Thus, the estimate (1.8) would also imply that $\frac{1}{2N}\mathrm{trace}(G_N(z, w))$ is bounded by a constant independent of N, z and w as long as we consider z with $\Im(z) \geq N^{-\varepsilon}$. Now, since the functions $\Im(z)\Im(m_{\mathcal{T}}^d(z, w))$ and $\Im(z)\frac{1}{2N}\Im(\mathrm{trace}(G_N(z, w)))$ are both monotone increasing in $\Im(z)$. Thus, we have that,

$$(1.12) \qquad \Im(m_{\mathcal{T}}^d(z, w)) \leq \frac{CN^{-\varepsilon}}{\Im(z)}, \frac{1}{2N}\Im(\mathrm{trace}(G_N(z, w))) \leq \frac{CN^{-\varepsilon}}{\Im(z)}$$

for any z with $\Im(z) \leq N^{-\varepsilon}$.

In addition, we have that,

$$(1.13) \qquad \left| m_{\mathcal{T}}^d(z, w) - \frac{1}{2N} \Im(\mathrm{trace}(G_N(z, w))) \right| \leq \frac{N^{-\delta}}{\Im(z)}, \forall z : 1 \geq \Im(z) \geq N^{-\varepsilon}.$$

If we define $\rho = \nu_N^w - \nu^w$ and let $f_{\hat{\eta}}$ be a positive twice differentiable function that is 1 on the region $(-E, E)$, 0 on $(-E - \hat{\eta}, E + \hat{\eta})$ and has first derivative bounded by $\frac{1}{\hat{\eta}}$ and second derivative bounded by $\frac{1}{\hat{\eta}^2}$, we can apply the Hellfer-Sjostrand formula as in [12, Lemma 11.2] to derive that,

$$
\begin{aligned}
(1.14) \quad \nu_N^w(-E, E) - \nu^w(-E, E) &= \int_{-E-\hat{\eta}}^{E+\hat{\eta}} f_{\hat{\eta}}(x) \mathrm{d}\rho(x) - \int_E^{E+\hat{\eta}} f_{\hat{\eta}}(x) \mathrm{d}\nu_N^w(x) \\
&\quad - \int_{-E-\hat{\eta}}^{-E} f_{\hat{\eta}}(x) \mathrm{d}\nu_N^w(x) + \int_E^{E+\hat{\eta}} f_{\hat{\eta}}(x) \mathrm{d}\nu^w(x) + \int_E^{E+\hat{\eta}} f_{\hat{\eta}}(x) \mathrm{d}\nu^w(x) \\
&\leq CN^{-\varepsilon} \log N + N^{-\delta} \frac{N^{-2\varepsilon}}{\hat{\eta}^2} + C\hat{\eta}
\end{aligned}
$$

We used the fact that $-\int_E^{E+\hat{\eta}} f_{\hat{\eta}}(x)$ is negative and that the density of ν^w is bounded (again via Lemma A.1) to control $\int_E^{E+\hat{\eta}} f_{\hat{\eta}}(x) \mathrm{d}\nu^w(x)$ for the last inequality. We can derive a lower bound for $\nu_N^w(-E, E) - \nu^w(-E, E)$ in a similar way.

If we set E to be any value $\gg N^{-\varepsilon}$, we can choose $\hat{\eta} = N^{-\varepsilon} \log N$ to assert that

$$(1.15) \qquad |\nu_N^w(-y, y) - \nu^w(-y, y)| \leq C[N^{-\delta} + N^{-\varepsilon} \log N], \ \forall \ y \gg N^{-\varepsilon}.$$

Using this estimate, we find that for $\chi = N^{-\varepsilon+\tau}$ for some small $\tau > 0$ and $\eta = d^2 + |w|$, we have that

$$(1.16) \qquad |\langle \mathrm{Log}, \nu_N^w \rangle_\chi^\eta - \langle \mathrm{Log}, \nu^w \rangle_\chi^\eta| \leq C[N^{-\delta} + N^{-\varepsilon} \log N] \log N.$$

If we also assume that we are on the event $\Omega_N(w)$, then we can also apply the estimates (1.11) and (1.10) to bound

$$(1.17) \quad |\langle \mathrm{Log}, \nu_N^w \rangle_0^\chi - \langle \mathrm{Log}, \nu^w \rangle_0^\chi| \leq |\langle \mathrm{Log}, \nu_N^w \rangle_0^\chi| + |\langle \mathrm{Log}, \nu^w \rangle_0^\chi|| \leq C(N^{\kappa-\varepsilon} + N^{-\varepsilon+\tau}) + CN^{-\varepsilon+\tau} \log N.$$

Now, since both $\nu_N^w(-y, y) = \nu^w(-y, y) = 1$ for $|y| \geq d^2 + |w|$, we have that

$$|\langle \mathrm{Log}, \nu_N^w \rangle_\eta^\infty - \langle \mathrm{Log}, \nu^w \rangle_\eta^\infty| = 0$$

Combining all these estimates, we see that on the event $\Omega_N(w)$, we have that,

$$(1.18) \qquad |\langle \mathrm{Log}, \nu^w \rangle - \langle \mathrm{Log}, \nu_N^w \rangle| \leq C[N^{\kappa-\varepsilon} \log N + N^{-\varepsilon+\tau} \log N + N^{-\delta} \log N + N^{-\varepsilon} (\log N)^2].$$

In what proceeds, we choose κ very small, τ very small, and $2a < \min(\varepsilon - \max[\kappa, \tau], \delta)$. Then, we have that on the event $\Omega_N(w)$ that,

$$(1.19) \qquad |\langle \mathrm{Log}, \nu^w \rangle - \langle \mathrm{Log}, \nu_N^w \rangle| \leq CN^{-2a} (\log N)^{-2}.$$

Recall the event $\Omega_N(w + U_N)$, which is the event that the smallest singular value of the matrix $A_N + U_N$ is larger than $\delta_N = \exp[-N^\kappa]$, we notice that the above estimates can be written as follows.

$$
\begin{aligned}
\left| \langle \mathrm{Log}, \nu_N^{w+U_N} \rangle - \langle \mathrm{Log}, \nu^{w+U_N} \rangle \right| &\leq C(\log N)^{-2} N^{-2a} \mathbb{1}[\Omega_N(w + U_N)] \\
&\quad + \mathbb{1}[\Omega_N(w + U_N)^c] \left[\frac{1}{N} \sum_{i=1}^N |\log |\lambda_i - w - U_N|| + |\langle \mathrm{Log}, \nu^{w+U_N} \rangle| \right].
\end{aligned}
$$

By putting this inequality back into Girko's formula and taking expectation, we notice that

$$\mathbb{E}\left[\int_{\mathbb{C}} \triangle\psi_{N^a}(w)\Gamma_N(w_o + w)dw\right] \leq N^{2a}\int_{\mathbb{C}} |\triangle\psi(\xi)|C(\log N)^{-2}N^{-2a}\mathbb{E}[\mathbb{1}[\Omega_N(w_0 + N^{-a}\xi + U_N)]]d\xi$$

$$+ N^{2a}\int_{\mathbb{C}} |\triangle\psi(\xi)|\mathbb{E}\left[\mathbb{1}[\Omega_N(w_0 + N^{-a}\xi + U_N)^c]\left[\frac{1}{N}\sum_{i=1}^{N}|\log|\lambda_i - w_0 + N^{-a}\xi - U_N||\right]\right]d\xi$$

$$(1.20) \quad + N^{2a}\int_{\mathbb{C}} |\triangle\psi(\xi)|\mathbb{E}\left[\mathbb{1}[\Omega_N(w_0 + N^{-a}\xi + U_N)^c]\right]|\langle\mathrm{Log}, \nu^{w_0 + N^{-a}\xi + U_N}\rangle|d\xi.$$

The first line of (1.20) is clearly $o(1)$. Furthermore, since pointwise we know that
$\mathbb{E}[\mathbb{1}[\Omega_N(w_0 + N^{-a}\xi + U_N)^c]] \leq N^{-\mathfrak{c}}$, the third line is also $o(1)$ provided $2a < \mathfrak{c}$, as otherwise all other quantities in the integral in the third line are constant in N.

To deal with the second line, we first notice that,

$$\mathbb{E}\left[\mathbb{1}[\Omega_N(w_0 + N^{-a}\xi + U_N)^c]\left[\frac{1}{N}\sum_{i=1}^{N}|\log|\lambda_i - w_0 + N^{-a}\xi - U_N||\right]\right]$$

$$\leq \mathbb{E}^{1/2}[\mathbb{1}[\Omega_N(w_0 + N^{-a}\xi + U_N)^c]]\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}^{1/2}\left[\log^2|\lambda_i - w_0 + N^{-a}\xi - U_N|\right]$$

Notice that, regardless of the value of λ_i, if we take the expectation over the random variable U_N, the value of $\mathbb{E}^{1/2}\left[\log^2|\lambda_i - w_0 + N^{-a}\xi - U_N|\right]$ will be bounded by $O(\log N)$. Now, $\mathbb{E}^{1/2}[\mathbb{1}[\Omega_N(w_0 + N^{-a}\xi + U_N)^c]] < N^{-\mathfrak{c}/2}$. Thus, as long as we set $a < \mathfrak{c}/2$, the integral on the second line of (1.20) is also $o(1)$. We find that $\mathbb{E}\left[\int_{\mathbb{C}} \triangle\psi_{N^a}(w)\Gamma_N(w_o + w)dw\right]$ is $o(1)$ and must converge to 0 in probability by Markov's inequality.

1.2. **The model and Green's function.** We start with a more formal definition of our model.

Definition 1.7 (Digraph). A digraph $\mathcal{G}$ has vertices V and directed edges E, where each oriented edges $e = (v, w) \in E$ has an initial vertex v and a terminal vertex $w \neq v$ (in particular, the (v, w) and (w, v) are distinct edges). In the adjacency matrix $\mathsf{A}_{\mathcal{G}}$ of a digraph $\mathcal{G}$, the entry $a_{i,j}$ indicates the number of edges that are oriented from the vertex i to the vertex j (and such adjacency matrix is, in general, not Hermitian), where a simple digraph has only $\{0, 1\}$-valued entries.

Definition 1.8 (Uniformly random d-regular digraph). A d-regular digraph is a graph such that for each vertex $v \in V$, exactly d edges of the graph have v as their initial vertex and exactly d edges have v as their terminal vertex. We let $\mathcal{G}_{N,d}$ denote the (finite) set of all d-regular, simple, digraphs on N vertices with $\mathbb{P}_{N,d}$ the corresponding uniform distribution on $\mathcal{G}_{N,d}$ (namely, selecting each specific instance from $\mathcal{G}_{N,d}$ with probability $\frac{1}{|\mathcal{G}_{N,d}|}$).

Hereafter we often use lighter notations, with the dimension N and the identity matrix I_N implicit, as are sometimes also the arguments w and z. Choosing $\mathcal{G}$ according to $\mathbb{P}_{N,d}$, WLOG we also replace $\mathsf{A}_{\mathcal{G}}$ hereafter by the normalized matrix $A := \frac{1}{\sqrt{d-1}}\mathsf{A}_{\mathcal{G}}$ and for $w \in \mathbb{C}$ and $z \in \mathbb{C}_+$, let

$$(1.21) \qquad G(z, w) := H(z, w)^{-1}, \qquad H(z, w) := \begin{bmatrix} -z & A - w \\ A^* - \overline{w} & -z \end{bmatrix}.$$

By definition, $G(z, w) = \sqrt{d-1}\mathsf{G}_N(\sqrt{d-1}z, \sqrt{d-1}w)$ and towards Theorem 1.3 our goal is to control for large N, the normalized trace of $G(z, w)$, where of most interest is the case of $\Im(z) \downarrow 0$.

To this end, recall the following expansion of the Green's function, at $|z|$ large enough,

$$(1.22) \qquad G(z, w) = -z^{-1}(I - z^{-1}H(0, w))^{-1} = -z^{-1}\sum_{k=0}^{\infty} z^{-k}H(0, w)^k.$$

Fixing w and using $H := H(0, w)$, we further have for any $a, b \leq N$ and $k \geq 0$,

$$[H^{2k}]_{a,b} = \sum_{v_1, v_2, \ldots, v_{2k-1}} \prod_{j=0}^{k-1} H_{v_{2j}, v_{2j+1}+N} H_{v_{2j+1}+N, v_{2j+2}},$$

where the sum is over $v_i \leq N$, with $v_0 = a$ and $v_{2k} = b$. Similarly,

$$[H^{2k}]_{a+N, b+N} = \sum_{v_1, v_2, \ldots, v_{2k-1}} \prod_{j=0}^{k-1} H_{v_{2j}+N, v_{2j+1}} H_{v_{2j+1}, v_{2j+2}+N},$$

again summing over $v_i \leq N$, with $v_0 = a$ and $v_{2k} = b$. Moreover,

$$[H^{2k+1}]_{a, b+N} = \sum_{v_1, \ldots, v_{2k}} \Big[\prod_{j=0}^{k-1} H_{v_{2j}, v_{2j+1}+N} H_{v_{2j+1}+N, v_{2j+2}} \Big] H_{v_{2k}, b+N}$$

summing over $v_i \leq N$, with $v_0 = a$, whereas $[H^{2k+1}]_{a+N, b}$ is the complex conjugate of $[H^{2k+1}]_{a, b+N}$.

Now, in case $v_{2j} \neq v_{2j+1}$, an entry $H_{v_{2j}, v_{2j+1}+N}$ represents a forward edge (agreeing with the orientation) from vertex v_{2j} to vertex v_{2j+1} while $H_{v_{2j+1}+N, v_{2j+2}} = H_{v_{2j+2}, v_{2j+1}+N}$ represents a traversal through a backward edge (opposing the orientation) from the vertex v_{2j+1} to v_{2j+2}. For this reason, we include in a local neighborhood of each vertex $v \in \mathcal{G}$ also all those vertices that are reachable through edges traversed in the reverse direction. When we have in the expansion for $[H^{2k}]_{a,b}$ some $v_{2j} = v_{2j+1}$, it follows that v_{2j-1} travels to v_{2j} via a backward edge, stays at v_{2j} with weight factor $-w$ (as there are no self-edges in $\mathcal{G}$), and proceeds from v_{2j+1} to v_{2j+2} via a backward edge (once again). Similarly, in case we have in such expansion some $v_{2j-1} = v_{2j}$, we enter v_{2j-1} via a forward edge, stay with weight factor $-\overline{w}$ and leave v_{2j} to v_{2j+1} again via a forward edge. Thus, the inclusion of such factors $-w$ or $-\overline{w}$ breaks the otherwise regular pattern of alternating between traveling on forward edges and on backward edges. More generally, we may travel a few forward edges in succession or a few backward edges in succession (each consecutive pair of forward edges then contributes the factor $-\overline{w}$ while each consecutive pair of backward edges contributes $-w$). This solidifies the assertion that a local neighborhood should consider traversal via forward and backward edges *in any type of succession*, measuring the neighborhood's size by the distance in the unoriented graph.

1.3. **The d-regular directed tree and its Green's function.** Our proof relies on relating entries G_{ij} of the Green's function for a d-regular digraph, to the Green's function of the local neighborhood of the graph. Typically, such local neighborhood should look like some subset of the infinite directed tree which we describe next.

Definition 1.9 (Construction of the d-regular K-neighborhoods in directed tree). First, fix a vertex $\hat{r}$ to be the root. From this root $\hat{r}$, create a $2d$ regular tree of K levels. The root will have $2d$ children while any other vertex that is not a root or a child at the K level will have $2d - 1$ children. At this point, we can start orienting the edges. For convenience, assume that we have drawn this tree on the plane so that the root is on the very top and any vertex at distance ℓ from the root will be below its parent at the $(\ell - 1)$-th level. With this planar drawing, the notion of j leftmost children of a vertex would be a meaningful notion. With these notions, we can now describe the orientation of our edges.

Our process proceeds level by level; first, we orient the edges that are adjacent to the root in the unoriented graph. Afterwards, we proceed to orient the edges that are adjacent to a vertex of distance 1 from the root in the unoriented graph and so on. At the root $\hat{r}$, we will orient the leftmost d edges away from the root and the rightmost d edges are oriented toward the root. Now assume that we have oriented the edges that are adjacent to a vertex of distances 1 to $\ell - 1$ from the root in the unoriented graph. We now proceed to orient the edges that are adjacent to a vertex at the ℓ-th level in the unoriented graph that have not been assigned an orientation as of yet. Observe for every vertex v at the ℓ-th level, there are still $2d - 1$ children edges that have not been assigned an orientation and these edges connect the vertex v to a vertex below it.

Depending on the orientation of the parent edge (i.e., the edge already assigned an orientation), there will be two ways to assign the orientation to the remaining children edges. If the parent is oriented away from the vertex in question, then orient the $d - 1$ leftmost children edges of the vertex away from the vertex and

the remaining d edges towards the vertex. If the parent edge is oriented towards the vertex, then we orient the d leftmost children away from the vertex and the remaining $d-1$ edges towards the vertex.

Definition 1.10 (d-regular infinite directed trees). We have three natural notions of d-regular directed trees (while only one is truly d-regular, we abuse the notation to make our descriptions more concise):

(a) The d-regular infinite directed tree of Definition 1.9 is denoted by $\mathcal{T}$. Ignoring orientation, it is a $2d$-regular infinite tree, and every vertex has d outedges and d inedges.
(b) The tree $\mathcal{T}_1$ is such that its unoriented version is a $(2d-1)$-ary tree; namely, the root has $2d-1$ edges while all other vertices have $2d$ edges in the unoriented version. When oriented, $\mathcal{T}_1$ has the property that the root has d outedges and $d-1$ inedges.
(c) The tree $\mathcal{T}_2$ is also such that its unoriented version is a $(2d-1)$-ary tree. The difference from $\mathcal{T}_1$ is that under the orientation, here the root vertex has $d-1$ outedges and d inedges.

1.3.1. *Self-consistent equation for the d-regular directed tree.* Our primary quantity of interest is the Green's function $G_\mathcal{T}$. However, in order to utilize the Schur complement formula to analyze this function, one must also analyze the Green's functions of $\mathcal{T}_1$ and $\mathcal{T}_2$. Our basic operation in the application of the Schur complement formula is to remove the root vertex and induct on the remaining graph. Indeed, removing the root of $\mathcal{T}$ yields d disjoint copies of $\mathcal{T}_1$ and d disjoint copies of $\mathcal{T}_2$. Similarly, removing the root of $\mathcal{T}_1$ yields d copies of $\mathcal{T}_1$ and $d-1$ copies of $\mathcal{T}_2$, with the reverse holding for $\mathcal{T}_2$. Utilizing these relations allows us to inductively compute entries of the Green's functions $G_\mathcal{T}$, $G_{\mathcal{T}_1}$ and $G_{\mathcal{T}_2}$, but there is one complication we must consider. In our convention each vertex $v \in \mathcal{T}$ corresponds to 2 rows in the corresponding Hermitized matrix H, which we shall call v and $v + \aleph$ (just as in an N vertex digraph, such vertex v corresponds to both rows v and $v + N$). Thus, when applying the Schur complement formula we must remove at the same time two vertices: v and $v + \aleph$. Specifically, our starting point in deriving Green's function identities is the block decomposition of a matrix and its inverse,

$$(1.23) \qquad \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

resulting with by elementary algebra with the Schur complement formulas

$$X = (A - BD^{-1}C)^{-1}, \qquad Z = -D^{-1}CX, \qquad W = D^{-1} + D^{-1}CXBD^{-1}.$$

Suppressing the common arguments z, w and setting $G^{(T)} := (H|_{T^c \times T^c})^{-1}$ for the Green's function with the rows and columns of T removed, these equations amount to,

$$(1.24) \qquad G|_{T \times T} = (H|_{T \times T} - H|_{T \times T^c} G^{(T)} H|_{T^c \times T})^{-1},$$

$$(1.25) \qquad G|_{T^c \times T} = -G^{(T)} H|_{T^c \times T} G|_{T \times T},$$

$$(1.26) \qquad G|_{T^c \times T^c} = G^{(T)} + G^{(T)} H|_{T^c \times T} G|_{T \times T} H|_{T \times T^c} G^{(T)}.$$

Hereafter, similarly to (1.21), we first scale the adjacency matrix of $\mathcal{T}$ by the factor $\frac{1}{\sqrt{d-1}}$, so our Green's function is given by $G_\mathcal{T}(z, w) := \sqrt{d-1}\mathsf{G}_\mathcal{T}(\sqrt{d-1}z, \sqrt{d-1}w)$ in terms of the original, non-scaled function $\mathsf{G}_\mathcal{T}$. Our analysis of $G_\mathcal{T}$ relies on the removal of both indices $\hat{r}$ and $\hat{r} + \aleph$ associated with the root vertex $\hat{r}$, to relate $\mathcal{T}$ to d copies of $\mathcal{T}_1$ and d copies of $\mathcal{T}_2$ (we use the same notation $\hat{r}, \hat{r} + \aleph$ for the root indices of $\mathcal{T}_1$ and $\mathcal{T}_2$, and it be clear from the context to which tree's root we refer each time). Likewise, removing the root of $G_{\mathcal{T}_1}$ yields d copies of $\mathcal{T}_1$ and $d-1$ copies of $\mathcal{T}_2$, where the index $\hat{r}$ of the original $\mathcal{T}_1$ was connected to the $\hat{r} + \aleph$ index of its children (namely, the d copies of the subtree $\mathcal{T}_1$), while the $\hat{r} + \aleph$ index of the original $\mathcal{T}_1$ was connected to the $\hat{r}$ index of its children (namely, the $d-1$ copies of the subtree $\mathcal{T}_2$). Moreover, having no contribution from $H|_{T \times T^c} G^{(T)} H|_{T^c \times T}$ to the entries indexed by $\hat{r}, \hat{r} + \aleph$ or $\hat{r} + \aleph, \hat{r}$, it thus follows from (1.24) that

$$(1.27) \qquad \begin{bmatrix} (G_{\mathcal{T}_1})_{\hat{r},\hat{r}} & (G_{\mathcal{T}_1})_{\hat{r},\hat{r}+\aleph} \\ (G_{\mathcal{T}_1})_{\hat{r}+\aleph,\hat{r}} & (G_{\mathcal{T}_1})_{\hat{r}+\aleph,\hat{r}+\aleph} \end{bmatrix}^{-1} = -\begin{bmatrix} z & w \\ \overline{w} & z \end{bmatrix} - \begin{bmatrix} \frac{d}{d-1}(G_{\mathcal{T}_1})_{\hat{r}+\aleph,\hat{r}+\aleph} & 0 \\ 0 & (G_{\mathcal{T}_2})_{\hat{r},\hat{r}} \end{bmatrix}.$$

Similarly, removing the root of $G_{\mathcal{T}_2}$ results with d subtrees $\mathcal{T}_2$ and $d-1$ subtrees $\mathcal{T}_1$. The index $\hat{r}$ of the original $\mathcal{T}_2$ was connected to the root $\hat{r} + \aleph$ of the $d-1$ subtrees $\mathcal{T}_1$, while $\hat{r} + \aleph$ was connected to the root $\hat{r}$ of the d subtrees $\mathcal{T}_2$, so by Schur complement formula,

$$(1.28) \qquad \begin{bmatrix} (G_{\mathcal{T}_2})_{\hat{r},\hat{r}} & (G_{\mathcal{T}_2})_{\hat{r},\hat{r}+\aleph} \\ (G_{\mathcal{T}_2})_{\hat{r}+\aleph,\hat{r}} & (G_{\mathcal{T}_2})_{\hat{r}+\aleph,\hat{r}+\aleph} \end{bmatrix}^{-1} = -\begin{bmatrix} z & w \\ \overline{w} & z \end{bmatrix} - \begin{bmatrix} (G_{\mathcal{T}_1})_{\hat{r}+\aleph,\hat{r}+\aleph} & 0 \\ 0 & \frac{d}{d-1}(G_{\mathcal{T}_2})_{\hat{r},\hat{r}} \end{bmatrix}.$$

Since, as mentioned earlier,

$$(1.29) \qquad \begin{bmatrix} (G_{\mathcal{T}})_{\hat{r},\hat{r}} & (G_{\mathcal{T}})_{\hat{r},\hat{r}+\aleph} \\ (G_{\mathcal{T}})_{\hat{r}+\aleph,\hat{r}} & (G_{\mathcal{T}})_{\hat{r}+\aleph,\hat{r}+\aleph} \end{bmatrix}^{-1} = -\begin{bmatrix} z & w \\ \overline{w} & z \end{bmatrix} - \begin{bmatrix} \frac{d}{d-1}(G_{\mathcal{T}_1})_{\hat{r}+\aleph,\hat{r}+\aleph} & 0 \\ 0 & \frac{d}{d-1}(G_{\mathcal{T}_2})_{\hat{r},\hat{r}} \end{bmatrix},$$

we see that all Green's function entries are derived from $(G_{\mathcal{T}_1})_{\hat{r}+\aleph,\hat{r}+\aleph}$ and $(G_{\mathcal{T}_2})_{\hat{r},\hat{r}}$. Further, the investigation of the path expansion of the Green's function at large $|z|$ suggests that $(G_{\mathcal{T}_1})_{\hat{r}+\aleph,\hat{r}+\aleph} = (G_{\mathcal{T}_2})_{\hat{r},\hat{r}}$. Proceeding under this ansatz and denoting the latter quantity by m_∞, we deduce from (1.27) by the inversion formula for 2×2 matrices, that

$$(1.30) \qquad m_\infty = \frac{z + \frac{d}{d-1}m_\infty}{|w|^2 - \left(z + \frac{d}{d-1}m_\infty\right)(z + m_\infty)}.$$

Fixing $w \in \mathbb{C}$, by the implicit function theorem in terms of (m_∞, z^{-1}) at $(0,0)$, there exists a unique analytic solution $m_\infty(\cdot, w) : \mathbb{C}_+ \mapsto \mathbb{C}_+$ of the cubic equation (1.30), such that $z\, m_\infty(z, w) \to -1$ when $|z| \to \infty$. Using hereafter this solution of (1.30), and suppressing the arguments z, w, we get from the central parameter m_∞ also the quantities

$$(1.31) \qquad \begin{aligned} m_\infty^{sd} &:= \frac{z + m_\infty}{z + \frac{d}{d-1}m_\infty}m_\infty, & m_{\mathcal{T}}^{d} &:= \frac{z + \frac{d}{d-1}m_\infty}{|w|^2 - \left(z + \frac{d}{d-1}m_\infty\right)^2}, \\ m_\infty^{uod} &:= \frac{-w}{z + \frac{d}{d-1}m_\infty}m_\infty, & m_{\mathcal{T}}^{uod} &:= \frac{-w}{z + \frac{d}{d-1}m_\infty}m_{\mathcal{T}}^{d}, \\ m_\infty^{lod} &:= \frac{-\overline{w}}{z + \frac{d}{d-1}m_\infty}m_\infty, & m_{\mathcal{T}}^{lod} &:= \frac{-\overline{w}}{z + \frac{d}{d-1}m_\infty}m_{\mathcal{T}}^{d}, \end{aligned}$$

where $m_\infty^{sd} := (G_{\mathcal{T}_1})_{\hat{r},\hat{r}}$, $m_\infty^{uod} := (G_{\mathcal{T}_1})_{\hat{r},\hat{r}+\aleph}$ and $m_\infty^{lod} := (G_{\mathcal{T}_1})_{\hat{r}+\aleph,\hat{r}}$ are the remaining entries of $G_{\mathcal{T}_1}$ (see the LHS of (1.27)). Upon exchanging $\hat{r}$ with $\hat{r} + \aleph$, these are also the entries of $G_{\mathcal{T}_2}$ (see (1.28)), whereas $m_{\mathcal{T}}^{d} := (G_{\mathcal{T}})_{\hat{r},\hat{r}} = (G_{\mathcal{T}})_{\hat{r}+\aleph,\hat{r}+\aleph}$, $m_{\mathcal{T}}^{uod} := (G_{\mathcal{T}})_{\hat{r},\hat{r}+\aleph}$ and $m_{\mathcal{T}}^{lod} := (G_{\mathcal{T}})_{\hat{r}+\aleph,\hat{r}}$ (see (1.29)).

In particular, it is not hard to verify that

$$(1.32) \qquad m_{\mathcal{T}}^{d}(z, w) = \sqrt{d-1}\, m_\star(\sqrt{d-1}z, \sqrt{d-1}w).$$

For details, one can refer to lemma A.9.

1.4. **Extension of graphs and main result.** A key ingredient in the asymptotic evaluation of the Green's function for our digraphs, is the following concept of graph extension (or more precisely, the extension of the associated weighted adjacency matrix).

Definition 1.11 (Extensions of graphs). Suppose $\mathcal{G}$ is a digraph of vertex set V, where each $v \in V$ has in-degree $d_{in}(v) \leq d$ and out-degree $d_o(v) \leq d$. Any given deficit functions $0 \leq def_{in}(v) \leq d - d_{in}(v)$, $0 \leq def_o(v) \leq d - d_o(v)$, and two functions Δ_1 and Δ_2 of z, w, induce the following extension of the matrix H which corresponds to $\mathcal{G}$ via (1.21) :

$$\text{Ext}_{\text{def}}(\mathcal{G}, \Delta_1, \Delta_2) =$$

$$(1.33) \qquad H - \sum_{v \in V} \left[\frac{d - d_o(v) - \text{def}_o(v)}{d-1}\Delta_1 e_{v,v} + \frac{d - d_{in}(v) - def_{in}(v)}{d-1}\Delta_2 e_{v+\aleph,v+\aleph} \right],$$

where $e_{v,v}$ and $e_{v+\aleph,v+\aleph}$ denote, respectively, the (v,v) and $(v+\aleph, v+\aleph)$ coordinates of the matrix (with isolated vertices in V also considered in the sum). The Green's function $G(\text{Ext}_{\text{def}}(\mathcal{G}, \Delta_1, \Delta_2))$ associated with such extension is merely the inverse of the RHS of (1.33). In particular, we use the notation Ext for

a graph extension where both def_{in} and def_o are identically zero. Similarly, Ext_i denotes the case where a specific vertex $\hat{r}$ (the root) has $def_{in}(\hat{r}) = 1$, while all other values of def_o and def_{in} are zero, whereas in Ext_o one has that $def_o(\hat{r}) = 1$, while all other values of the deficit functions are zero. That is,

$$(1.34) \qquad \text{Ext}_i(\mathcal{G}, \Delta_1, \Delta_2) = H - \sum_{v \in V} \left[\frac{d - d_o(v)}{d - 1} \Delta_1 e_{v,v} + \frac{d - d_{in}(v) - \delta_{v,\hat{r}}}{d - 1} \Delta_2 e_{v+\aleph, v+\aleph} \right].$$

$$(1.35) \qquad \text{Ext}_o(\mathcal{G}, \Delta_1, \Delta_2) = H - \sum_{v \in V} \left[\frac{d - d_o(v) - \delta_{v,\hat{r}}}{d - 1} \Delta_1 e_{v,v} + \frac{d - d_{in}(v)}{d - 1} \Delta_2 e_{v+\aleph, v+\aleph} \right].$$

We denote by $B_\ell(v)$ the unoriented ball in digraph $\mathcal{G}$ of radius ℓ centered at vertex v, with $\partial B_\ell(v)$ denoting the set of vertices of $\mathcal{G}$ having an undirected distance ℓ from v. More generally, for a subset U of vertices of $\mathcal{G}$, we denote by $B_\ell(U)$ the union of such balls $B_\ell(v)$ for $v \in U$. With these quantities at hand, our main result, stated next, is a refinement of Theorem 1.3, providing also a uniform over (i,j), control on the entries $G_{i,j}$ of $G(z,w)$.

Theorem 1.12. *Fix a constant $0 < \mathfrak{c} < 1$ and $d \geq 3$. Given these terms, we can find parameters $\delta = \delta(\mathfrak{c}, d) > 0$, $\varepsilon = \varepsilon(\mathfrak{c}, d) > 0$ and $\mathfrak{r} = \mathfrak{r}(\mathfrak{c}, d)$ such that for $r = r_N = \mathfrak{r} \log_{d-1} N$, some set of d-regular random digraphs $\mathcal{G}$ occurring with probability $\mathbb{P}_{N,d}$ at least $1 - O(N^{-\mathfrak{c}})$ and any compact $\mathbb{D} \subset \mathbb{C}$,*

$$(1.36) \qquad \sup_{w \in \mathbb{D}, \Im(z) \geq N^{-\varepsilon}} \max_{ij} |G_{i,j} - G_{i,j}(Ext(B_r(\{i,j\}), m_\infty, m_\infty))| \leq N^{-\delta},$$

$$(1.37) \qquad \sup_{w \in \mathbb{D}, \Im(z) \geq N^{-\varepsilon}} \left| \frac{1}{N} \sum_{i=1}^{N} G_{i,i} - m_{\mathcal{T}}^d \right| \leq N^{-\delta}.$$

Theorem 1.12 will be shortly deduced from our next two results, the first of which is a perturbation theory on infinite d-regular digraphs.

Theorem 1.13. *Let $\mathcal{G}$ be a graph with excess of at most 1. Assume further that $diam(\mathcal{G})^3 [|Q_I - m_\infty| + |Q_O - m_\infty|] \ll 1$. Then, we have,*

$$(1.38) \qquad \begin{aligned} &|G_{ij}(Ext(\mathcal{G}, Q_I, Q_O)) - G_{ij}(Ext(\mathcal{G}, m_\infty, m_\infty))| \\ &\lesssim \left(\frac{1}{d-1} \right)^{dist(i,j)} (1 + dist(i,j))[|Q_I - m_\infty| + |Q_O - m_\infty|] \end{aligned}$$

We will elaborate on the meanings of Q_I and Q_O in the next section. For now, just understand these terms as a tool that help mediate the derivation of an appropriate self-consistent equation. Our next result states the relationship between G_{ij} and the d-regular infinite tree extension, except with parameters Q_I and Q_O. The specific definitions of the error terms ($\epsilon(z,w)$ and S_g^1), can also be found in the next section.

Theorem 1.14. *Fix $d \geq 3$, $0 < \mathfrak{c} < 1$ and a compact domain $\mathbb{D} \subset \mathbb{C}$. There exist $\varepsilon = \varepsilon(\mathfrak{c}, d) > 0$ and $r = r_N = \mathfrak{r} \log_{d-1} N$, such that on a set of d-regular, random digraphs occurring with probability $1 - O(N^{-\mathfrak{c}})$, all $w \in \mathbb{D}$ and $\Im(z) \geq N^{-\varepsilon}$ we have*

$$(1.39) \qquad \begin{aligned} &|Q_{I/O} - m_\infty(z,w)| \leq \frac{\epsilon(z,w)}{S_g^1}, \\ &\max_{i,j} |G_{ij} - G_{ij}(Ext(B_\ell(\{i,j\}, \mathcal{G}), Q_I, Q_O))| \leq \epsilon(z,w). \end{aligned}$$

(See Part 2 of Section 2 for a formal definition of $\epsilon(z,w)$, which are independent of $\mathcal{G}$. We remark also that S_g^1 of (3.1) is a deterministic constant depending only on z and w.)

Proof of Theorem 1.12 (from Theorems 1.13 and 1.14). With $|Q_{I/O} - m_\infty(z,w)| \leq \frac{\epsilon(z,w)}{S_g^1}$, we can apply the perturbation estimate from Lemma 3.5 to ensure that for $r = r_N = \mathfrak{r} \log N$,

$$|G_{ij}(Ext(B_r(\{i,j\}, \mathcal{G}), m_\infty(z,w), m_\infty(z,w))) - G_{ij}(Ext(B_r(\{i,j\}, \mathcal{G}), Q_I, Q_O))| \lesssim (\log N) \frac{\epsilon}{S_g^1}.$$

This combined with the second estimate in (1.39), ensures that,

$$(1.40) \qquad |G_{ij} - G_{ij}(Ext(B_r(\{i,j\},\mathcal{G}), m_\infty(z,w), m_\infty(z,w)))| \lesssim \epsilon + (\log N)\frac{\epsilon}{S_g^1}.$$

Now, if i has a radius $\mathfrak{R}$ tree-like neighborhood, we would know that

$$G_{ii}(Ext(B_r(\{i,j\},\mathcal{G}), m_\infty(z,w), m_\infty(z,w))) = m_{\mathcal{T}}^d.$$

There are also at most $O(N^{\mathfrak{c}})$ vertices without a radius $\mathfrak{R}$ tree-like neighborhood. For these vertices, we can instead use the bound that $|G_{ii}| = O(1)$.

This again comes from Lemma 3.5 to bound $G_{ij}(Ext(B_r(\{i,j\},\mathcal{G}), Q_I, Q_O))$ by $O(1)$ and the triangle inequality. This ensures that,

$$(1.41) \qquad \Big|\frac{1}{N}\sum_{i=1}^{N} G_{ii} - m_{\mathcal{T}}^d\Big| \lesssim N^{\mathfrak{c}-1} + \epsilon + (\log N)\frac{\epsilon}{S_g^1}.$$

Finally, recall (A.4) that here $S_g^1 \geq \frac{1}{2}N^{-\varepsilon}$ and (2.9) that $\epsilon(z,w) \leq N^{-\delta'}$ for any $\delta' < \min(\mathfrak{r}, \frac{1-\varepsilon}{2})$. We thus deduce the stated uniform bound of order $N^{-\delta}$, upon choosing $\varepsilon(\mathfrak{c},\mathfrak{r}) > 0$ small enough. $\qquad\square$

2. Challenges and outline for proving Theorem 1.14

Sharp estimates on the Green's function are key to many works in random matrix theory. A standard technique to establish those is by first deriving a self-consistent equation for this Green's function, which under suitable stability properties, allows for a multi-scale argument transferring the estimates on the Green's function from one value of $\Im[z]$ to a lower value of $\Im[z]$ (see [12]).

This strategy is used in [18], where the control parameter Q of [18, Eqn. (1.6)], which represents the averaged Green's function of the graph with one edge removed, satisfies the self-consistent equation of [18, Eqn. (1.7)]. It roughly says that one can determine the Green's function by treating a local neighborhood of large size K around each vertex as part of an infinite tree. We proceed to relate our work with [18] describing also the new challenges that appear in the context of our analysis. First, note that to deal with digraphs, we must take into account whether an in-edge or an out-edge has been removed. To this end, we associate two vertices v and $v + N$ to every vertex v in the original graph, resulting with a pair of control parameters

$$(2.1) \qquad \begin{aligned} Q_I &:= \frac{1}{Nd} \sum_{(x\to y)\in E} G_{y+N,y+N}^{(x,x+N)}, \\ Q_O &:= \frac{1}{Nd} \sum_{(y\to x)\in E} G_{y,y}^{(x,x+N)}. \end{aligned}$$

Further, our basic removal operation via the Schur-complement formula, is now the removal of a two-dimensional submatrix rather than a single entry, resulting with a pair of self-consistent equations

$$(2.2) \qquad \begin{aligned} Q_I &= Y_{i,K}(Q_I, Q_O), \\ Q_O &= Y_{o,K}(Q_I, Q_O), \end{aligned}$$

representing the treatment of a neighborhood of size K of our digraph as part of the d-regular directed tree.

The general outline for deriving (2.2) is as in [18]. Namely, apply the Schur complement formula to describe the effects of removing a large neighborhood of a vertex, then perform a switching transformation and finally add back the removed neighborhood. However, the intermediate estimates required for implementing this strategy differ drastically between our digraph case and the non-oriented case. In particular, the task of proving many estimates in [18] is much simplified by having the Green's function estimates depend only on the graph distance between vertices. For example, utilizing the identity (4.4), namely

$$(2.3) \qquad G^{(\mathbb{T})} - P^{(\mathbb{T})} = G - P - (G-P)(G|_{\mathbb{T}})^{-1}G - P[(G|_{\mathbb{T}})^{-1} - (P|_{\mathbb{T}})^{-1}]G - P(P|_{\mathbb{T}})^{-1}(G-P),$$

one aims to show that the LHS is small by relying on a-priori control on the size of the terms $P_{i,j}$ on its RHS. Here $P_{i,j}$ roughly corresponds to the Green's function of vertices i and j in the infinite tree, which are of distance $O(\log \log N)$. Having removed a large neighborhood, one is to sum over $O((\log N)^p)$ such terms and it is crucial to have a contractive effect of the total contribution, so the total error on the RHS will not get exponentially worse as we remove larger and larger pieces (when N grows).

In the digraph case, when bounding $P_{i,j}$ the contribution of each path between the two vertices depends on the orientations of the various edges along that path. Applying the worst-case value over such orientations yield bounds which grow exponentially in the size of the neighborhood removed and thereby to a complete failure of the general strategy. To address this challenge and ensure that the perturbation estimates on the Green's function are contractive, we leverage on having certain Green's function terms which are better than the worst case bound and carefully enumerate over all types of factors that may appear. Specifically, see Lemma A.1, where the contractive factor is $X + Y < 1$ but a less careful analysis of the perturbation estimates would lead to an expansive factor of the form $2X > 1$. For this reason, we also can not apply the intermediate lemmas of [18] which rely on the simplification offered by the Green's function structure of the infinite unoriented regular tree (that we lack here).

Note also another major difficulty in our self-consistent equations preventing us from obtaining estimates on the Green's function reaching the optimal $\Im[z] \asymp N^{-1+\epsilon}$. Indeed, the linearization (3.33) of our self-consistent equations is very singular as $\Im[z]$ gets close to this critical value. This problem is circumvented in [18] by a second order expansion of their equation for Q and explicitly solving for the solutions of the resulting quadratic equation. Applying such approach for our (2.2) yields a pair of quadratic equations, reduced to a single quartic equation, whose solutions are not analytically tractable. As already shown, our primary goal of estimating the Green's function well enough for deriving the global limit law of d-regular digraphs does not require descending to the critical threshold of $\Im[z]$, and we thus settle for the linearized self-consistent equations.

2.1. **Part 1: Heuristics of the switching argument.** We proceed to provide the heuristics behind our derivation of the self-consistent equations for the Green's function of the random d-regular digraph. To this end, recall the important control quantities Q_I and Q_O of (2.1) (and as before, we suppress the arguments z, w whenever their specific value is irrelevant). Such quantities deal with the effect at of removing an edge on the Green's function of the digraph. Specifically, $G^{(x,x+N)}_{y+N,y+N}$ measures the effect at y of removing the incoming edge $x \to y$, and $G^{(x,x+N)}_{y,y}$ likewise measures the effect at y of removing the outgoing edge $y \to x$. Thus, our control quantities Q_I and Q_O measure the average effect on the Green's function due to the removal of a uniformly chosen incoming edge, or outgoing edge, respectively. As outlined next, our proof strategy consists of the following two steps:

(a) Our first step relates the Green's function of $\mathcal{G}$ around a vertex v to the Green's function of the $\ell(= \ell_N)$ neighborhood $\mathcal{N}_\ell(v)$ of v for $\ell = O(\log \log N)$.
(b) Performing a switching of the edges connecting $\mathcal{N}_\ell(v)$ to $\mathcal{N}_\ell^c(v)$ averages out the effects of specific edges, leading to Q_I and Q_O. We then show that the latter functions concentrate for large N around the non-random m_∞.

2.1.1. *Reduction to the local K neighborhood.* Ignoring for the time being the specific value of ℓ, consider the relation via the Schur complement formula between entries of the Green's function of $\mathcal{G}$ within $\mathcal{N}_K(v)$ and the matrix H restricted to this local neighborhood. Specifically, denote by $\tilde{\mathcal{N}}_K(v)$ the collection of indices in $\mathcal{N}_K(v)$ and $\mathcal{N}_K(v) + N$, with $O_{\mathcal{N}_K(v)}$ denoting the collection of directed edges that start in $\mathcal{N}_K(v)$ and terminate in $\mathcal{N}_K^c(v)$, while $I_{\mathcal{N}_K(v)}$ denotes the directed edges that start in $\mathcal{N}_K^c(v)$ and terminate in $\mathcal{N}_K(v)$. Considering (1.24) when $T = \tilde{\mathcal{N}}_K(v)$, yields that for any pair of indices $c, d \in \tilde{\mathcal{N}}_K(v)$

$$(2.4) \qquad G_{c,d} = \left[\left(H|_{\tilde{\mathcal{N}}_K(v)} - \frac{1}{d-1} \left[\sum_{(b \to a) \in I_{\mathcal{N}_K(v)}} G^{(\tilde{\mathcal{N}}_K(v))}_{b,b} e_{a+N,a+N} + \sum_{(a \to b) \in O_{\mathcal{N}_K(v)}} G^{(\tilde{\mathcal{N}}_K(v))}_{b+N,b+N} e_{aa} \right] \right)^{-1} \right]_{cd}.$$

We now proceed with the following approximations of the two sums on the RHS. First one would expect that since our d-regular digraph is locally tree-like around v (when K does not grow fast with N), the value of

$G_{b,b}^{(\tilde{\mathcal{N}}_K(v))}$ should only depend on the local behavior around b of the restricted graph $\mathcal{G}|\mathcal{N}_K^c(v)$. Moreover, by the same reasoning, locally around b, the digraph $\mathcal{G}|\mathcal{N}_K^c(v)$ should be similar to the original graph $\mathcal{G}$ with only the vertex a removed. We thus expect that, at least for most edges involved in the first sum of (2.4),

$$G_{b,b}^{(\tilde{\mathcal{N}}_K(v))} \approx G_{b,b}^{(a,a+N)}.$$

After such replacement, assuming K is large enough, we may expect to be able to do even further averaging, thereby replacing each $G_{bb}^{(a,a+N)}$ by the total average of terms of this form along all Nd directed edges of our d-regular digraph. Namely, fixing $a \in \mathcal{N}_K(v)$,

$$\sum_{\{b:(b \to a) \in I_{\mathcal{N}_K(v)}\}} G_{b,b}^{(a,a+N)} \to (d - d_i(a))\frac{1}{Nd}\sum_{(y \to x) \in E} G_{y,y}^{(x,x+N)} = (d - d_i(a))Q_O,$$

where $d_i(a)$ denotes the number of in-edges of a within $\mathcal{G}|\mathcal{N}_K(v)$. By the same reasoning, one can also expect to have per fixed $a \in \mathcal{N}_K(v)$,

$$\sum_{\{b:(a \to b) \in O_{\mathcal{N}_K(v)}\}} G_{b+N,b+N}^{(a,a+N)} \to (d - d_o(a))\frac{1}{Nd}\sum_{(x \to y) \in E} G_{y+N,y+N}^{(x,x+N)} = (d - d_o(a))Q_I.$$

Taken together, this would posit an almost equality

$$G_{c,d} \approx G(Ext(\mathcal{G}|\mathcal{N}_K(v), Q_I, Q_O))_{c,d}$$

for any pair of indices $c, d \in \tilde{\mathcal{N}}_K(v)$, thereby 'reducing' the evaluation of the Green's function to the study of the two order parameters Q_I and Q_O of (2.1).

Turning to the latter task, we proceed to derive an equation for the terms $G_{y,y}^{(x,x+N)}$ by applying a similar reduction to the local K neighborhood of y, now for the graph $\mathcal{G}$ after the removal of $(y \to x)$. Specifically, let $\mathcal{N}_K^{(y \to x)}(y)$ denote the K neighborhood of the digraph, rooted at y, after removing the edge $(y \to x)$. Applying the same heuristics as done earlier, we arrive at

$$(2.5) \qquad\qquad G_{y,y}^{(x,x+N)} \approx G(Ext_o(\mathcal{G}|\mathcal{N}_K^{(y \to x)}(y), Q_I, Q_O))_{y,y}$$

(where the removal of $(y \to x)$ yields the deficiency $def_o(y) = 1$ at the root). Similarly, one would propose that

$$(2.6) \qquad\qquad G_{y+N,y+N}^{(x,x+N)} \approx G(Ext_i(\mathcal{G}|\mathcal{N}_K^{(x \to y)}(y), Q_I, Q_O))_{y+N,y+N}.$$

2.1.2. *Tree-like neighborhoods and self-consistency.*

For any $K \geq 1$, let $\mathcal{T}_1^K$ denote the $(2d-1)$-ary directed tree of K levels, whose root vertex has d out-edges and $d-1$ in-edges. Similarly, let $\mathcal{T}_2^K$ denote the $(2d-1)$-ary directed tree of K levels, whose root vertex has $d-1$ out-edges and d in-edges. Define for functions $\Delta_1(z,w)$, $\Delta_2(z,w)$ on $\mathbb{C}_+ \times \mathbb{C}$ the functions

$$(2.7) \qquad \begin{aligned} Y_{i,K}(\Delta_1, \Delta_2) &:= G(Ext_i(\mathcal{T}_1^K), \Delta_1, \Delta_2)_{r+\aleph, r+\aleph}, \\ Y_{o,K}(\Delta_1, \Delta_2) &:= G(Ext_o(\mathcal{T}_2^K), \Delta_1, \Delta_2)_{r,r}, \end{aligned}$$

where r and $r + \aleph$ are the two indices corresponding to the root of the relevant tree.

To determine the asymptotic of Q_I and Q_O we average the approximations (2.5) and (2.6) over all directed edges of our digraph. While doing so, we note that most neighborhoods $\mathcal{N}_K^{(x \to y)}(y)$ should look like the $(2d-1)$-ary directed tree $\mathcal{T}_1^K$ of K levels, whose root vertex has d out-edges and $d-1$ in-edges. Similarly, most neighborhoods $\mathcal{N}_K^{(y \to x)}(y)$ should look like the $(2d-1)$-ary directed tree $\mathcal{T}_2^K$ of K levels, whose root vertex has $d-1$ out-edges and d in-edges. Thus, after averaging, we expect Q_I and Q_O to approximately satisfy the following self-consistency equations

$$(2.8) \qquad \begin{aligned} Q_I &= Y_{i,K}(Q_I, Q_O), \\ Q_O &= Y_{o,K}(Q_I, Q_O). \end{aligned}$$

Further, by definition the pair (Q_I, Q_O) should not depend on $K \geq 1$, and for any $w \in \mathbb{C}$, either function should be analytic from $\mathbb{C}_+$ to $\mathbb{C}_+$ and such that $zQ_I(z,w) \to -1$ and $zQ_O(z,w) \to -1$ when $|z| \to \infty$.

In view of our expectation that $Q_I \approx Q_O$, we verify that the unique solution $Q_I = Q_O$ of (2.2) with such analytic properties is $m_\infty(z,w)$ of (1.30).

2.2. **Part 2: The structure of the proof of Theorem 1.14.** The proof of Theorem 1.14 is based on careful and technical estimates of Green's functions after performing a switching. In order to even discuss our proof, it is first necessary to devote some time to defining appropriate notation and the correct order parameters. The first two sub-subsections here are devoted to introducing the appropriate notation. After defining all necessary notions, we then state our important results on Green's function estimates for graphs after switching. These are Theorems 2.9 and 2.11. After some computations, we will show that Theorem 1.14 follows from these substeps.

2.2.1. *Notation related to Graph Structure.* The goal of this section is to introduce the notation that we will use later as well as give a broad overview of the strategy of the proof. The sections that will appear later will go over many of the detailed estimates.

We will start with definitions related to the graph.

Definition 2.1 (Ball of Radius R). Let $\mathcal{G} = (V, E)$ be a digraph and let S be some set of vertices in V; here, we will let $\mathcal{G}^u$ denote the unoriented version of $\mathcal{G}$. We let $B_R(S, \mathcal{G})$ denote the set of vertices that are of distance R from the vertices V in the unoriented graph $\mathcal{G}^u$. If S consists of a single element v, then we will use $B_R(v, \mathcal{G})$ as a shorthand for $B_R(\{v\}, \mathcal{G})$; $B_R(S, \mathcal{G})$ will also be used to denote the subgraph induced by the vertices in the set; which version is used will be clear in context. In what follows later in the paper, we will always use distance between vertices to refer to the distance between vertices in the unoriented graph $\mathcal{G}^u$. As this is the only reason to refer to the unoriented version $\mathcal{G}^u$; we will not use this notation later and let the reader understand that all distances refer to the unoriented distance.

We will say that the vertex v has a radius R tree-like neighborhood in $\mathcal{G}$ is the unoriented structure of the subgraph $B_R(v, \mathcal{G})$ is a tree. If this neighborhood is not tree-like, then we can define the excess of the neighborhood as the number of edges that need to be removed from the subgraph $B_R(v, \mathcal{G})$ that need to be removed in order to make the neighborhood tree-like.

Our method also requires the introduction of many scales of order parameters.

Definition 2.2 (Order Parameters).   (1) We define the value $\mathfrak{R} := \frac{\mathfrak{c}}{4} \log_{d-1} N$. This value is chosen so that nearly all vertices will have a radius $\mathfrak{R}$ tree-like neighborhood. We will show that with high probability, almost all vertices have a radius $\mathfrak{R}$ tree-like neighborhood.
   (2) We introduce the value r which satisfies the inequalities $\frac{\mathfrak{R}}{16} \leq r \leq \frac{\mathfrak{R}}{8}$. Alternatively, $r := \mathfrak{r} \log_{d-1} N$ with $\frac{\mathfrak{c}}{64} \leq \mathfrak{r} \leq \frac{\mathfrak{c}}{32}$. The purpose of this parameter r is to ensure that the Green's functions of two graphs that agree on a neighborhood of size r will be very close to one another. This is the statement of Lemma 3.7.
   (3) Finally, we have the value $\ell \in [\mathfrak{a} \log_{d-1} \log N, 2\mathfrak{a} \log_{d-1} \log N]$, with $\mathfrak{a} \geq 12$. The specific value of ℓ is chosen so that the stability estimates in the self-consistent equation from Lemma 3.8 hold. ℓ represents the size of the neighborhood on which we perform a switching argument.

We have the following order of scales on our constants: $\mathfrak{R} > r \gg l$.

Indeed with our order parameters, we can make the following definition of good d-regular digraphs.

Definition 2.3 (Good d-regular digraphs). We define the set Ω (the set of good d-regular digraphs) to be the set of d-regular digraphs on N vertices that satisfy the following two properties.
   (1) All radius $\mathfrak{R}$ neighborhoods have excess at most 1.
   (2) All except $N^{\mathfrak{c}}$ vertices have neighborhoods of size $\mathfrak{R}$ that are tree-like.

Fix some parameter $\mathfrak{q} > 0$. This constant $\mathfrak{q}$ dictates the probability that a 'good' switching event occurs with probability $1 - O(N^{-\mathfrak{q}})$. Depending on this parameter $\mathfrak{q}$, we can define the set $\Omega_{S,\mathfrak{q}}$ of the graphs that have sufficiently nice structural properties after switching. A graph belongs in $\Omega_{S,\mathfrak{q}}$ if it satisfies the following two properties,
   (1) All except $2N^{\mathfrak{c}}$ vertices have neighborhoods of size $\mathfrak{R}$ that are tree-like.

(2) The $\mathfrak{R}/8$ neighborhood of any vertex has excess at most 1 and the radius $\mathfrak{R}/2$ neighborhood of any vertex has excess at most $C_{\mathfrak{q}}$. $C_{\mathfrak{q}}$ is a constant depending on $\mathfrak{q}$ but not on N.

2.2.2. *Notation related to the Switching Procedure.* With our order parameters defined, we can now define our set of switching data. In order to avoid an overly long definition, we split the discussion of the switching procedure into two parts. The first part will introduce useful notation, while the second part will actually describe the procedure used in the switching.

Definition 2.4 (Switching Data for a Radius ℓ Neighborhood). Consider a d-regular digraph $\mathcal{G}$ with N vertices. Consider a vertex o and its radius ℓ neighborhood $\mathbb{T} := B_\ell(o, \mathcal{G})$.

We define the boundary of $\mathbb{T}$ to be the vertices of $\mathbb{T}$ that are connected to a vertex in $\mathcal{G}$. The boundary edges of $\mathbb{T}$ are the set of edges that connect a vertex in the boundary of $\mathbb{T}$ to a vertex in $\mathbb{T}^c$. We can index these boundary edges as follows: $\{(l_\alpha, a_\alpha)\}$; l_α is chosen to be the vertex that lies in $\mathbb{T}$, while a_α is chosen to be the vertex that is in $\mathbb{T}^c$. We will use μ to denote the number of such edges connecting $\mathbb{T}$ to $\mathbb{T}^c$. We remark that if $\mathbb{T}$ were a purely treelike neighborhood, then $\mu = 2d(2d-1)^{\ell-1}$. Furthermore, in a treelike neighborhood then $\mu/2$ of these edges are oriented in the direction $l_\alpha \to a_\alpha$ and $\mu/2$ of these edges are oriented in the direction $a_\alpha \to l_\alpha$.

Since the edges can have different orientations, we have to be careful regarding the choice of the switching when constructing the switching data. Our switching data $S = (S_1, \ldots, S_\mu)$ will consist of a choice of $|\mu|$ random edges that lie strictly inside $\mathbb{T}^c$ in $\mathcal{G}$; these will be indexed as (b_α, c_α), but the specific value of b_α and c_α will depend on the orientation. If the edge were oriented in the direction l_α to a_α, then the indices are chosen so that (b_α, c_α) were oriented in the direction b_α to c_α. If instead the vertices we had the orientation a_α to l_α, then the edge (b_α, c_α) will be oriented c_α to b_α. The purpose of this is to ensure that a_α will naturally be switched with c_α.

The collection S is chosen so that for each $\alpha \leq |\mu|$, each edge b_α, c_α is chosen independently and uniformly over all edges that lie in $\mathbb{T}^c$. Note that this will allow for repetition, but such collisions are highly unlikely and will not make any difference in the ultimate analysis. The set of all possible switching data corresponding to a graph $\mathcal{G}$ around the vertex o will be denoted as $\mathbb{S}_{\mathcal{G},o}$. The vertices o are naturally identified in $\mathcal{G}$ and $\tilde{\mathcal{G}}$.

We continue our discussion of the switching procedure by actually describing the graphs that are constructed by switching.

Definition 2.5 (The Switching Construction). Recall the digraph $\mathcal{G}$, the neighborhood $\mathbb{T}$ around o, as well as the switching data S from Definition 2.4.

For each α, we define the indicator function variable χ_α which will take values 1 or 0 based on the following conditions. $\chi_\alpha = 1$ if and only if the subgraph $B_{\mathfrak{R}/4}(\{a_\alpha, b_\alpha, c_\alpha\}, \mathcal{G}\backslash\mathbb{T})$ is a tree even after adding the edge connecting a_α to b_α and , furthermore, the vertices $\{a_\alpha, b_\alpha, c_\alpha\}$ is of distance at least $\mathfrak{R}/4$ from all other switching vertices $\{a_\beta, b_\beta, c_\beta\}$ for $\beta \neq \alpha$.

The switching map T can be thought of as a bijective map from pairs of d-regular digraphs and switching data $T : (\mathcal{G}, S) \to (\tilde{\mathcal{G}}, T(S))$ with $S \in \mathbb{S}_{\mathcal{G},o}$ and $T(S) \in \mathbb{S}_{\tilde{\mathcal{G}},o}$. There is a natural way to associate the vertices o in both $\mathcal{G}$ and $\tilde{\mathcal{G}}$.

To construct $\tilde{\mathcal{G}}$, we perform the following procedure. If $\chi_\alpha = 1$, then we remove the edge (l_α, a_α) and connect l_α to c_α as well as b_α to a_α. The orientation will be such that the orientation of l_α to c_α is the same as the orientation of l_α to a_α; in addition, the orientation of b_α to a_α will match the orientation of b_α to c_α. We also set $\tilde{a}_\alpha := c_\alpha$. However, if $\chi_\alpha = 0$, we make no change to the graph and set $\tilde{a}_\alpha := a_\alpha$.

If $\chi_\alpha = 1$, then we change the switching data S_α to the edge (a_α, b_α) with the proper orientation. However, if $\chi_\alpha = 0$, we make no change. It is easy to check that T is an involution and preserves probability.

We now give the statement of the following proposition. This will show that all except an $O(1)$ indices α will satisfy $\chi_\alpha = 1$.

Proposition 2.6. *Fix some constant $\mathfrak{q} > 0$ and some graph $\mathcal{G}$ in either Ω or $\Omega_{S,\mathfrak{q}}$. Fix a vertex o and its ℓ neighborhood $\mathbb{T}$. Then with probability $1 - O(N^{-\mathfrak{q}})$, one has that for all except $O_{\mathfrak{q}}(1)$ indices α, that $\chi_\alpha = 1$. Here, $O_{\mathfrak{q}}(1)$ means that this is bounded by a constant depending on the parameter $\mathfrak{q}$. Furthermore, for any*

vertex x in $\mathbb{T}^c$, the set of indices α such that x is of distance less than $\mathfrak{R}/4$ from $\{a_\alpha, b_\alpha, c_\alpha\}$ is less than $O_{\mathfrak{q}}(1)$.

Proof. This is [18, Prop. 5.10]. $\qquad\square$

2.2.3. *The inductive method.* At this point, we will now introduce our preliminary induction hypothesis.

Definition 2.7 (Induction Hypothesis). Let us define $\epsilon = \epsilon(z, w)$ as our control parameter,

$$(2.9) \qquad \epsilon = (\epsilon(z,w)) := \frac{(\log N)^{K\mathfrak{a}}}{N^{\mathfrak{r}}} + \frac{(\log N)^{K\mathfrak{a}}\sqrt{\text{Im}[m_\infty]}}{\sqrt{N\eta}},$$

where K is a large positive constant $K \geq 100$.

We let $\Omega^0(z, w)$ denote the set of graphs inside Ω from Definition 2.3 such that the following estimates hold at the point (z, w).

$$|Q_I + Q_O - 2m_\infty| \leq \frac{\epsilon}{S_g^1},$$

$$(2.10) \qquad |Q_I - Q_O| \leq \frac{\epsilon}{S_g^2},$$

$$|G_{ij} - G_{ij}(Ext(B_r(\{i,j\}, \mathcal{G}), Q_I, Q_O)| \leq \epsilon.$$

where the latter estimate holds for all pairs i and j. The singularity parameters, $S_g^{1,2}$, are formally defined in Definition 3.1.

Now, we make a definition designed to capture the properties of graphs produced after one switching on the neighborhood $\mathbb{T}$ around a vertex o.

Definition 2.8. We define the new error parameter ϵ' ($= \epsilon'(z, w)$) as,

$$(2.11) \qquad \epsilon' := (\log N)^{\mathfrak{j}(d)\mathfrak{a}}\epsilon,$$

where $\mathfrak{j}(d) = \log_{d-1}\left[\frac{1}{2} + \frac{1}{2}\sqrt{\frac{2d-1}{d-1}}\right] < 1/2$.

Let $\mathcal{G}$ be a graph inside $\Omega_{S,\mathfrak{q}}$ from Definition 2.3, and we fix a vertex $\hat{r}$ in $\mathcal{G}$ with radius ℓ neighborhood, $\mathbb{T}$. We will include $\mathcal{G}$ inside the set $\Omega^1_{S,\mathfrak{q},\hat{r}}(z, w)$ if it satisfies the following estimates at the point (z, w),

$$|Q_I + Q_O - 2m_\infty| \lesssim \frac{\epsilon}{S_g^1},$$

$$|Q_I - Q_O| \lesssim \frac{\epsilon}{S_g^2},$$

$$(2.12) \qquad |G_{i+o_i, j+o_j} - G_{i+o_i, j+o_j}(Ext(B_r(\{i,j\}, S(\mathcal{G})), Q_I, Q_O))| \lesssim \frac{1}{\log N}, \forall i, j \in \mathcal{G}$$

$$|G^{(\mathbb{T})}_{i+o_i, j+o_j} - G_{i+o_i, j+o_j}(Ext(B_r(\{i,j\}, S(\mathcal{G}))\backslash\mathbb{T}, Q_I, Q_O))| \lesssim \epsilon', \forall i, j \in \mathbb{T}^c.$$

Here, o_i and o_j are symbols that can take the value 0 or N, so that $i + o_i$ can correspond to one of the two labels corresponding to the vertex i. The constant implicit in $\lesssim$ is uniform and does not depend on N.

We have the following relating $\Omega^0(z, w)$ to $\Omega^1_{S,\mathfrak{q},o}(z, w)$.

Theorem 2.9. *Fix a graph $\mathcal{G}$ in $\Omega^0(z, w)$, as well as a vertex o with radius ℓ neighborhood $\mathbb{T}$. In addition, fix the parameter $\mathfrak{q} > 0$. There is an event $S_G(\mathcal{G})$ of switching events S in $\mathbb{S}_{\mathcal{G},o}$ with probability $\mathbb{P}(S_G(\mathcal{G})) = 1 - O(N^{-\mathfrak{q}})$, such that for every switching event $S \in S_G(\mathcal{G})$, we have that $\tilde{\mathcal{G}}$ is in $\Omega^1_{S,\mathfrak{q},o}(z, w)$.*

The issue with the above statement is that it does not take into account improved concentration estimates from the switching. Our issue is that we cannot ensure that after a single switching, the event $\Omega^0(z, w)\backslash T(\Omega^0(z, w))$ is of vanishingly small probability. In order to avoid this issue, we use the fact that T is an involution and two evaluations of T on the set $\Omega^0(z, w)$ will result in a graph that belongs to $\Omega^0(z, w)$.

To this end, we introduce a new definition that describes the graphs that we would obtain after two switching.

Definition 2.10. Let $\mathcal{G}$ be d-regular digraph with N vertices, and let $\hat{r}$ be a vertex in $\mathcal{G}$ with radius ℓ neighborhood $\mathbb{T}$. For all vertices i in $\mathbb{T}$, we define ℓ_i to be the distance from $\hat{r}$ to i. We set that the graph $\mathcal{G}$ belongs to the set $\Omega^2_{S,\mathfrak{q},\hat{r}}(z,w)$ if it satisfies the following conditions at the point (z,w):

$$
|Q_I + Q_O - 2m_\infty| \lesssim \frac{\epsilon}{S^1_g},
$$

(2.13)

$$
|Q_I - Q_O| \lesssim \frac{\epsilon}{S^2_g}.
$$

(2.14)
$$
|G_{\hat{r},i} - G_{\hat{r},i}(Ext(B_r(\{\hat{r},i\},\mathcal{G}),Q_I,Q_O))| \lesssim \left(\frac{\log N}{(d-1)^{\ell/2}} + \frac{(\log N)^2}{(d-1)^{\ell-\ell_i/2}} \right)\epsilon'
$$
$$
+ \log N(|S^1_g||Q_I + Q_O - 2m_\infty| + |S^2_g||Q_I - Q_O| + |Q_I - m_\infty|^2 + |Q_O - m_\infty|^2), \forall i \in \mathbb{T}
$$

(2.15)
$$
|G_{\hat{r},i} - G_{\hat{r},i}(Ext(B_r(\{\hat{r},i\},\mathcal{G}),Q_I,Q_O))| \lesssim \frac{(\log N)^2 \epsilon'}{(d-1)^{\ell/2}}
$$
$$
+ \log N(|S^1_g||Q_I + Q_O - 2m_\infty| + |S^2_g||Q_I - Q_O| + |Q_I - m_\infty|^2 + |Q_O - m_\infty|^2), \forall i \in \mathbb{T}^c
$$

Furthermore, if the vertex $\hat{r}$ has a radius $\mathfrak{R}$ tree-like neighborhood in $\mathcal{G}$, then we also have,

(2.16)
$$
\left| \frac{1}{d} \sum_{(i \to \hat{r})} G^{(i)}_{\hat{r}+N,\hat{r}+N} - Y_{i,\ell}(Q_I,Q_O) \right| \lesssim \frac{\log N \epsilon'}{(d-1)^{\ell/2}},
$$
$$
\left| \frac{1}{d} \sum_{(\hat{r} \to i)} G^{(i)}_{\hat{r},\hat{r}} - Y_{o,\ell}(Q_I,Q_O) \right| \lesssim \frac{\log N \epsilon'}{(d-1)^{\ell/2}}.
$$

As always, $\lesssim$ represents a bound by a universal constant that does not depend on N.

The improvement of the last estimate is found in the decay factors $\frac{1}{(d-1)^{\ell/2}}$ found in the denominator. Our main relationship between $\Omega^1_{S,\mathfrak{q},o}(z,w)$ and $\Omega^2_{S,\mathfrak{q},o}(z,w)$ is the following theorem.

Theorem 2.11. *Fix a graph $\mathcal{G}$ in $\Omega^1_{S,\mathfrak{q},o}(z,w)$. This graph comes with a specially denoted vertex o and neighborhood $\mathbb{T}$. There is a collection of switches $S_C(\mathcal{G}) \in \mathbb{S}_{\mathcal{G},o}$ with probability $\mathbb{P}(S_C(\mathcal{G})) = 1 - O(N^{-\mathfrak{q}})$ such that the image of $(\mathcal{G},s)$ under T for $s \in S_C(\mathcal{G})$ is in $\Omega^2_{S,\mathfrak{q},o}(z,w)$.*

2.3. **A Multiscale Induction Scheme.** Our central Theorem 1.14 is proven via a multiscale induction scheme.Namely, we start by proving the estimates in equation (1.39). We start by establishing these estimates at some z_0 where $|z_0|$ is large and one can derive the inequalities by basic computations with expansions. Define $z_k = z_0 - kN^{-4}$. Now, if we establish the equations (1.39) at some scale z_k, the goal is to prove these inequalities at z_{k+1}. On a high level, the key input is Theorem 2.11; this shows that, with very high probability, a switched graph would satisfy stronger Green's function estimates at the scale z_k. Afterward, one can use basic Lipschitz continuity estimates to take these improved Greens function estimates at z_k to obtain the inductive hypothesis at z_{k+1}. We can more precisely quantify the improvement from switching int he following definition.

Definition 2.12. We let $\Omega_g(z,w) \subset \Omega$ be the set of d-regular digraphs $\mathcal{G}$ that satisfy the following estimates at the point (z,w),

$$
|Q_I + Q_O - m_\infty| \leq \frac{\epsilon}{2S^1_g},
$$

(2.17)
$$
|Q_I - Q_O| \leq \frac{\epsilon}{2S^2_g},
$$
$$
|G_{ij} - G_{ij}(Ext(B_r(\{i,j\},\mathcal{G}),Q_I,Q_O))| \leq \frac{\epsilon}{2}, \qquad \forall i,j.
$$

Our first goal is to show that for $|z|$ large enough, we know that any graph lies in $\Omega_g(z,w)$ deterministically.

Proposition 2.13. *Let w lie in some compact domain $\mathbb{D}$ of $\mathbb{C}$ and choose z with $|z| \geq 2d^2 + \max_{a \in \mathbb{D}} d|a|$. Then, any d-regular digraph $\mathcal{G}$ on N vertices belongs to $\Omega_g(z, w)$.*

Proof. We know that the largest eigenvalue of a Hermitized version H of the adjacency matrix A of a d-regular digraph can be bounded by $\sqrt{d} + |w|$; this can be obtained by taking the maximum sum of the elements in a row or in a column. This shows us that we may expand the Green's function of our Hermitized matrix as,

$$(2.18) \qquad (H - z)^{-1} = -z^{-1} - \sum_{k=1}^{\infty} H^k z^{-(k+1)}.$$

Now, one can see that the computation of H_{ab}^k involves counting the contribution of all length k paths between a and b in the graph $\mathcal{G}$ with possible self-loops weighted by w or $\overline{w}$. From this logic, we can assert the following statement on graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ that each have vertices o that are identified to each other. If the radius r neighborhood of o in $\mathcal{G}_1$ is isomorphic to the radius r neighborhood of o in $\mathcal{G}_2$, then $\sum_{k=1}^{r} H^k z^{-(k+1)}$ agree with each other in both $\mathcal{G}_1$ and $\mathcal{G}_2$.

Thus, we would know that,

$$(2.19) \qquad |G_{ij} - G_{ij}(Ext(B_r(\{i, j\}, \mathcal{G}), m_\infty, m_\infty))| \leq 2 \sum_{k=r+1}^{\infty} \|H\| |z|^{-(k+1)} \leq 4 \left(\frac{1}{d}\right)^r \lesssim \frac{\epsilon}{(\log N)^2}$$

In particular, if i has a radius $\mathfrak{R}$ tree-like neighborhood, then

$$(2.20) \qquad |G_{ii} - m_{\mathcal{T}}^d| \leq 4 \left(\frac{1}{d}\right)^r.$$

Here, we only used the fact that $G_{ii}(Ext(B_r(i, \mathcal{G}), m_\infty, m_\infty))$ will be equal to $m_{\mathcal{T}}^d$.

These perturbation estimates will also give us information on $G_{ii}^{(j)}$ via an application of the following identity,

$$(2.21) \qquad G_{ii}^{(j)} = G_{ii} - G|_{i, \{j, j+N\}} (G|\{j, j+N\})^{-1} G|_{\{j, j+N\}, i},$$

as well as a similar one for the extension $G_{ii}(Ext(B_r(i, \mathcal{G}^{(j)}), m_\infty, m_\infty))$. This will show that,

$$(2.22) \qquad |G_{ii}^{(j)} - m_\infty| \lesssim \frac{\epsilon}{(\log N)^2},$$

for those i with a tree-like neighborhood. Otherwise, we can assert that $G_{ii}^{(j)}$ is $O(1)$ for the $N^{\mathfrak{c}}$ many vertices that do not have a tree-like neighborhood.

This will show that,

$$(2.23) \qquad \left| \frac{1}{Nd} \sum_{(j \to i)} G_{ii}^{(j)} - m_\infty \right| \lesssim N^{\mathfrak{c}-1} + \frac{\epsilon}{(\log N)^2} \lesssim \frac{\epsilon}{(\log N)^2}.$$

We have a similar statement for Q_O.

Finally, we can use Lemma 3.5 to replace

$$G_{ij}(Ext(B_r(\{i, j\}, \mathcal{G}), m_\infty, m_\infty))$$

with

$$G_{ij}(Ext(B_r(\{i, j\}, \mathcal{G}), Q_I, Q_O)).$$

$\square$

Our next Proposition discusses how the switching procedure will allow us to take an element from $\Omega^0(z, w)$ from Definition 2.7 and return an element of the improved set $\Omega_g(z, w)$ after two random switching.

Proposition 2.14. *Fix w in a compact domain $\mathbb{D} \in \mathbb{C}$. There exists some $\varepsilon > 0$ such that for any z with $Im[z] \geq N^{-\varepsilon}$ and $d \geq 3$, we have that,*

$$(2.24) \qquad \mathbb{P}(\Omega^0(z, w) \backslash \Omega_g^0(z, w)) = O(N^{-\mathfrak{q}+1}).$$

Proof. The proof will be divided into two steps:

(1) We will show for each vertex o that $\mathbb{P}(\Omega^0(z,w)\backslash\Omega^2_{S,\mathfrak{q},o}(z,w)) = O(N^{-\mathfrak{q}})$. $\Omega^2_{S,\mathfrak{q},o}$ is the set from Definition 2.10. A union bound will then show that
$$\mathbb{P}(\Omega^0(z,w)\backslash\bigcap_{o\in\mathcal{G}}\Omega^2_{S,\mathfrak{q},o}(z,w)) \lesssim N^{1-\mathfrak{q}}.$$

(2) After this, we show that a graph that is in $\Omega^0(z,w)\bigcap_{o\in\mathcal{G}}\Omega^2_{S,\mathfrak{q},o}(z,w)$ is in $\Omega_g(z,w)$ via some self-consistent estimates.

Proof of Step 1

Just to distinguish notation, we will let $\mathbb{P}$ denote the probability distribution over graphs, while we let $\tilde{\mathbb{P}}$ denote the joint probability distribution over graphs and switching data. Furthermore, for a set of graphs $\mathbb{G}$, we will use $\tilde{\mathbb{G}}$ to denote the set of the form $(\mathcal{G}, S)$ with $\mathcal{G}\in\mathbb{G}$ and S can be any switching data in $\mathbb{S}_{\mathcal{G},o}$. (We assume that each graph has a specially marked vertex o).

Theorem 2.9 shows that,
$$\tilde{\mathbb{P}}(T(\tilde{\Omega}^0(z,w))\backslash\tilde{\Omega}^1_{S,\mathfrak{q},o}(z,w)) \lesssim N^{-\mathfrak{q}}.$$

Since T both preserves measure and is an involution, this will also show that

(2.25)
$$\tilde{\mathbb{P}}(\tilde{\Omega}^0(z,w)\backslash T(\tilde{\Omega}^1_{S,\mathfrak{q},o}(z,w)) \lesssim N^{-\mathfrak{q}}$$

Theorem 2.11 shows that,

(2.26)
$$\tilde{\mathbb{P}}(T(\tilde{\Omega}^1_{S,\mathfrak{q},o})\backslash\tilde{\Omega}^2_{S,\mathfrak{q},o}(z,w)) \lesssim N^{-\mathfrak{q}}.$$

Combining these two estimates will give us,

(2.27)
$$\tilde{\mathbb{P}}(\tilde{\Omega}^0(z,w)\backslash\tilde{\Omega}^2_{S,\mathfrak{q},o}(z,w)) \lesssim N^{-\mathfrak{q}}.$$

A simple projection shows that this will imply that $\mathbb{P}(\Omega^0(z,w)\backslash\Omega^2_{S,\mathfrak{q},o}(z,w)) \lesssim N^{-\mathfrak{q}}$.

Proof of Step 2

We first form a self-consistent equation to understand the values of Q_I and Q_O. If we sum up the estimates in equation (2.16), we see that we have that,

(2.28)
$$\left|\frac{1}{Nd}\sum_{(i\to j)\in\mathcal{G}}G^{(i)}_{j+N,j+N} - Y_{i,\ell}(Q_I,Q_O)\right| \lesssim \frac{\log N\epsilon'}{(d-1)^{\ell/2}},$$
$$\left|\frac{1}{Nd}\sum_{(j\to i)\in\mathcal{G}}G^{(i)}_{j,j} - Y_{i,\ell}(Q_I,Q_O)\right| \lesssim \frac{\log N\epsilon'}{(d-1)^{\ell/2}}.$$

This is the self-consistent equation for Q_I and Q_O that has been analyzed in Theorem 3.8.

Now, the estimates on G_{oi} from (2.14) and (2.15) are good enough for the Green's functions estimates in Theorem 3.8. The main benefit is the presence of the $(d-1)^{\ell/2} = (\log N)^{\mathfrak{a}/2}$ factor in the denominator, which is sufficient to cancel the $(\log N)$ factor in the numerator as well as the factor $(\log N)^{\mathfrak{j}(d)\mathfrak{a}}$ factor in the definition of ϵ' when $d \geq 3$.

This shows that,

(2.29)
$$|Q_I + Q_O - 2m_\infty| \lesssim \frac{\epsilon'}{S^1_g(\log N)^k},$$
$$|Q_I - Q_O| \lesssim \frac{\epsilon'}{S^2_g(\log N)^k}.$$

where k can be arbitrarily large if $\mathfrak{a}$ is chosen large enough.

Now, this bound is good enough to use in the other Green's function estimates found in equation (2.14) and (2.15). As we have mentioned before, the term $\frac{\log N\epsilon'}{(d-1)^{\ell/2}}\epsilon'$ is $\lesssim \frac{\epsilon}{\log N}$ due to the presence of the factor $(d-1)^{\ell/2}$ in the denominator. Now, our improved bounds on $|(Q_I - m_\infty)\pm(Q_O - m_\infty)|$ can be substituted into the right hand sides of these expressions in order to deal with the remaining terms on the right hand sides of equations (2.14) and (2.15). This completes the proof of the proposition. $\qquad\square$

At this point, one is left with preforming a standard multi-scale analysis with continuity estimates.

Proof of Theorem 1.14. Consider the domain

$$\{(w, z) : w \in \mathbb{D}, \operatorname{Im}[z] \geq N^{-\epsilon}, |z| \leq d^2 + d \max_{y \in \mathbb{D}} |y| + 1\}.$$

Now consider a dense discrete grid $\mathbb{L}$ inside this domain where the spacing between elements is $\frac{1}{N^5}$. By basic derivative estimates of the trace of a Green's function, we know that if $|z - z'| \leq \frac{1}{N^5}, |w - w'| \leq \frac{1}{N^5}$. Then, $\left|\frac{1}{N}\operatorname{Tr}[G(z, w)] - \frac{1}{N}\operatorname{Tr}[G(z', w')]\right| \lesssim \frac{1}{N}$, deterministically. Thus, it suffices to prove the local law with high probability on the elements of the discrete grid under question.

For fixed value of $\operatorname{Re}[z]$ and w, we can order the values of z in the grid with these values of $\operatorname{Re}[z]$ and this w in order of decreasing imaginary part as $\operatorname{Im}[z_1] \geq \operatorname{Im}[z_2] \geq \operatorname{Im}[z_3] \geq \ldots \geq \operatorname{Im}[z_M]$ with $M = O(N^5)$.

Now, we know that $\mathbb{P}(\Omega_g(z_1, w)) = 1$ due to Proposition 2.13. Out continuity bound shows that $\Omega_g(z_k, w) \subset \Omega^0(z_{k+1}, w)$. Now, Proposition 2.14 shows that $\mathbb{P}(\Omega^0(z_{k+1}, w)\backslash\Omega_g(z_{k+1}, w)) \lesssim N^{-\mathfrak{q}}$. Thus, we know that $\mathbb{P}(\Omega_g(z_{k+1}, w) \geq \mathbb{P}(\Omega_g(z_k, w)) - O(N^{1-\mathfrak{q}})$. This is iterated up to at most N^5 steps so $\mathbb{P}(\Omega_g(z_M, w)) \geq 1 - O(N^{16-\mathfrak{q}})$. We can now take a union bound over our dense grid. Choosing $\mathfrak{q}$ sufficiently large will show that the intersection $\mathbb{P}(\bigcap_{(z,w)\in\mathbb{L}} \Omega_g(z, w)) = 1 - o(1)$ as desired. $\square$

2.4. **The strategy of proving Theorems 2.9 and 2.11.** Now that we have introduced the notation necessary to discuss the switching arguments and the proofs of the Green's function estimates, we will begin to give an overview of the series of estimates that are required in order to derive the theorems that were previously mentioned in this section. However, before we detail the strategy, we will discuss on a high level the differences that arise when considering the Green's functions of the adjacency matrices of directed graphs as compared to undirected graphs. First of all, since we now have to deal with edges with orientations, this means that when performing a switching, we have to preserve the orientation; furthermore, since we now have two distinct classes of edges, we get two different order parameters Q_I and Q_O corresponding to the different orientations of the edges. Furthermore, it will be observed in the proof that many estimates of Green's function involve understanding the values of Green's functions for the infinite d-regular tree and obtaining appropriate l^2 and l^1 controls for the sum of Green's function for the infinite tree. In the undirected case, these Green's function values were a simple function of the distance. However, when considering the directed graph, the specific path and the orientation of edges along the path will drastically change the value of the Green's function between two vertices; keeping track of these differences occurring from the path structure is the main technical challenge that occurs in the directed case.

The proofs of these theorems will come a series of estimates. To simplify the organization, the proofs are divided into multiple parts. We will briefly describe the structure here. In the statements that follow, if we let i be a vertex, then it could represent either i or $i+\aleph$ when it appears as a sub-index of a Green's function value.

Steps to Prove Theorem 2.9

(1) Our first step is to relate estimates of G to $G^{(\mathbb{T})}$. This involves applying the Schur complement formula; a major ingredient in intermediate estimates are the Green's function estimates on tree-like neighborhoods from Section 3. This is the goal of Section 4.

The main result is the following:

Proposition 2.15. *Consider a graph $\mathcal{G}$ in $\Omega^0(z, w)$. Fix a vertex o and let $\mathbb{T}$ be the radius ℓ neighborhood around $\mathbb{T}$. Assume, in addition, that o has a radius $\mathfrak{R}$ neighborhood with excess at most 1. Then, we have the following estimates on the Green's function of the graph $\mathcal{G}$ with the neighborhood $\mathbb{T}$ removed. Let i and j be vertices outside of $\mathbb{T}$ in $\mathcal{G}$ and let P denote $G(Ext(B_r(\mathbb{T} \cup \{i, j\}, \mathcal{G}), Q_I, Q_O))$. Then, we have that,*

(2.30) $$|G_{ij}^{(\mathbb{T})} - P_{ij}^{(\mathbb{T})}| \leq \epsilon'.$$

(2) Our second step is to then remove the vertices W_S that take part in the switching. Again, this involves applying the Schur complement formula. However, this time we need to introduce a notion called Green's function distance. This uses the fact that the average case bound on the Green's

functions $|G_{ij}|$ is usually better than the worst case bound; Green's function distance quantifies information on when we can apply the average case bound rather than the worst case bound. This is the goal of Section 5. Though the specific value of ϕ and other assorted definitions will be defined in the section, we state the Proposition we will prove here for convenience.

Proposition 2.16. *Consider a graph $\mathcal{G}$ that belongs to the setting of Proposition 2.15. Thus, we consider graphs that satisfy the important conclusion (2.30). Given a switching event S involving exchanging the edges on the boundary of $\mathbb{T}$ with arbitrary edges given by $\{b_1, c_1\} \ldots \{b_\mu, c_\mu\}$ as in Definition 2.5, then W_S is the set of those b_α with $\chi_\alpha = 1$. Then, there is an event of switching, $S_G(\mathcal{G})$ that holds with probability $1 - O(N^{-q})$ such that the following estimates hold uniformly.*

Let i and j be vertices in of $\mathcal{G}\backslash(\mathbb{T} \cup W_S)$ and let P be a shorthand for $G(Ext(B_r(\{i,j\}) \cup W_S, \mathcal{G}\backslash\mathbb{T}, Q_I, Q_O))$.

$$|\tilde{G}_{ij}^{(\mathbb{T}\cup W_S)} - P_{ij}| \lesssim \phi. \tag{2.31}$$

Otherwise, we have the worst case bound,

$$|\tilde{G}_{ij}^{(\mathbb{T}\cup W_S)} - P_{ij}| \lesssim \epsilon'. \tag{2.32}$$

(3) Our third step is to reintroduce the vertices in W_S, but with adjacency matrix given by the switching procedure. This section will give us most of the estimates on $\tilde{G}^{(\mathbb{T})}$ for the switched graph; namely, we derive the third inequality of equation (2.12). This is the goal of Section 6. Our formal statement is as follows:

Proposition 2.17. *Consider the setting of the previous Proposition 2.16. Thus, we are considering a graph $\mathcal{G}$ along with a switching event $S \in S_G(\mathcal{G})$ such that we satisfy both of the estimates (2.31) and (2.32). Construct the switched graph $\tilde{\mathcal{G}}$ as in Definition 2.5. Let $\tilde{G}$ denote the Green's function of the switched graph $\tilde{G}$. Let i and j be arbitrary vertices in $\tilde{\mathcal{G}}\backslash\mathbb{T}$ and let P be a shorthand for $G(Ext(B_r(\{i,j\} \cup (\mathbb{T} \cup W_S), \tilde{\mathcal{G}}\backslash T), Q_I, Q_O)$.*

Then, we have the following estimates: If i and j are not Green's function connected and either both i and j belong to W_S or both are outside W_S, then we have the bound,

$$|\tilde{G}_{ij}^{(\mathbb{T})} - P_{ij}| \lesssim \phi. \tag{2.33}$$

Recall our notation for $\{b_\alpha, c_\alpha\}$ for the set of vertices that participate in the switching. In the specific case where we let i be of the form b_β and j be of the form c_α with $\alpha \neq \beta$, then we also have the estimate,

$$|\tilde{G}_{b_\beta c_\alpha}^{(\mathbb{T}\cup W_S)} - P_{b_\beta c_\alpha}| \lesssim \phi + (\epsilon')^2. \tag{2.34}$$

Otherwise, we have the worst case bound,

$$|\tilde{G}_{ij}^{(\mathbb{T})} - P_{ij}| \lesssim \epsilon'. \tag{2.35}$$

(4) Our fourth step is to finally re-introduce the vertices of $\mathbb{T}$ to the switched graph. This is the goal of Section 7. Formally, our main result is the following,

Proposition 2.18. *Consider the setting of the previous Proposition 2.17; especially the important estimates (2.33), (2.34), and (2.35). When we add back the neighborhood $\mathbb{T}$ to the switched graph $\tilde{G}$, we will get the following Green's function estimates. Let i and j be arbitrary vertices in the graph $\tilde{G}$ and let P be a shorthand for $G(Ext(B_r(\{i,j\} \cup \mathbb{T}, \tilde{\mathcal{G}}), \tilde{Q}_I, \tilde{Q}_O))$. Here, $\tilde{Q}_I$ and $\tilde{Q}_O$ are the corresponding values of Q_I and Q_O corresponding to the switched graph $\tilde{G}$.*

If i and j both belong to the neighborhood $\mathbb{T}$, then we have,

$$|\tilde{G}_{ij} - P_{ij}| \lesssim \epsilon'(\log N)^2 \left(\frac{1}{\sqrt{d-1}}\right)^{2\ell - d_i - d_j} + (1 + |d_j + d_i|)\left(\frac{1}{d-1}\right)^{\frac{|d_j - d_i|}{2}} \epsilon'. \tag{2.36}$$

If instead one of i belongs to $\mathbb{T}$ and the other belongs to $\mathbb{T}^c$, then we have the estimates,

$$(2.37) \qquad |\tilde{G}_{ij} - P_{ij}| \lesssim (logN)^2 \epsilon' \left(\frac{1}{\sqrt{d-1}} \right)^{\ell - d_i}.$$

In addition, we also have,

$$(2.38) \qquad |Q_I + Q_O - 2m_\infty| \lesssim \frac{\epsilon}{S_g^1}, \quad |Q_I - Q_O| \lesssim \frac{\epsilon}{S_g^2}.$$

Given this final statement, Theorem 2.9 is a collection of estimates proved already in the previous propositions.

Proof of Theorem 2.9. Consider a graph $\mathcal{G}$ in $\Omega^0(z, w)$. To this graph, we can apply the conclusion of Proposition 2.18; from this we have a set of good switching events $S_G(\mathcal{G})$. On these good switching events, we can derive equations (2.36) and (2.37). This corresponds to the middle equation of (2.12). Additionally, (2.38) correspond to the first equation of (2.12). Finally, equation (2.30) is the final equation of (2.12). $\square$

Steps to Prove Theorem 2.11

Most of the same estimates used to prove Theorem 2.9 can be used to prove many of the estimates necessary to prove Theorem 2.11. The only new ingredient are new concentration estimates, that involve the fact that the switch used random edges. This step will be discussed in Section 8.

3. Perturbation theory for almost tree-like neighborhoods: The proof of Theorem 1.13

The basic underpinning of our strategy is the ability to deal with the Green's function of graphs with almost tree-like neighborhoods via perturbations of the one for the infinite tree (and when the neighborhood is not tree-like, we use the relation between the Green's function of a graph and the Green's function of the covering graph). In this section we detail the relevant estimates one needs for this purpose.

Hereafter, for digraphs of a generic size we assign one standard index per vertex, denoted by lowercase letters, such as i. Pairing each such index i with a second index, denoted $i + \aleph$, any forward edge $(i \to j)$ of the digraph yields two non-zero entries $H_{i,j+\aleph}$ and $H_{j+\aleph,i}$ of the matrix $H(z, w)$. As we did in case of trees, the entry $H_{i,j+\aleph}$ is interpreted as traversing a forward edge, and $H_{j+\aleph,i}$ as traversing a backward edge.

3.1. General perturbation bounds.

The goal of this section if to prove general perturbations bounds for Green's functions of almost tree-like neighborhoods. Theorem 1.13 is an example of a result of this type and is used directly in the proof of our main result. However, there are times where we need to use more sophisticated estimates and, thus, this subsection will also contain more delicate estimates for these cases. Before we actually go on and prove Theorem 1.13, we will collect some necessary preliminary results; we remark here that many of these useful combinatorial results also appear in the Appendix. We start with some useful definitions and lemmas; our first definition will discuss the nature of singularities that appear in the self-consistent equations.

Definition 3.1 (Singularities of the self-consistency equation). Singularities that may occur in computations with our (distinguished) solution m_∞ of (1.30), are parametrized by

$$(3.1) \qquad S_g^1 := 1 - X - Y \leq 1 + X - Y := S_g^2,$$

for X, Y of (A.1).

Lemma 3.2. *Suppose a tree-like digraph $\mathcal{G}$ with root $\hat{r}$ is such that*

$$(1 + diam(\mathcal{G}))^3 [|Q_+| + |Q_-|] \ll 1,$$

where $Q_+ := \frac{Q_I + Q_O - 2m_\infty}{\sqrt{2}}$ and $Q_- := \frac{Q_I - Q_O}{\sqrt{2}}$. then we have in terms of S_g^1, S_g^2 and $Ext_i(\mathcal{G}, Q_I, Q_O)$ of Definition 1.11, that

$$(3.2) \qquad \begin{aligned} |G_{\hat{r}+\aleph, \hat{r}+\aleph}(Ext_i(\mathcal{G}, Q_I, Q_O)) - Q_I| \lesssim \\ (1 + diam(\mathcal{G}))^2 \left(S_g^1 |Q_+| + S_g^2 |Q_-| \right) + (1 + diam(\mathcal{G}))^3 [|Q_+|^2 + |Q_-|^2]. \end{aligned}$$

One can obtain a similar expression comparing $G_{\hat{r},\hat{r}}(Ext_o(\mathcal{G}, Q_I, Q_O))$ *and* Q_O.

Remark 3.3. We include the factors $1 \pm X - Y$ in order to match the singularity factors we expect in the stability equation for $\tilde{Q}_1 - \sqrt{2}m_\infty$ and $\tilde{Q}_2$.

In order to prove this lemma, we need to introduce the following definition that aids in counting.

Definition 3.4 (Basal Tree). Consider the d-regular infinite digraph tree $\mathcal{T}$ with root $\hat{r}$. For k a positive integer, let $v_1, \ldots, v_k$ be vertices on this infinite tree. If we let P be the Green's function corresponding to this infinite tree, consider the product $|P_{\hat{r}v_1}||P_{v_1v_2}||P_{v_2v_3}|\ldots|P_{v_n\hat{r}}|$. This product can be associated with what we call a basal tree, T_k, that reflects the general structure of this product and is convenient for summation.

If $k = 1$, the basal tree, T_1, consists of a single edge between two vertices. The interpretation of this is as follows. The two vertices of the edge should correspond to the vertices r and v_1, while the path between r and v_1 corresponds to the edge between r and v_1. In general, our construction will correspond vertices of the basal tree to some vertices in $\mathcal{T}$ and edges of the basal tree to paths in $\mathcal{T}$.

For $k > 1$, we construct the basal tree inductively. Let T_{k-1} denote the basal tree corresponding to the product $|P_{\hat{r}v_1}||P_{v_1v_2}||P_{v_2v_3}|\ldots|P_{v_{n-1}\hat{r}}|$. Consider the subgraph in $\mathcal{T}$ consisting of the union of paths between v_i to v_{i+1} for $i = 1, \ldots, n-2$ as well as the paths between $\hat{r}$ and v_1 and v_{n-1}; call this subgraph S. Now consider the path from r to v_n. This path has some final ancestral vertex A_n, which is the last vertex in common between the path between r and v_n and S.

In the correspondence between the edges of T_{k-1} and paths in $\mathcal{T}$, the vertex A_n must like on one of these paths; call this path p and the edge corresponding to it in T_{k-1} e. To construct T_k from T_{k-1}, add a vertex in the middle of the edge e; this vertex in T_n corresponds to A_n. Finally, add a new leaf vertex at this new vertex. The other endpoint of this leaf corresponds to the vertex v_k while the leaf edge corresponds to the path between A_n and v_n. This completes the inductive construction of the tree.

Proof. As in the proof of the stability equation, we consider the difference between $G := G_{\hat{r}+\aleph,\hat{r}+\aleph}(Ext_i(\mathcal{G}, Q_I, Q_O))$ and $P := G_{\hat{r}+\aleph,\hat{r}+\aleph}(Ext_i(\mathcal{G}, m_\infty, m_\infty))$. The function P gives the Green's function entries of the infinite tree $\mathcal{T}_1$.

By applying the resolvent identity, we have,

$$(3.3) \qquad G - P = \sum_{k=1}^{\infty} P[(P^{-1} - G^{-1})P]^k.$$

Notice that the terms of $P^{-1} - G^{-1}$ are of the form $Q_I - m_\infty$ or $Q_O - m_\infty$. Thus, we see that higher order terms appearing in $P[(P^{-1} - G^{-1})P]^k$ will give us errors of order $O((Q. - m_\infty)^k)$. To make this rigorous, we need to determine the coefficient.

Part 1: Controlling the contribution of the higher order parameters

Observe that the coefficients of $P[(P^{-1} - G^{-1})P]^k$ can be bounded in absolute value by,

$$(3.4) \qquad \sum_{v_1,\ldots,v_k} |P_{\hat{r}v_1}||P_{v_1v_2}||P_{v_2v_3}|\ldots|P_{v_k\hat{r}}|,$$

where the points v_i are on the boundary of the graph G. This is too difficult to sum in each variable v_k independently, so we introduce a new decomposition.

Step 1: An alternative decomposition

If we are considering the product $P_{\hat{r}v_1}P_{v_1v_2}\ldots P_{v_kr}$, we can naturally associate this product to the subgraph of the tree consisting of all the unions of paths between v_i to v_{i+1}. Since each term $P_{v_iv_{i+1}}$ is a function of the edges of the path between them, knowing the form of this subgraph is sufficient to bound the value of the desired product. We remark that this is the main point of Lemma A.3. For example, we see that we can bound,

$$|P_{\hat{r}v_1}| \leq \prod_{i=1}^{k_1-1} K(\{e_i^1, e_{i+1}^1\}).$$

Notice that, depending on the specific orientation of the edges, this path will have varying factors of X and Y. Naively applying a worst case bound by using $X, Y \leq \max(X, Y)$ will lead to a bound of

$\sum_{v_1,\ldots,v_k} |P_{\hat{r}v_1}||P_{v_1v_2}||P_{v_2v_3}|\ldots|P_{v_k\hat{r}}|$ that is useless. Instead, one must use the condition $X + Y \leq 1$ by grouping together terms that have similar structures, but roughly exchange the roles of X and Y. This grouping is formally encoded by finding all terms that map to the same basal tree structure.

To make this procedure clearer, we will first consider the case where we only have to deal with two vertices v_1 and v_2. The construction used here is the linchpin of our method. Trees with more edges are constructed inductively.

Two Vertex Case: We let A_1 be the last common ancestor of v_1 and v_2. Pictorially, to arrive at v_2 from v_1, we first proceed along the path p^1 to A_1. From the point A_1 we start branching out to the point v_2 on the boundary; we let $p^2 = e_1^2, e_2^2, \ldots, e_{k_2}^2$ be the path from A_1 to v_2. If the path from A_1 to v_1 is $e_{b_1}^1, e_{b_1+1}^1, \ldots, e_{k_1}^1$, then we see that the value of $|P_{v_1v_2}|$ can be bounded as follows:

$$|P_{v_1v_2}| \leq C \prod_{i=b_1}^{k_1-1} K_{e_i^1, e_{i+1}^1} \prod_{j=1}^{k_2-1} K_{e_j^2, e_{j+1}^2}$$

.

The path back from v_2 to $\hat{r}$ will first traverse v_2 to A_1 and then A_1 to $\hat{r}$. We can visualize this as a star with three edges. This star is the tree associated with our subgraph. A_1 is the center of this star and it is connected to v_1, v_2 and $\hat{r}$. If we let $K_{A_1, \hat{r}/v_i}$ be a shorthand for the product of K factors between A_1 and the respective vertex in question, then we have that,

(3.5) $$|P_{\hat{r}v_1}| \leq C K_{A_1\hat{r}} K_{A_1 v_1}, \quad |P_{v_1v_2}| \leq C K_{A_1 v_1} K_{A_1 v_2}, \quad |P_{rv_2}| \leq C K_{A_1 v_1} K_{A_1 v_2}.$$

Thus, we would have that,

(3.6) $$\sum_{v_1,v_2} |P_{\hat{r}v_1}||P_{v_1v_2}||P_{v_2\hat{r}}| \leq \sum_{A_1,v_1,v_2} C^3 K_{A_1\hat{r}}^2 K_{A_1 v_1}^2 K_{A_1 v_2}^2.$$

Step 2: The summation along the tree

The purpose of the tree structure is to introduce the new summation variables $A_1 \ldots A_{k-1}$. This also introduces a new way to sum along the edges. As we mentioned before, we use Lemma A.3 the fact that the products $|P_{v_i v_{i+1}}|$ decomposes as a product of K factors along the edges of the path connecting v_i and v_{i+1}.

Assume that the path from v_i to v_{i+1} in the tree T constructed earlier involves edges $\mathfrak{e}_1, \ldots, \mathfrak{e}_k$ of the tree T. Recall that any edge $\mathfrak{e}$ of T corresponds to a path $p = (e_1, e_2, e_3, \ldots, e_m)$ of edges on our original graph. Given an edge $\mathfrak{e}$ of T that corresponds to a path $p = (e_1, e_2, e_3, \ldots, e_m)$ in our original graph, we let $K_{\mathfrak{e}} = \prod_{i=1}^{m-1} K_{e_i, e_{i+1}}$. We thus have that,

$$|P_{v_i, v_{i+1}}| \leq C^k \prod_{j=1}^{k} P_{\mathfrak{e}_j},$$

$$\prod_{i=0}^{k} |P_{v_i v_{i+1}}| \leq C^{4k} \prod_{\mathfrak{e} \in T} K_{\mathfrak{e}}^2,$$

$$\sum_{v_1, v_2, \ldots, v_k} \prod_{i=0}^{k} |P_{v_i v_{i+1}}| \leq \sum_{\mathcal{T}} \sum_{T \in \mathcal{T}} C^{4k} \prod_{\mathfrak{e} \in T} K_{\mathfrak{e}}^2$$

where v_0 and v_{k+1} are both the root $\hat{r}$. The constant C^{2k} comes since we get a factor of at most C each time we introduce a factor of the form $K_{\mathfrak{e}}$ for each edge $\mathfrak{e} \in \mathbb{R}$. Each edge $\mathfrak{e}$ will appear at most two times. In the last line $\mathcal{T}$ represents all possible tree structures that could occur, while $T \in \mathcal{T}$ is a sum over all actual trees that have said tree structure.

Fix a tree structure $\mathcal{T}$; this will have a leaf edge l, and we can represent $\mathcal{T} = \mathcal{T}' \cup l$ where $\mathcal{T}'$ is the tree structure with the leaf edge removed. Then, each tree T in $\mathcal{T}$ can be decomposed as the union of a tree T' of the form $\mathcal{T}'$ and some path p that will project down to the edge l.

(3.7) $$\sum_{T \in \mathcal{T}} \prod_{\mathfrak{e} \in T} K_{\mathfrak{e}}^2 = \sum_{T' \in \mathcal{T}'} \prod_{\mathfrak{e} \in T'} K_{\mathfrak{e}}^2 \sum_{p \to l} K_l^2 \leq (1 + \text{diam}(\mathcal{G})) \sum_{T' \in \mathcal{T}'} \prod_{\mathfrak{e} \in T'} K_{\mathfrak{e}}^2$$

The last line sums up K_l^2 over all paths p that could possibly project down to l. For this quantity, the bound $(1 + \mathrm{diam}(\mathcal{G}))$ is an application of Lemma A.5. By iteration, since there are $2k - 1$ leaves, one can see that upon fixing the tree, one can derive a bound for the tree sum that is $\lesssim (\mathrm{diam}(\mathcal{G}) + 1)^{2k-1}$. All that is left is to compute the sum over all possible tree structures $\mathcal{T}$.

Since, we know that the distance between any two vertices v_i and v_{i+1} is less than $\mathrm{diam}(\mathcal{G})$; we use the very generous bound that there are no more than $1 + \mathrm{diam}(\mathcal{G})$ ways to chose the position of A_i once the structure of the vertices from $\{v_0, \ldots, v_i\} \cup \{A_1, \ldots, A_{i-1}\}$ has been decided. This shows that there are no more than $(1 + \mathrm{diam}(\mathcal{G}))^{k-1}$ possible tree structures to consider. In total, this bounds the sum appearing in (3.4) by $\lesssim (1 + \mathrm{diam}(\mathcal{G}))^{3k}$.

Part 2: Deriving the correct coefficient structure of the first order term

We have more exact expressions for the coefficient of the first order terms.

$$(3.8) \qquad \sum_{v \in \delta(G)} \frac{d - d_{\mathrm{out}}(v)}{d - 1} P_{\hat{r},v} P_{v,\hat{r}} (Q_I - m_\infty) + \sum_{v \in \delta(G)} \frac{d - d_{\mathrm{in}}(v)}{d - 1} P_{\hat{r},v+\aleph} P_{v+\aleph,\hat{r}} (Q_O - m_\infty).$$

Here, $\delta(G)$ consists of the vertices that lie on the boundary of our graph G.

We can imbed our original graph G into a tree T_{diam} of depth $\mathrm{diam}(G)$. For each vertex v, we let S_v^{out} denote the vertices of T_{diam} that are children of v from an out-edge from v, while S_v^{in} denotes the vertices of T_{diam} that are children of v from an in-edge of v.

An alternative way to write the expression in (3.8) is as,

$$(3.9)$$
$$\sum_{v \in \delta(G)} \Big[\frac{d - d_{\mathrm{out}}(v)}{d - 1} P_{\hat{r},v} P_{v,\hat{r}} [Q_I - m_\infty] - \sum_{w \in S_v^{out}} P_{\hat{r},w} P_{w,\hat{r}} [Q_I - m_\infty] - \sum_{w \in S_v^{out}} P_{\hat{r},w+\aleph} P_{w+\aleph,\hat{r}} [Q_O - m_\infty] \Big]$$
$$+ \sum_{v \in \delta(G)} \Big[\frac{d - d_{\mathrm{in}}(v)}{d - 1} P_{\hat{r},v} P_{v,\hat{r}} [Q_I - m_\infty] - \sum_{w \in S_v^{in}} P_{\hat{r},w} P_{w,\hat{r}} [Q_I - m_\infty] - \sum_{w \in S_v^{in}} P_{\hat{r},w+\aleph} P_{w+\aleph,\hat{r}} [Q_O - m_\infty] \Big]$$
$$+ \sum_{v \in \delta(T_{\mathrm{diam}})} \frac{d - d_{\mathrm{out}}(v)}{d - 1} P_{\hat{r},v} P_{v,\hat{r}} (Q_I - m_\infty) + \frac{d - d_{\mathrm{in}}(v)}{d - 1} P_{\hat{r},v+\aleph} P_{v+\aleph,\hat{r}} (Q_O - m_\infty).$$

The main point is that the last line can be computed exactly via a standard recursion, while the terms in the first two lines should correspond to a smaller order error term with the correct singularity we have to consider. We now treat bounding the last two lines. Indeed, for $w \in S_v^{out}$, we can first try to write $P_{\hat{r}w} P_{w\hat{r}}$ as $P_{\hat{r}v} P_{v\hat{r}} (G_{\mathcal{T}_1})_{vw} (G_{\mathcal{T}_1})_{wv}$. Here, we treat v as the root of a new tree of type $\mathcal{T}_1$ and used the fact that the product $P_{\hat{r}w}$ splits as the product $P_{\hat{r}v}$ times this $(G_{\mathcal{T}_1})_{vw}$. The main benefit now is the fact that we have explicit ways to compute $\sum_{w \in S_v^{out}} (G_{\mathcal{T}_1})_{vw} (G_{\mathcal{T}_1})_{wv}$.

Namely, the matrix $\begin{bmatrix} Y & X \\ X & Y \end{bmatrix}$ satisfies the property that we can write

$$\Big[\frac{d - d_{\mathrm{out}}(v)}{d - 1} P_{\hat{r},v} P_{v,\hat{r}} [Q_I - m_\infty] - \sum_{w \in S_v^{out}} P_{\hat{r},w} P_{w,\hat{r}} [Q_I - m_\infty] - \sum_{w \in S_v^{out}} P_{\hat{r},w+\aleph} P_{w+\aleph,\hat{r}} [Q_O - m_\infty] \Big]$$

as,

$$(3.10) \qquad \frac{d - d_{\mathrm{out}}(v)}{d - 1} P_{\hat{r},v} P_{v,\hat{r}} \begin{bmatrix} 1 & 0 \end{bmatrix} \Big(I - \begin{bmatrix} Y & X \\ X & Y \end{bmatrix}^{\mathrm{diam}(\mathcal{G}) - \mathrm{dist}(v,\hat{r})} \Big) \begin{bmatrix} Q_I - m_\infty \\ Q_O - m_\infty \end{bmatrix}.$$

Now, we can diagonalize the matrix $\begin{bmatrix} Y & X \\ X & Y \end{bmatrix}$ and show that we can rewrite (3.10) as,

$$(3.11) \qquad \frac{d - d_{\mathrm{out}}(v)}{d - 1} P_{\hat{r},v} P_{v,\hat{r}} \Big[\frac{1}{\sqrt{2}} [1 - (X + Y)^k] Q_+ + \frac{1}{\sqrt{2}} [1 - (Y - X)^k] Q_- \Big].$$

In absolute value, this can be bounded by,

$$(3.12) \qquad \frac{d - d_{\text{out}(v)}}{d-1}|P_{\hat{r},v}||P_{v,\hat{r}}|(1 + \text{diam}(\mathcal{G}))\left[|1 - (X+Y)||Q_+| + |1 - (Y-X)||Q_-|\right]$$

Finally, we see that,

$$(3.13) \qquad \begin{aligned} &\sum_v \frac{d - d_{\text{out}(v)}}{d-1}|P_{\hat{r},v}||P_{v,\hat{r}}|(1 + \text{diam}(\mathcal{G}))\left[|1 - (X+Y)||Q_+| + |1 - (Y-X)||Q_-|\right] \\ &\leq O((1 + \text{diam}(\mathcal{G})^2)\left[|1 - (X+Y)||Q_+| + |1 - (Y-X)||Q_-|\right] \end{aligned}$$

We can treat the second line of (3.9) in a similar way.

Combining these error estimates gives us the desired result (3.2). $\qquad \square$

Our next lemma gives us a perturbation analysis of $G_{ij}(Ext(\mathcal{G}, Q_I, Q_O))$ with respect to $G_{ij}(Ext(\mathcal{G}, m_\infty, m_\infty))$.

Lemma 3.5. *Let $\mathcal{G}$ be a graph with excess of at most 1. Assume further that $\text{diam}(\mathcal{G})^3[|Q_I - m_\infty| + |Q_O - m_\infty|] \ll 1$. Then, we have,*

$$(3.14) \qquad \begin{aligned} &|G_{ij}(Ext(\mathcal{G}, Q_I, Q_O)) - G_{ij}(Ext(\mathcal{G}, m_\infty, m_\infty))| \\ &\lesssim \left(\frac{1}{d-1}\right)^{dist(i,j)}(1 + dist(i,j))[|Q_I - m_\infty| + |Q_O - m_\infty|] \end{aligned}$$

Proof. As a shorthand, we let P_{ij} denote the Green's function $G_{ij}(Ext(\mathcal{G}, m_\infty, m_\infty))$, while G_{ij} will denote the Green's function of $G_{ij}(Ext(\mathcal{G}, Q_I, Q_O))$. We can determine the value of $|G_{ij} - P_{ij}|$ via the resolvent formula. Let l_α denote the indices corresponding to vertices on the boundary of the graph $\mathcal{G}$. We have the following,

$$(3.15) \qquad |G_{ij} - P_{ij}| = \sum_k \sum_{l^1, \ldots, l^k} P_{i,l^1} \prod_{m=1}^{k}[Q_{j(l^m)} - m_\infty]P_{l^m, l^{m+1}},$$

where $j(l) = I$ or O if l^m is of the form $v + \aleph$ or v, respectively. We use the convention that, upon fixing k, the final l index l^{k+1} is the same as $\hat{r}$.

As in the proof of Lemma 3.2, our biggest issue is to determine the value of the k product,

$$(3.16) \qquad Ex_k(i,j) := \sum_{l^1, \ldots, l^k} |P_{i,l^1}| \prod_{m=1}^{k} |P_{l^m, l^{m+1}}|.$$

Now, we cannot directly apply the analysis detailed in the aforementioned lemma, as the proof there assumed a purely tree-like structure.

What we have do instead is appeal to the fact that our graph has a tree-like covering and relate the values of P_{l^i, l^j} to the Green's functions of the covering. Now, the results of Lemma A.6 show that, up to a constant factor, $|P_{l^i, l^{i+1}}|$ can be bounded by $2\max(|P_{\tilde{l}^i, \tilde{l}_1^{i+1}}|, |P_{\tilde{l}^i, \tilde{l}_2^{i+1}}|)$. Here, $\tilde{l}^i$ is some chosen covering of l^i and $\tilde{l}_1^{i+1}$ and $\tilde{l}_2^{i+1}$ are the two closest chosen coverings of l^{i+1} to $\tilde{l}^i$. Thus, when we consider the lift to the covering tree,

$$\sum_{l^1, \ldots, l^k} |P_{i,l^1}| \prod_{m=1}^{k} |P_{l^m, l^{m+1}}|,$$

can be upper bounded by the sum

$$2^{k+1} \sum_{\substack{\text{dist}(\tilde{l}^n, \tilde{l}^{n+1}) \leq 3\,\text{diam}(\mathcal{G}) \\ \tilde{l}^{k+1} \text{ covers } j}} |P_{\tilde{i}, \tilde{l}^1}| \prod_{m=1}^{k} |P_{\tilde{l}^m, \tilde{l}^{m+1}}|.$$

Here, $\tilde{i}$ is a fixed lift of i. The restriction that $\text{dist}(\tilde{l}^n, \tilde{l}^{n+1}) \le 3 \, \text{diam}(\mathcal{G})$ comes from the fact that the closest and the second closest cover of any point differ by distance at most $|C| \le 2 \, \text{diam}(\mathcal{G})$. To perform the above summation, we first fix the final cover of the point j, and then compute the above sum in the same manner as in the proof of Lemma 3.2.

Namely, we decompose the product $|P_{l^m, l^{m+1}}|$ into its K factors, introduce the least common ancestor variables, and construct the associated tree. The only difference with the constructed tree is that the K factors associated with the vertices appearing on the path p_{ij} between the lift of i and j may not appear twice. To deal with these vertices, we remove the edges in the tree that correspond to vertices that are connected via this path. The remaining forest graph can be bounded via the same methods of the proof of Lemma 3.2 and will give the same bound.

We only need to reintroduce the K factors that we did not consider on the path between i and j. Observe that at least $dist(\tilde{i}, \tilde{j}) - 3k$ of the vertices have K factors associated with them, since each factorization from (A.13) can lose no more than 3 natural K factors. Once can see that the K factors corresponding to this path between i and j will give a factor bounded by $\left(\frac{1}{\sqrt{d-1}}\right)^{\text{dist}(\tilde{i},\tilde{j})} (\sqrt{d-1})^{3k}$. We see that the last $(\sqrt{d-1})^{3k}$ can be absorbed by $|Q_I - m_\infty|^k$. This latter factor can allow us again to sum over the possible lifts of $\tilde{j}$, since the successive lifts of $\tilde{j}$ increase by distance $|C|$ from the lift $\tilde{i}$. This allows the term $\left(\frac{1}{\sqrt{d-1}}\right)^{\text{dist}(\tilde{i},\tilde{j})}$ to be summable over successive lifts.

Now, if we want to be careful, we can improve the first order coefficient from $(1 + \text{diam}(\mathcal{G}))^3$ to $1 + \text{dist}(i, j)$. The method of proof is similar, one only needs to more careful in the discussion of branching. First, consider the path p_{ij} between the lifts of i and j. To choose $\tilde{l}_1$, one first picks a point Q_i on the path p_{ij} and then starts branching from there. There are $1 + \text{dist}(\tilde{i}, \tilde{j})$ choices for the point Q_i. Now, if one takes into account that $\tilde{l}_1$ is the lift of a boundary point, one can use the superior analysis of the first order coefficient from Lemma 3.2 to get rid of the factor of $(1 + \text{diam}(\mathcal{G}))^2$ and replace it with an order 1 coefficient. This gives us our final answer. $\qquad\square$

Because of the particular dependence of pointwise estimates on Green's functions from i to j, the worst case bound on pointwise Green's function are generally worse than the averaged case bound. By a similar analysis, one can derive average case bounds.

Corollary 3.6. *Consider the setting of the previous Lemma 3.5. Assume also that we have the improved bound $(\log N)^9 [|Q_I - m_\infty| + |Q_O - m_\infty|] \ll 1$. Let $\hat{r}$ be a distinguished vertex of $\mathcal{G}$, which we will call the 'root'. We have that,*

$$(3.17) \qquad \sum_{x \in \delta(\mathcal{G})} \sum_{o_x} |G_{\hat{r}+o_r, x+o_x}(Ext(\mathcal{G}, Q_I, Q_O))|^2| \lesssim 1.$$

Here, o_r and o_x are allowed to be either 0 or $\aleph$.

Proof. Without loss of generality we may assume that $o_r = 0$.

Let $P_{\hat{r}x} := G_{\hat{r}x}(Ext(\mathcal{G}, m_\infty, m_\infty))$ and use G_{rx} as a shorthand for the Green's function $G_{\hat{r}x}(Ext(\mathcal{G}, Q_I, Q_O))$. We have that,

$$(3.18) \qquad \sum_{x \in \delta(\mathcal{G})} \sum_{o_x} |G_{\hat{r}, x+o_x}|^2 \le 2 \sum_{x \in \delta(\mathcal{G})} \sum_{o_x} [|P_{\hat{r}, x+o_x}|^2 + |G_{\hat{r}, x+o_x} - P_{\hat{r}, x+o_x}|^2].$$

Now, we have the correct bound for $\sum_{x \in \delta(\mathcal{G})} \sum_{o_x} |P_{\hat{r}, x+o_x}|^2$ by Corollary A.8. We perform the perturbation expansion for $G_{\hat{r}, x+o_x} - P_{\hat{r}, x+o_x}$.

Recall Ex_k from equation (3.16). We have,

$$
\begin{aligned}
|G_{\hat{r}, x+o_x} - P_{\hat{r}, x+o_x}|^2 &\le \left(\sum_{k=1}^{\infty} Ex_k(\hat{r}, x + o_x)[|Q_I - m_\infty| + |Q_O - m_\infty]^k \right)^2 \\
&\le \sum_{k=1}^{\infty} Ex_k(\hat{r}, x+o_x)^2 (2[|Q_I - m_\infty| + |Q_O - m_\infty|])^{2k} \sum_{k=1}^{\infty} 2^{-2k}.
\end{aligned}
$$
(3.19)

Now, we see our mail goal is to understand the asymptotics of the averaged sum,

$$(3.20) \qquad \sum_x \sum_{o_x} Ex_k(\hat{r}, x + o_x)^2 \leq (1 + \text{diam}(\mathcal{G}))^k \sum_{l^1,\ldots,l^k,l^{k+1}} |P_{il^1}|^2 \prod_{m=1}^{k} |P_{l^m,l^{m+1}}|^2$$

Now, we can write,

$$(3.21) \qquad \sum_x \sum_{o_x} Ex_k(\hat{r}, x + o_x)^2 = \sum_{l^1,\ldots,l^k,x,o_x,\hat{l}^1,\ldots,\hat{l}^k} |P_{i,l^1}||P_{i,\hat{l}^1}| \prod_{m=1}^{k-1} |P_{l^m,l^{m+1}}||P_{\hat{l}^m,\hat{l}^{m+1}}||P_{l^k,j}||P_{\hat{l}^k,\hat{j}},$$

However, one can relabel $x + o_x$ as a new summation variable l^{k+1} and relabel $\hat{l}^j$ to $l^{k+1+(k-j)+1}$ to understand this as $Ex_{2k+1}(\hat{r},\hat{r})$, which we can bound by $\lesssim (1 + \text{diam}(\mathcal{G})^{6k+3}$. Under our assumption that $(1 + \text{diam}(\mathcal{G}))^9[|Q_I - m_\infty| + |Q_O - m_\infty|] \ll 1$, this shows that the sum in (3.20) is $\ll 1$. This gives us our desired result. $\qquad\square$

The stability result detailed above shows, as one would expect, that the approximation of $\text{Ext}_i(\mathcal{G}, Q_I, Q_O)$ to the infinite tree would get better as $\mathcal{G}$ becomes closer to the tree. This is an essential input to our next lemma, which details how the Green's functions change as the graph neighborhood around two points change.

Lemma 3.7. *Let $\mathcal{G}$ be a graph containing two vertices i and j. Let $\mathcal{H}$ be a subgraph of $\mathcal{G}$ that contains the radius r neighborhood of i and j. Then, we have the following estimate,*

$$(3.22) \qquad \begin{aligned} |G_{ij}(Ext(\mathcal{G}, Q_I, Q_O)) - G_{ij}(Ext(\mathcal{H}, Q_I, Q_O))| &\leq \left(\frac{1}{d-1}\right)^r \\ &+ (1 + diam(\mathcal{G})^2[S_g^1|Q_I + Q_O - 2m_\infty(z,w)| + S_g^2|Q_I - Q_O|] \\ &+ (\log N)^2[|Q_I - m_\infty|^2 + |Q_O - m_\infty|^2], \end{aligned}$$

provided we have that $(\log N)^9|Q_{I/O} - m_\infty| \ll 1$.

Proof. We can write the adjacency matrices of $Ext(\mathcal{G}, Q_I, Q_O)$ as,

$$(3.23) \qquad \begin{bmatrix} A & B \\ B^* & C \end{bmatrix},$$

Here, A is the adjacency matrix for the subgraph $\mathcal{H}$. B contains information on how $\mathcal{H}$ connects to the complement and C is the adjacency information of the complement.

By the Schur complement formula, we have,

$$(3.24) \qquad G_{ij}(Ext(\mathcal{G}, Q_I, Q_O)) = [(A - BC^{-1}B^*)^{-1}]_{ij}.$$

At this point, we can apply the resolvent estimates,

$$(3.25) \qquad \begin{aligned} &|G_{ij}(Ext(\mathcal{G}, Q_I, Q_O)) - G_{ij}(Ext(\mathcal{H}, Q_I, Q_O))| \\ &= \sum_{a,b,c,d} G_{ia}(Ext(\mathcal{G}, Q_I, Q_O))B_{ab}(C_{bc}^{-1} - \delta_{b,c}Q_k)B_{cd}^* G_{dj}(Ext(\mathcal{H}, Q_I, Q_O)) \end{aligned}$$

Here, $Q_k = Q_I$ if b is of the form $v + \aleph$ for a vertex v in the complement of $\mathcal{H}$ while $Q_k = Q_O$ if b is of the form v for a vertex v in the complement of $\mathcal{H}$. Let us first consider the off-diagonal terms with $b \neq c$.

In order to get non-zero terms, we see from the above that a and d must be vertices on the boundary of $\mathcal{H}$, and are thus distance greater than r from a and b. Furthermore, b and c must be vertices on the boundary of $\mathcal{G}\backslash\mathcal{H}$ that are adjacent to vertices in $\mathcal{H}$. Our combinatorial assumption on the graph $\mathcal{G}$ allows us to make the following assertion: there are only finitely many pairs of vertices (b, c) with $b \neq c$ such that $C_{bc}^{-1} \neq 0$.

Just to give a summary for the reasoning behind this point, we remark that C^{-1} is a diagonal matrix on the connected components of $\mathcal{G}\backslash\mathcal{H}$. Now, the fact that $\mathcal{G}$ has at most one cycle will show that all but one connected neighborhood of $\mathcal{G}\backslash\mathcal{H}$ is purely treelike. Furthermore, the only connected neighborhood that is not tree-like can contain no more than $O(1)$ many vertices that like on the boundary of $\mathcal{G}\backslash\mathcal{H}$ lest there be

too many cycles. For this $O(1)$ many off-diagonal entries, we can apply Lemma 3.5 to show that the diagonal entry $|C_{bc}^{-1}|$ itself is $O(1)$.

We also have the Green's function bound $|B_{ab}| \leq \left(\frac{1}{\sqrt{d-1}}\right)^r$ from Lemma 3.5. Thus, our bound on the off-diagonal terms is $O\left(\frac{1}{\sqrt{d-1}}\right)^{2r}$.

It is now left to understand the contribution of the terms that are on the diagonal. Due to Lemma 3.2, we can argue that the size of the differences $|C_{bc}^{-1} - Q_k| \lesssim (\log N)^2[|S_g^1||Q_I + Q_O - 2m_\infty(z,w)| + |S_g^2||Q_I - Q_O|] + (\log N)^3[|Q_I - m_\infty|^2 + |Q_O - m_\infty|^2]$. Now, given b, the only entries B_{ab} and B_{bd} that are non-zero correspond to vertices adjacent to b.

Furthermore, Corollary 3.6 can show that the contribution $\sum_a |G_{ia}||G_{aj}| \lesssim 1$. One can only gain a constant factor when considering $\sum_{a,b,d} |G_{ia}||B_{ab}||B_{bd}^*||G_{dj}|$. This shows that the contribution of the off-diagonal entries is

$$\lesssim (\log N)^2[|S_g^1||Q_I + Q_O - 2m_\infty(z,w)| + |S_g^2||Q_I - Q_O|] + (\log N)^3[|Q_I - m_\infty|^2 + |Q_O - m_\infty|^2]$$

$\square$

3.2. Stability Bounds.

Theorem 3.8. *Consider the following pairs of equations,*

$$(3.26) \qquad \begin{aligned} Q_I &= Y_{i,K}(Q_I, Q_O) + E_I \\ Q_O &= Y_{i,K}(Q_I, Q_O) + E_O, \end{aligned}$$

Assume further that it is known that $|Q_{I,O} - m_\infty|^2 \lesssim \frac{S_g^1|E_O|}{(\log N)^4}$, and $N^\delta|Q_{I,O} - m_\infty| \ll 1$ for some $\delta > 0$. Recall the singularity functions from Definition 3.1.

Fix some parameter $\mathfrak{a} > 0$. There exists K in the domain $[\mathfrak{a} \log_{d-1} \log N, 2\mathfrak{a} \log_{d-1} \log N]$ such that, we have the stability estimates,

$$(3.27) \qquad \begin{aligned} |Q_I - Q_O| &\leq \frac{|E_I| + |E_O|}{|S_g^2|} + O\left(\frac{|E_I| + |E_O|}{\log N}\right) \\ |Q_I + Q_O - 2m_\infty| &\leq \frac{|E_I| + |E_O|}{|S_g^1|} + O\left(\frac{(|E_I| + |E_O|)}{\log N}\right) \end{aligned}$$

Proof. This is mainly a consequence of perturbation theory using the resolvent identity.

$$(A - z)^{-1} = (B - z)^{-1} + \sum_{p=1}^\infty (B - z)^{-1}[(B - A)(B - z)^{-1}]^p.$$

We apply this identity to the compute the difference of the resolvents,

$$Y_{i,K}(Q_I, Q_O) - Y_{i,K}(m_\infty, m_\infty) = G(\mathcal{T}_K^1, Q_I, Q_O)_{\hat{r}+\aleph, \hat{r}+\aleph} - G(\mathcal{T}_K^1, m_\infty, m_\infty)_{\hat{r}+\aleph, \hat{r}+\aleph}.$$

The main benefit is that the second quantity in the line above can be explicitly computed; these are the Green's function entries for the infinite tree.

Using that $N^\delta|Q_I - m_\infty|, N^\delta|Q_O - m_\infty| \ll 1$ for some $\delta > 0$ and furthermore that we have the a-priori bound that $|Q_I - m_\infty|^2 \ll \frac{S_g^1}{(\log N)^4}[|E_I| + |E_O|]$, then it would suffice to consider only the first order term in the resolvent expansion.

By applying the resolvent identity and the more detailed computations of Lemma 3.2, we have that,

$$
\begin{aligned}
G(\mathcal{T}_K^1, Q_I, Q_O)_{r+\aleph, r+\aleph} &- G(\mathcal{T}_K^1, m_\infty, m_\infty)_{\hat{r}+\aleph, \hat{r}+\aleph} \\
&= \frac{d-1}{d-1}(Q_I - m_\infty) \sum_{v \in V_O^K} G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, v} \overline{G_{\mathcal{T}_1}(\overline{z}, w)_{\hat{r}+\aleph, v}} \\
&+ \frac{d}{d-1}(Q_O - m_\infty) \sum_{v \in V_O^K} G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, v+\aleph} \overline{G_{\mathcal{T}_1}(\overline{z}, w)_{\hat{r}+\aleph, v+\aleph}} \\
&+ \frac{d}{d-1}(Q_I - m_\infty) \sum_{v \in V_I^K} G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, v} \overline{G_{\mathcal{T}_1}(\overline{z}, w)_{\hat{r}+\aleph, v}} \\
&+ \frac{d-1}{d-1}(Q_O - m_\infty) \sum_{v \in V_I^K} G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, v+\aleph} \overline{G_{\mathcal{T}_1}(\overline{z}, w)_{\hat{r}+\aleph, v+\aleph}} \\
&+ O((\log N)^3(|Q_I - m_\infty|^2 + |Q_O - m_\infty|^2))
\end{aligned}
\tag{3.28}
$$

Here, V_O^K are the leaves of the tree $\mathcal{T}_K^1$ whose only edge is an outgoing edge (so it has $d-1$ out-edges missing) while V_I^K are those leaves whose only edge is an ingoing edge (so it has d out-edges missing). When going to the second line, we used the fact that $G(\mathcal{T}_K^1, m_\infty, m_\infty)$ are equal to the Green's functions of the infinite tree $\mathcal{T}_1$. We also used the fact that $G_{ab}(z,w) = \overline{G_{ba}(\overline{z}, w)}$.

Provided we know that the coefficient matrix corresponding to the first order terms are sufficiently well controlled, we can derive stability estimates on $Q_i - m_\infty$ by merely inverting the coefficients. We can derive the first order coefficients by a direct recursion as follows:consider the root $\hat{r}$ and the unique path $\hat{r} \to v_1 \to v_2 \to \ldots \to v_n$ to some vertex v_n (the path from $\hat{r}$ to v_n does not care about the orientation of the edges.)

If $\hat{r} \to v_1$ is a forward edge, then $G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, v_n+o_n} = G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, \hat{r}} \frac{1}{\sqrt{d-1}} G_{\mathcal{T}_1}(z,w)_{v_1+\aleph, v_n+o_n}$, where o_n can either be 0 or $\aleph$. If instead $\hat{r} \to v_1$ is a backward edge, then $G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, v_n+o_n} = G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, \hat{r}+\aleph} \frac{1}{\sqrt{d-1}} G_{\mathcal{T}_2}(z,w)_{v_1, v_n+o_n}$.

Combining these estimates can give us a recursion for the value of the coefficients of the linear term in the matrix. Indeed, we have the following recursion to determine the coefficients.

$$
\begin{aligned}
\sum_{v \in V_O^K} &G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, v+\aleph} \overline{G_{\mathcal{T}_1}(\overline{z}, w)_{\hat{r}+\aleph, v+\aleph}} \\
&= \frac{d}{d-1} G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, \hat{r}} \overline{G_{\mathcal{T}_1}(\overline{z}, w)_{r+\aleph, r}} \sum_{v \in V_O^{K-1}} G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, v+\aleph} \overline{G_{\mathcal{T}_1}(\overline{z}, w)_{\hat{r}+\aleph, v+\aleph}} \\
&+ \frac{d-1}{d-1} G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, \hat{r}+\aleph}(z,w) \overline{G_{\mathcal{T}_1}(z,w)_{\hat{r}+\aleph, \hat{r}+\aleph}(\overline{z}, w)} \sum_{v \in V_O^{K-1}} G_{\mathcal{T}_2}(z,w)_{\hat{r}, v+\aleph} \overline{G_{\mathcal{T}_2}(\overline{z}, w)_{r, v+\aleph}}.
\end{aligned}
\tag{3.29}
$$

We will more generally define the quantity,

$$
A_j^{K,l,o} := \sum_{v \in V_l^K} G_{T_j}(z,w)_{\hat{r}+o(j), v+o} \overline{G_{T_j}(\overline{z}, w)_{\hat{r}+o(j), v+o}}.
\tag{3.30}
$$

l can be either I or O for in or for out vertices. K is the depth of the nodes we are considering on the tree. j can take the values either 1 or 2 (representing either the tree $\mathcal{T}_1$ or $\mathcal{T}_2$). o can be either $\aleph$ or 0 (representing the index of the vertex we are considering). $o(j)$ is either $\aleph$ or 0, when $j = 1$, then $o(j) = \aleph$, otherwise $o(j) = 0$ for $j = 2$. $(3-j)$ will denote the opposite o assignment. (Namely, it is 0 when $j = 1$ and $\aleph$ when $j = 2$.)

We see that one way to write the recursive equation considered above is as follows:

$$
\begin{aligned}
(3.31) \quad A_j^{K,l,o} &= \frac{d}{d-1} G_{T_j}(z,w)_{\hat{r}+o(j),\hat{r}+o(3-j))} \overline{G_{T_j}(\overline{z},w)_{\hat{r}+o(j),\hat{r}+o(3-j)}} A_j^{K-1,l,o} \\
&+ \frac{d-1}{d-1} G_{T_j}(z,w)_{\hat{r}+o(j),\hat{r}+o(j)} \overline{G_{T_j}(\overline{z},w)_{\hat{r}+o(j),\hat{r}+o(j)}} A_{3-j}^{K-1,l,o}
\end{aligned}
$$

From looking at the infinite tree, we would know that

$$
\frac{d}{d-1} G_{T_j}(z,w)_{\hat{r}+o(j),\hat{r}+o(3-j))} \overline{G_{T_j}(\overline{z},w)_{\hat{r}+o(j),\hat{r}+o(3-j)}}
$$

is a constant in j, precisely matching Y of (A.1). Similarly,

$$
\frac{d-1}{d-1} G_{T_j}(z,w)_{\hat{r}+o(j),\hat{r}+o(j)} \overline{G_{T_j}(z,w)_{\hat{r}+o(j),\hat{r}+o(j)}}
$$

is a constant in j, precisely matching X of (A.1).

We see we have the general recursions,

$$
(3.32) \quad \begin{bmatrix} A_1^{K,l,o} \\ A_2^{K,l,o} \end{bmatrix} = \begin{bmatrix} Y & X \\ X & Y \end{bmatrix} \begin{bmatrix} A_1^{K-1,l,o} \\ A_2^{K-1,l,o} \end{bmatrix}.
$$

The matrix in the middle has eigenvalues $X+Y$ with eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $Y-X$ with eigenvector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

In fact, we observe that we have the same recursion matrix for the actual coefficients $\frac{d-1}{d-1}A_1^{K,O,0} + \frac{d}{d-1}A_1^{K,I,0}$ and $\frac{d}{d-1}A_1^{K,O,\aleph} + \frac{d-1}{d-1}A_1^{K,I,\aleph}$ that we actually need.

Indeed, we see we have the coefficient matrix,

$$
\begin{aligned}
&\begin{bmatrix} G(\mathcal{T}_K^1, Q_I, Q_O)_{\hat{r}+\aleph,\hat{r}+\aleph} - G(\mathcal{T}_K^1, m_\infty, m_\infty)_{\hat{r}+\aleph,\hat{r}+\aleph} \\ G(\mathcal{T}_K^2, Q_I, Q_O)_{\hat{r},\hat{r}} - G(\mathcal{T}_K^2, m_\infty, m_\infty)_{\hat{r},\hat{r}} \end{bmatrix} \\
&= \begin{bmatrix} \frac{d-1}{d-1}A_1^{K,O,0} + \frac{d}{d-1}A_1^{K,I,0} & \frac{d}{d-1}A_1^{K,O,\aleph} + \frac{d-1}{d-1}A_1^{K,I,\aleph} \\ \frac{d-1}{d-1}A_2^{K,O,0} + \frac{d}{d-1}A_2^{K,I,0} & \frac{d}{d-1}A_2^{K,O,\aleph} + \frac{d-1}{d-1}A_2^{K,I,\aleph} \end{bmatrix} \begin{bmatrix} Q_I - m_\infty \\ Q_O - m_\infty \end{bmatrix} + O(|Q_I - m_\infty|^2 + |Q_O - m_\infty|^2) \\
&= \begin{bmatrix} Y & X \\ X & Y \end{bmatrix}^K \begin{bmatrix} \frac{d-1}{d-1}A_1^{0,0,0} + \frac{d}{d-1}A_1^{0,I,0} & \frac{d}{d-1}A_1^{0,O,\aleph} + \frac{d-1}{d-1}A_1^{0,I,\aleph} \\ \frac{d-1}{d-1}A_2^{0,0,0} + \frac{d}{d-1}A_2^{0,I,0} & \frac{d}{d-1}A_2^{0,O,\aleph} + \frac{d-1}{d-1}A_2^{0,I,\aleph} \end{bmatrix} \begin{bmatrix} Q_I - m_\infty \\ Q_O - m_\infty \end{bmatrix} + O(|Q - m_\infty|^2)
\end{aligned}
$$

Now we can apply our initial data conditions. This is the part where our coefficient dependence on the terms I, O, o are seen. For example, we have $A_1^{0,O,o} = 0$ and $A_2^{0,I,o} = 0$. In addition, $A_1^{0,I,\aleph} = X$, $A_1^{0,I,0} = \frac{d-1}{d}Y$, $A_2^{0,O,\aleph} = \frac{d-1}{d}Y$, and $A_2^{0,O,0} = X$. Substituting these values inside the initial data matrix will give us the following equation.

$$
(3.33) \quad \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} Y & X \\ X & Y \end{bmatrix}^{K+1} \right) \begin{bmatrix} Q_I - m_\infty \\ Q_O - m_\infty \end{bmatrix} = \begin{bmatrix} E_I + O(|Q - m_\infty|^2) \\ E_O + O(|Q - m_\infty|^2) \end{bmatrix}
$$

If we let C denote the coefficient matrix appearing in front of the vector of $Q_i - m_\infty$ in the above equation, then we see that C has eigenvalues $1 - (X+Y)^{K+1}$ and $1 - (Y-X)^{K+1}$.

By an appropriate choice of K, one can try to ensure that the singularity of these eigenvalues is no worse than the singularity of $1 - (X+Y)$ or $1 - (Y-X)$. $\qquad \square$

4. Estimates on $G^{(\mathbb{T})}$: Proof of Proposition 2.15

Recall we use $\mathbb{T}$ to represent the neighborhood of size ℓ around a point. Our main strategy to prove Proposition 2.15 would be to apply the Schur complement formula to relate the values of G to $G^{(\mathbb{T})}$. To better visualize this, let us first write the adjacency matrix for the graph $\mathcal{G}$ as follows.

$$
(4.1) \quad G^{-1} = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}
$$

A is a $2|\mathbb{T}|$ by $2|\mathbb{T}|$ matrix, B is a $2|\mathbb{T}|$ by $2N - 2|\mathbb{T}|$ matrix, and D is a $2N - 2|\mathbb{T}|$ by $2N - 2|\mathbb{T}|$ matrix.

If we let $\{v_1, \ldots, v_{|\mathbb{T}|}\}$ be the indices of the vertices in $\mathbb{T}$. Then, we find A by considering the principal submatrix given by the indices $\{v_1, v_1 + N, \ldots, v_{|\mathbb{T}|}, v_{|\mathbb{T}|} + N\}$. B is the adjacency graph of the edges connecting $\mathbb{T}$ to $\mathbb{T}^c$. D is the adjacency graph of $\mathbb{T}^c$.

Remark 4.1 (Notational convention for adjacency matrix decompositions). We remark here for the reader that though we will use the same notation A, B, and D to denote the upper-left diagonal block, the upper right off-diagonal block, and the lower-right diagonal blocks respectively for some adjacency matrix, the specific adjacency matrices will change based on the section. The exact matrix referred to by the notation A, B, or D will be clear in context.

Remark 4.2 (Notational Convention for Perturbation Theory). In all of the following few sections, our goal is to understand the Green's function of the Hermitized adjacency matrix of the graph $\mathcal{G}$. These matrices will be denoted as G or some modifications of G such as $G^{(\mathbb{T})}$, or $\tilde{G}$ based on the context; for example, $\tilde{G}$ will correspond to the Green's function of a switched graph.

These Green's functions will always be compared to a Green's function that will be constructed by understanding the local neighborhood around the vertices. For example, we will compare the Green's function G_{xy} with $G(Ext(B_r(\mathbb{T} \cup \{x, y\}, \mathcal{G}), Q_I, Q_O))$. The specific comparison will highly depend on the context, but we will always denote this secondary matrix by P for simplicity of presentation. We warn the reader here that the specific matrix P found in the estimates will not be the same in different sections; however, their role as a characterizing the local neighborhood will always be the same.

In this section we let P be a shorthand for $G(Ext(B_r(\mathbb{T} \cup \{i, j\}, \mathcal{G}), Q_I, Q_O))$. Since our induction hypothesis is that $|G_{ij} - G(Ext(B_r(\{i, j\}, \mathcal{G}), Q_I, Q_O))| \lesssim \epsilon$, we can use Lemma 3.7 to assert that we would still have $|G_{ij} - P_{ij}| \lesssim \epsilon$. We only used the fact here that we contain the radius r neighborhood around i and j. The benefit of using P_{ij} instead of $G(Ext(B_r(\{i, j\}, \mathcal{G}), Q_I, Q_O)$ is that we have a more convenient adjacency matrix to apply perturbation theory with in P_{ij}. In later sections, we will freely adjust the neighborhood to which we consider the extension to be more useful in perturbation theory without any explicitly mention; the justification of this change will always be Lemma 3.7.

As we have mentioned before, the benefit of the matrix P is that we have the adjacency matrix decomposition,

$$
(4.2) \qquad P^{-1} = \begin{bmatrix} A & B \\ B^* & D_0 \end{bmatrix}.
$$

A is the same adjacency matrix when of the graph restricted to the neighborhood $\mathbb{T}$. We technically abuse notation when we use the matrix B, but the upper-left and lower-right blocks of P^{-1} have the same non-zero elements as those of G^{-1} (these are the elements that connect the vertices of $\mathbb{T}$ to the complement $\mathbb{T}^c$.)

Applying the Schur complement formula, we see that,

$$
(4.3) \qquad \begin{aligned} G^{(\mathbb{T})} &= G - G(G|_{\mathbb{T}})^{-1} G, \\ P^{(\mathbb{T})} &= P - P(P|_{\mathbb{T}})^{-1} P. \end{aligned}
$$

If we take the difference of these expressions, we see that,

$$
(4.4) \qquad \begin{aligned} G^{(\mathbb{T})} - P^{(\mathbb{T})} &= G - P - (G - P)(G|_{\mathbb{T}})^{-1} G \\ &\quad - P[(G|_{\mathbb{T}})^{-1} - (P|_{\mathbb{T}})^{-1}]G - P(P|_{\mathbb{T}})^{-1}(G - P). \end{aligned}
$$

This equation is the basis of the proof of Proposition 2.15, which we reproduce here.

Proposition 4.3. *Consider a graph $\mathcal{G}$ in $\Omega^0(z, w)$. Fix a vertex o and let $\mathbb{T}$ be the radius ℓ neighborhood around $\mathbb{T}$. Assume, in addition, that o has a radius $\mathfrak{R}$ neighborhood with excess at most 1. Then, we have the following estimates on the Green's function of the graph $\mathcal{G}$ with the neighborhood $\mathbb{T}$ removed. Let i and j be vertices outside of $\mathbb{T}$ in $\mathcal{G}$ and let P denote $G(Ext(B_r(\mathbb{T} \cup \{i, j\}, \mathcal{G}), Q_I, Q_O))$. Then, we have that,*

$$
(4.5) \qquad |G_{ij}^{(\mathbb{T})} - P_{ij}^{(\mathbb{T})}| \le \epsilon'(z, w).
$$

Before we can make assertions on the difference between G and P, we must first make assertions on the values of $(G|_{\mathbb{T}})^{-1}$. This will be compared to the quantity $(P|_{\mathbb{T}})^{-1}$; we start with a combinatorial lemma whose goal is to understand $(P|_{\mathbb{T}})^{-1}$ using the fact that we are performing a computation on a treelike neighborhood.

Lemma 4.4. *We have the following estimates of $(P|_{\mathbb{T}})^{-1}$. For any index x corresponding to a vertex in $\mathbb{T}$, we have,*

$$(4.6) \qquad \sum_y |(P|_{\mathbb{T}})^{-1}_{xy}| \lesssim 1.$$

Proof. As is standard, we will apply the Schur complement formula. We have,

$$(4.7) \qquad (P|_{\mathbb{T}})^{-1}_{xy} = A_{xy} - \sum_{a_\alpha, a_\beta} (B^*)_{xa_\alpha} [(D_0)^{-1}]_{a_\alpha, a_\beta} B_{a_\beta, y}.$$

Here, a_α and a_β vary over the indices of vertices in $\mathbb{T}^c$ that are adjacent to vertices l_α on the boundary of $\mathbb{T}$. We see that the matrix product is trivial unless $x = l_\alpha$ and $y = l_\beta$ are indices corresponding to some vertices on the boundary of $\mathbb{T}$.

If we fix x not of the form l_α and we sum up over y, we do not have to worry about the last term on the RHS of equation (4.7). Furthermore, A_{xy} can only be nonzero for $O(1)$ many terms (the indices y must either correspond to vertices adjacent to the one corresponding to x, or they would correspond to the same index as y itself).

Let us now consider the case that x corresponds to some vertex l_α on the boundary. Since $\mathbb{T}$ is centered around a neighborhood of excess at most 1, this means that $[(D_0)^{-1}]_{a_\alpha, a_\beta} = 0$ when $a_\alpha \neq a_\beta$, except for at most $O(1)$ many pairs. Recall D_0 is a block diagonal matrix with the blocks corresponding to neighborhoods of $\mathbb{T}^c$. This further gives only $O(1)$ many options for y to get a nonzero quantity. $\qquad \square$

For the sake of completeness, we also present the corresponding combinatorial bound for $|P|_{iy}$.

Lemma 4.5. *If we let i be a vertex in $\mathbb{T}^c$, we have the following estimates of $P|_{\mathbb{T}}$.*

$$(4.8) \qquad \sum_{y \in \mathbb{T}} |P_{iy}| \lesssim (\log N)^{\mathrm{j}(d)\mathfrak{a}},$$

where $\mathrm{j}(d) = \log_{d-1}\left[\frac{1}{2} + \frac{1}{2}\sqrt{\frac{2d-1}{d-1}}\right] < 1/2$ is a constant that decays to 0 as d increases to ∞.

Remark 4.6. The main difference between the case with $|w| = 0$ and $|w| \neq 0$ is that our combinatorial estimates on the tree-like graph now are bounded by $(\log N)$ to a power depending on $\mathfrak{a}$, but the coefficient behind this $\mathfrak{a}$ can be made arbitrarily small by choosing d large, but still finite.

Proof. For this lemma, we have to apply the estimate $|P_{iy}| \lesssim \left(\frac{1}{\sqrt{d-1}}\right)^{\mathrm{dist}(i,y)}$ coming from perturbation theory from the infinite directed d-regular graph and as in Lemma A.6.

Let p denote the shortest path between i and the center o of the neighborhood $\mathbb{T}$. We let a_k denote the vertex on the path p that is of distance k from i. We also let $\mathcal{C}_j$ denote the vertices on $\mathbb{T}$ that are reachable from a_k without using an edge in p. Note that it is possible that some sets $\mathcal{C}_{j_1} = \mathcal{C}_{j_2}$ for $j_1 \neq j_2$, but this is not possible in $\mathbb{T}$ is purely tree-like.

Now, let us consider the computation of $\sum_{y \in \mathbb{T}} |P_{yj}|$ is tree-like. If $\mathbb{T}$ were not tree-like, then Lemma A.6 would show that we would get the same bound up to a constant factor. Since we are only concerned about the correct size up to a constant factor, this is not a major loss for us.

By the factorization lemma A.3, we see that $\sum_{y \in \mathcal{C}_j} |P_{iy}| \lesssim \max\left\{\sqrt{X}, \sqrt{\frac{d-1}{d}}\sqrt{Y}\right\}^{j-1} \sum_{y \in \mathcal{C}_j} |P_{a_j y}|$.

Now, the computation of $\sum_{y \in \mathcal{C}_j} |P_{a_j y}|$ is like a computation performed if a_j were the root of a j level tree. If we consider the quantities,

$$(4.9) \qquad L_{1,k} := \sum_{\text{dist}(y,r)=k} |(G_{\mathcal{T}_1})_{r+\aleph, y+o_y}|, L_{2,k} := \sum_{\text{dist}(y,r)=k} |(G_{\mathcal{T}_2})_{r, y+o_y}|,$$

and r is the root of $\mathcal{T}_1$ and $\mathcal{T}_2$, then we have the recursion,

$$(4.10) \qquad \begin{bmatrix} L_{1,k} \\ L_{2,k} \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{d}{d-1}}\sqrt{Y} & \sqrt{X} \\ \sqrt{X} & \sqrt{\frac{d}{d-1}}\sqrt{Y} \end{bmatrix} \begin{bmatrix} L_{1,k-1} \\ L_{2,k-1} \end{bmatrix}.$$

Thus, we see that $L_{1,k}, L_{2,k} \lesssim \left[\sqrt{\frac{d}{d-1}}\sqrt{Y} + \sqrt{X} \right]^k$.

We thus see that

$$\sum_{y \in \mathcal{C}_j} |P_{iy}| \lesssim \ell \max\left\{ \sqrt{X}, \sqrt{\frac{d-1}{d}}\sqrt{Y} \right\}^{j-1} \left[\sqrt{\frac{d}{d-1}}\sqrt{Y} + \sqrt{X} \right]^{j-1}.$$

Thus,

$$(4.11) \qquad \sum_{y \in \mathbb{T}} |P_{iy}| \lesssim (\ell)^2 \left\{ \sqrt{X}, \sqrt{\frac{d-1}{d}}\sqrt{Y} \right\}^{\ell-1} \left[\sqrt{\frac{d}{d-1}}\sqrt{Y} + \sqrt{X} \right]^{\ell-1}.$$

We now let $a^2 := X$ and $b^2 := Y$.

We solve the maximization problem

$$(4.12) \qquad \max\left\{ \sqrt{\frac{d}{d-1}}ab + a^2 : a^2 + b^2 = 1 \right\}.$$

We remark here that the other relevant maximization problem would be,

$$(4.13) \qquad \max\left\{ ab + \sqrt{\frac{d-1}{d}}b^2 : a^2 + b^2 = 1 \right\},$$

which is manifestly smaller.

By Lagrange muiltipliers, this amounts to maximizing,

$$(4.14) \qquad \sqrt{\frac{d}{d-1}}ab + a^2 - \lambda(a^2 + b^2).$$

Taking derivatives in b gives us,

$$(4.15) \qquad a = 2\lambda\sqrt{\frac{d-1}{d}}b.$$

Taking derivatives in a give us,

$$(4.16) \qquad 2a - 2\lambda a + \sqrt{\frac{d}{d-1}}b = 0.$$

Substituting in our previous relation between a and b gives us the following relationship on λ.

$$(4.17) \qquad 4\lambda^2 - 4\lambda + 1 - \frac{d}{d-1} - 1 = 0.$$

Thus, λ is given by,

$$(4.18) \qquad \lambda = \frac{1}{2} + \frac{1}{2}\sqrt{\frac{2d-1}{d-1}}.$$

We can now substitute this into the relation $a^2 + b^2 = 1$ to derive,

$$(4.19) \qquad b = \frac{1}{\sqrt{1 + 4\frac{d-1}{d}\lambda^2}}, a = \frac{2\lambda\sqrt{\frac{d-1}{d}}}{\sqrt{1 + 4\frac{d-1}{d}\lambda^2}}$$

The value of $\sqrt{\frac{d}{d-1}}ab + a^2$ is now given by,

$$(4.20) \qquad \frac{2\lambda\left[1 + 2\lambda\frac{d-1}{d}\right]}{1 + 4\frac{d-1}{d}\lambda^2}.$$

Now the equation for λ gives,

$$(4.21) \qquad 4\frac{d-1}{d}\lambda^2 + 1 = 4\frac{d-1}{d}\lambda + 2.$$

Thus, the value of $\sqrt{\frac{d}{d-1}}ab + a^2$ is now given by λ. This λ is our upper bound on the sum $\sum_{y \in \delta(\mathbb{T})}|P_{iy}|$.

$\square$

Now, with combinatorial estimates for $(P|_{\mathbb{T}})^{-1}$ in hand, we will be able to return to compute estimates on the difference $(G|_{\mathbb{T}})^{-1} - (P|_{\mathbb{T}})^{-1}$. This is encompassed in the following lemma,

Lemma 4.7. *Recall the setting of the previous Lemma.*
   Then

$$(4.22) \qquad |(G|_{\mathbb{T}})^{-1} - (P|_{\mathbb{T}})^{-1}|_{xy} \lesssim \epsilon.$$

Proof. Let E be the matrix of differences,

$$(4.23) \qquad \begin{aligned} (G|_{\mathbb{T}})^{-1} &= (P|_{\mathbb{T}})^{-1} + E \\ G|_{\mathbb{T}} &= P|_{\mathbb{T}} + W. \end{aligned}$$

By our inductive assumption, the entries of W are bounded by $W_{ab} \leq \epsilon(z, w)$ for all a, b found in $\mathbb{T}$.
   Multiplying these equations together give us,

$$(4.24) \qquad I = I + EP|_{\mathbb{T}} + (P|_{\mathbb{T}})^{-1}W + EW.$$

Manipulating this expression gives us,

$$(4.25) \qquad E = -(P|_{\mathbb{T}})^{-1}W(P|_{\mathbb{T}})^{-1} - EW(P|_{\mathbb{T}})^{-1}.$$

Considering the appearance of E on the right hand side, our hope is to use a maximum principle argument to show that the entries of E are bounded.
   Define $\Lambda := max_{c,d \in \mathbb{T}}|E_{cd}|$. Then, we see that,

$$(4.26) \qquad |[EW(P|_{\mathbb{T}})^{-1}]_{cd}| \leq \sum_{ef}|E_{ce}||W_{ef}||[(P|_{\mathbb{T}})^{-1}]_{fd}| \leq 2(2d-1)^l \Lambda\epsilon \sum_f |[(P|_{\mathbb{T}})^{-1}]_{fd}|.$$

Here, we used the fact that the number of possible indices e in $\mathbb{T}$ is $2(2d-1)^l$, which satisfies $(2d-1)^l \leq (\log N)^2$. Furthermore, one can use Lemma 4.4 to assert

$$(4.27) \qquad \sum_f |[(P|_{\mathbb{T}})^{-1}]_{fd}| \lesssim 1.$$

Thus, we can bound the second term as,

$$(4.28) \qquad |[EW(P|_{\mathbb{T}})^{-1}]_{cd}| \leq o(1)\Lambda.$$

Now, we can try to bound the first term, we see that,

$$|[(P|_{\mathbb{T}})^{-1}W(P|_{\mathbb{T}})^{-1}]_{cd}| \leq \sum_{ef} |[(P|_{\mathbb{T}})^{-1}]_{ce}||W_{ef}||[(P|_{\mathbb{T}})^{-1}]_{fd}|$$

(4.29)
$$\leq \epsilon(\sum_e |[(P|_{\mathbb{T}})^{-1}]_{ce}|)(\sum_f |[(P|_{\mathbb{T}})^{-1}]_{fd}|)$$

$$\lesssim \epsilon.$$

The last line is again based on Lemma 4.4.

Combining these estimates together, we see that,

(4.30)
$$\Lambda \leq o(1)\Lambda + C\epsilon.$$

This gives us our desired statement. $\qquad\square$

We now return to the proof of Proposition 2.15

Proof of Proposition 2.15. To compare $|G_{ij}^{(\mathbb{T})} - P_{ij}^{(\mathbb{T})}|$, we use equation (4.4) and deal with each term separately. We can bound $|G_{ij} - P_{ij}|$ by ϵ by our induction hypothesis. We need to perform a little bit more manipulation on our first term $(G - P)(G|_{\mathbb{T}})^{-1}G$.

We see that we have,

(4.31)
$$(G - P)(G|_{\mathbb{T}})^{-1}G = (G - P)(P|_{\mathbb{T}})^{-1}G + (G - P)[(G|_{\mathbb{T}})^{-1} - (P|_{\mathbb{T}})^{-1}]G.$$

We now bound the first term appearing above,

(4.32)
$$|[(G - P)(P|_{\mathbb{T}})^{-1}G]_{ij}| \leq \sum_{x,y \in \mathbb{T}} |[G - P]_{ix}||[(P|_{\mathbb{T}})^{-1}]_{xy}||G_{yj}|.$$

We can bound $[[G - P]_{ix}|$ by ϵ. If we fix y and sum over x, we observe that,

(4.33)
$$\sum_{x \in \mathbb{T}} |[(P|_{\mathbb{T}})^{-1}]_{xy}| \lesssim 1,$$

uniformly over x.

Thus, the preceding equation (4.32) is bounded by,

(4.34)
$$\epsilon \sum_{y \in \mathbb{T}} |G_{yj}| \leq \epsilon \sum_{y \in \mathbb{T}} [\epsilon + |P_{yj}|]$$

We have, from Lemma 4.5, that

$$\sum_{y \in \mathbb{T}} |P_{yj}| \leq (\log N)^{\mathfrak{j}(d)\mathfrak{a}}.$$

The term $(G - P)[(G|\mathbb{T})^{-1} - (P|\mathbb{T})^{-1}]G$ will give an error of the form $(\log N)^{8\mathfrak{a}}\epsilon^2$. This is of even smaller order than our proposed main error term.

Similar logic can be used to estimate the other terms remaining in (4.4). Ultimately, we can get our desired bound,

(4.35)
$$|G_{ij} - P_{ij}| \lesssim \epsilon'.$$

$\qquad\square$

5. Removal of the switched vertices W_S: Proof of Proposition 2.16

As one can see from the statement of Proposition 2.16, one needs many auxiliary notions for optimal estimates. As we will describe later, it is not good enough to simply use the worst case bound at all time. In general, most elements will be much smaller than the worst case bound ϵ'. It is this property which we exploit in what follows.

5.1. **Introduction of Green's function distance metric.** In the past section, we used ϵ as a control parameter to determine the maximum value of the distance of the Green's functions $|G_{ij} - P_{ij}| \leq \epsilon$. However, we find that there is a small difference between this worst case estimate and the average case estimate.

By the Ward Identity, we have that,

$$(5.1) \qquad \sum_j |G_{ij}^{(\mathbb{T})}|^2 = \frac{\text{Im}[G_{ii}^{(\mathbb{T})}]}{\eta}.$$

Thus, for any function $f(N)$, we see that no more than $\frac{N}{f(N)}$ indices j (for a fixed i) can satisfy the property that,

$$(5.2) \qquad |G_{ij}^{(\mathbb{T})}| \geq \frac{\sqrt{f(N)\text{Im}[G_{ii}^{(\mathbb{T})}]}}{\sqrt{N\eta}}.$$

By replacing $\text{Im}[G_{ii}]$ with the deterministic value m_∞ with small error, we see that the pigeonhole principle will assert the concrete estimate that,

$$(5.3) \qquad |G_{ij}^{(\mathbb{T})}| \geq \frac{\sqrt{f(N)\text{Im}[m_\infty]}}{\sqrt{N\eta}},$$

for no more than $2\frac{N}{f(N)}$ indices j (This merely used the fact that $\text{Im}[G_{ii}^{(\mathbb{T})}] \leq 2\text{Im}[m_\infty]$ from our inductive local law estimate on $\text{Im}[G_{ii}^{(\mathbb{T})}]$.)

It might look like this gives very little gain at first; however, we will soon see that improperly applied stability estimates would ruin the stability estimates by vastly increasing the prefactor of $(\log N)$ in front of the epsilon. Namely, instead of $(\log N)^{j(d)\mathfrak{a}}$, we may get $(\log N)^{C\mathfrak{a}}$ for very large C after our stability iteration. Unfortunately, the concentration estimate we apply later is not good enough to cancel out the increase of this multiplicative prefactor.

Remark 5.1. the concentration estimate which we will prove in Section 8 can kill a factor of say $(\log N)^{\mathfrak{a}}$. Thus, it unfortunately cannot kill a large multiplicative factor of the form $(\log N)^{K\mathfrak{a}}$. However, choosing $\mathfrak{a}$ large enough can kill any power of the form $(\log N)^p$ for some constant power p. The notions in this section allow us to reduce the decay in the stability estimate from $(\log N)^{K\mathfrak{a}}$ to $(\log N)^p$.

Definition 5.2. Fix the parameter

$$(5.4) \qquad \phi(z,w) := \frac{(\log N)^{\frac{K}{2}\mathfrak{a}}}{N^{1-\mathfrak{c}}} + \frac{(\log N)^{\frac{K}{2}\mathfrak{a}}\sqrt{\text{Im}[m_\infty]}}{\sqrt{N\eta}}.$$

The important part of ϕ is that $\phi \ll (\log N)^{-\frac{K}{2}\mathfrak{a}}\epsilon(z,w)$. We also use $(\epsilon')^2 \ll \phi$ many times for convenience of the final expression.

We say two indices i and j corresponding to vertices in $(\mathbb{T})^c$ are Green's function connected if there is any neighbor i' of i and a neighbor j' of j such that $|G_{i'j'}^{(\mathbb{T})}| \geq \phi$. We will use the notation $i \sim j$ if this is the case. Note that this is not an equivalence relation due to lack of transitivity.

We will use the fact that most Green's function would have the slightly better error estimate $|G_{ij}^{(\mathbb{T})} - P_{ij}^{(\mathbb{T})}| \leq \phi$ so that we do not need to worry as much about the decay of the multiplicative factor. In order to see this, we prove the following lemma, which shows how most vertices in $\mathbb{T}^c$ cannot be Green's function connected to too many (more than $O(\log N)$) of the vertices that we want to switch.

Lemma 5.3. *Fix our random dregular digraph $\mathcal{G}$ as well as a vertex v whose local $\mathfrak{R}$-like neighborhood is treelike. Consider the local l neighborhood $\mathbb{T}$ around the vertex v. Let $e_1 = (v_1, w_1), \ldots, e_\mu = (v_\mu, w_\mu)$ be randomly selected edges that lie outside $\mathbb{T}$ in $\mathcal{G}$ and μ is the number of edges connecting $\mathbb{T}$ to $\mathbb{T}^c$. Then, with probability $1 - O(N^{-\mathfrak{q}})$, we have the following two conditions:*

(1) *Any index v in $\mathbb{T}^c$ is Green's function connected to no more than $O(\log N)$ of the indices corresponding to vertices $\{b_1, c_1, \ldots, b_\mu, c_\mu\}$.*

(2) The set α such that $\{b_\alpha, c_\alpha\}$ is Green's function connected to some $\{b_\beta, c_\beta\}$ for $\beta \neq \alpha$ is $O(\log N)$.

Any vertex v in $\mathbb{T}^c$ is Green's function connected to no more than $O(\log N)$ vertices in $\cup_{i=1}^m \{v_i, w_i\}$.

Proof. Let X be any vertex in $\mathbb{T}^c$. For the vertex X, the set X_{GF} of the vertices that are Green's function connected to X is of size less than $\frac{N}{(\log N)^{K\mathfrak{q}}}$. For each i in $\{1, \ldots, m\}$, the chance that X is Green's function connected to one of v_i, w_i is an independent event. At this point, one can obtain the desired result by straightforward counting arguments. One can see [18, Prop. 5.14] for details. □

At this point, we can finally define the event $S_G(\mathcal{G})$ of good switching events.

Definition 5.4. Consider a graph $\mathcal{G}$ in $\Omega_g(z, w)$ as well as a distinguished vertex o along with its radius ℓ neighborhood $\mathbb{T}$. Recall $\mathbb{S}_{\mathcal{G}, o}$, the switching events that correspond to the graph $\mathcal{G}$ around vertex o, as well as the characteristic variables χ_α. We define the event $S_G(\mathcal{G}) \subset \mathbb{S}_\mathcal{G}$ as follows: a switch S will belong to $S_G(\mathcal{G})$ if it satisfies the following conditions,

(1) All except for $O_\mathfrak{q}(1)$ of the vertices $\{c_1, \ldots, c_\mu\}$ have radius $\mathfrak{R}$ tree neighborhoods in $\mathcal{G}^{(\mathbb{T})}$.
(2) $\chi_\alpha = 1$ for all except $O_\mathfrak{q}(1)$ many vertices.
(3) The Green's function connectivity conditions of Lemma 5.3 hold.

Due to Lemmas 5.3 and Proposition 2.6, the event $S_G(\mathcal{G})$ holds with probability $\mathbb{P}(S_G(\mathcal{G})) = 1 - O(N^{-\mathfrak{q}})$.

In what follows in Sections 5 ,6, and 7, we will always implicitly assume we are considering a switching event S found in $S_G(\mathcal{G})$.

5.2. **Switching estimates involving the Green's function distance.** We can apply a very similar type of analysis for the removal of the vertices in W_S. Given a switching event S, where we elect to switch the vertices on the boundary of $\mathbb{T}$ with arbitrary edges, the vertices W_S consist of those vertices $\{b_1, \ldots, b_\mu\}$. Again, we introduce, as a shorthand, the matrix $P := G(Ext(B_r(\{i, j\} \cup W_S, \mathcal{G} \backslash \mathbb{T}), Q_I, Q_O))$. As before, the point of this matrix is that we can write $G^{(\mathbb{T})}$ as,

$$(5.5) \qquad (G^{(\mathbb{T})})^{-1} = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix},$$

and

$$(5.6) \qquad P^{-1} = \begin{bmatrix} A & B \\ B^* & D' \end{bmatrix}.$$

Here, A is the adjacency matrix of the graph restricted to W_S and B will be the adjacency matrix of the edges connecting W_S to W_S^c. At this point, we remark that we can restrict to the event that all the edges that are chosen for switching are far away from each other; more specifically, two vertices in b_α, b_β in W_S are of distance at least $2r$ from each other.

Due to this restriction, the matrix $P|_{W_S}$ is much easier to understand. It is a block diagonal matrix (with 2×2 blocks) where the blocks consist of the indices b_α and $b_\alpha + N$ of the vertices that border the edge used in the switching. For this reason, it is also easy to describe the matrix $(P|_{W_S})^{-1}$.

Another remark that we will have to make is that in addition to proving the stability of the worst case bound $|G_{ij}^{(T)} - P_{ij}^{(T)}| \leq \epsilon$, we would further need to prove the stability of the average case bound for pairs of vertices that are not Green's function connected. With this remark in mind, we now return to the proof.

By applying the Schur complement formula, we have that,

$$(5.7) \qquad G^{(\mathbb{T} \cup W_S)} = G^{(\mathbb{T})} - G^{(\mathbb{T})}(G^{(\mathbb{T})}|_{W_S})^{-1}G^{(\mathbb{T})}, P^{(W_S)} = P - P(P|_{W_S})^{-1}P.$$

We can explicitly write out our difference as,

$$(5.8) \qquad \begin{aligned} G^{(\mathbb{T} \cup W_S)} - P^{(W_S)} &= [G^{(\mathbb{T})} - P] - [G^{(\mathbb{T})} - P](G^{(\mathbb{T})}|_{W_S})^{-1}G^{(\mathbb{T})} \\ &\quad - P[(G^{(\mathbb{T})}|_{W_S})^{-1} - (P|_{W_S})^{-1}]G^{(\mathbb{T})} - P(P|_{W_S})^{-1}[G^{(\mathbb{T})} - P]. \end{aligned}$$

As before, one main issue would be to try to estimate the value of $(G^{(\mathbb{T})}|_{W_S})^{-1}$

Lemma 5.5. *We have the following estimates on*

$$\left|(P|_{W_S})_{ij}^{-1} - (G^{(\mathbb{T})}|_{W_S})_{ij}^{-1}\right| \lesssim \epsilon'.$$

Furthermore, if we know that i and j are not Green's function connected, then we are able to get the superior estimate,

(5.9) $$|(P|_{W_S})_{ij}^{-1} - (G^{(\mathbb{T})}|_{W_S})_{ij}^{-1}| \lesssim \phi,$$

for some constant C. We remark here that if i and j are not Green's function connected, then they would belong to different blocks of $P|_{W_S}$. In this case, $[P|_{W_S}]_{ij}$ and $(P|_{W_S})_{ij}^{-1}$ are both zero.

Proof. Again, we write,

(5.10)
$$(G^{(\mathbb{T})}|_{W_S})^{-1} = (P|_{W_S})^{-1} + E$$
$$G^{(\mathbb{T})}|_{W_S} = P|_{W_S} + J.$$

Multiplying these equations, we again can derive that,

(5.11) $$E = -(P|_{W_S})^{-1} J (P|_{W_S})^{-1} - E J (P|_{W_S})^{-1}.$$

Let us estimate the first term. We have,

(5.12) $$|[(P|_{W_S})^{-1} J (P|_{W_S})^{-1}]_{ij}| \leq \sum_{a,b \in W_S} |(P|_{W_S})_{ia}^{-1}||J_{ab}||(P|_{W_S})_{bj}^{-1}|.$$

We can bound $|J_{ab}| \leq \epsilon'$. Furthermore, observe that there are at most 2 values a such that $(P|_{W_S})_{ia}^{-1}$ is non-zero and 2 values b such that $(P|_{W_S})_{bj}^{-1}$; this is due to our assumption that the switched edges in W_S are far apart from one another. Furthermore, these values are $O(1)$ quantities; this can be shown be a direct application of the Schur complement formula on a pure tree-like neighborhood combined with some perturbation computations.

Let us define

$$P_\infty := G(Ext(B_r(\{i,j\} \cup W_S, \mathcal{G}\backslash\mathbb{T}), m_\infty, m_\infty)).$$

By Lemma 3.5, one can show that

$$||P|_{W_S} - P_\infty|_{W_S}||_\infty \ll 1.$$

By ∞ norm here, we refer to the supremum over all entries. Furthermore, since the vertices in W_S are chosen so that they have tree-like neighborhoods, we additionally have that $(P_\infty|_{W_S})^{-1}$ is explicitly given by

$$\begin{bmatrix} -z - \frac{d}{d-1}m_\infty & w \\ \overline{w} & -z - \frac{d}{d-1}m_\infty \end{bmatrix},$$

on each of its diagonal blocks. Clearly, all of these terms are $O(1)$. By the same analysis we have done earlier, we can argue that $||(P|_{W_S})^{-1} - (P_\infty|_{W_S})^{-1}||_\infty \ll 1$; this is especially simple since we only have to consider only 2×2 diagonal blocks. This shows that the entries of $(P_\infty|_{W_S})^{-1}$ are of $O(1)$, as desired.

Thus, the sum from equation (5.12) can be bounded by $C|J_{ab}|$ for the unique possible values of a and b (if they exist), that return a non-zero quantity in the computation $|(P|_{W_S})_{ia}^{-1}||J_{ab}||(P|_{W_S})_{bj}^{-1}|$. Since we have earlier shown that $|J_{ab}| \leq \epsilon'$ for any pair a and b, we know that (5.12) $\leq C\epsilon$.

Now, we estimate the second term from (5.11),

(5.13) $$|\sum_{a,b \in W_S} E_{ia} J_{ab} (P|_{W_S})_{bj}^{-1}| \leq 2(\sup_{a,b \in W_S} E_{ab})|W_S|\epsilon'.$$

We applied the triangle inequality and used the fact that $(P|_{W_S})_{bj}^{-1}$ is zero for all values of b except for 2 of them. We can bound E_{ia} by the supremum over all values and J_{ab} by ϵ'. Now, $|W_S|\epsilon'(z,w) \ll 1$. Combining this inequality and the previous one, we can derive the inequality,

(5.14) $$\sup_{ab} E_{ab} \leq (\sup_{ab} E_{ab})|W_S|\epsilon' + \epsilon'.$$

This would imply that $\sup_{a,b \in W_S} E_{ab} \leq 2\epsilon'$ as desired.

With the estimate on the supremum in hand, we can improve our estimate for the difference on terms that are not Green's function connected. Observe, that we can bound

$$(5.15) \qquad |\sum_{a,b} E_{ia} J_{ab} (P|_{W_S})^{-1}_{bj}| \leq (\sup_{ab} E_{ab})|W_S|\epsilon' \leq (\log N)^{4\mathfrak{a}}(\epsilon')^2 \ll \phi.$$

Furthermore, we have the following improvement when considering the first term from equation (5.11).

$$(5.16) \qquad \sum_{a,b \in W_S} |(P|_{W_S})^{-1}_{ia}||J_{ab}||(P|_{W_S})^{-1}_{bj}|,$$

Let i correspond to one of $b_\alpha, b_\alpha + N$ and and j correspond to one of $b_\beta, b_\beta + N$, where b_α and b_β are vertices in W_S. The only non-zero terms $[(P|_{W_S})^{-1}]_{ia}$ has $a = b_\alpha$ or $b_\alpha + N$. Similarly, the only nonzero terms $(P|_{W_S})^{-1}_{bj}$ has $b = b_\beta$ or $b_\beta + N$. For these indices, we can apply the bound $|J_{ij}| \leq \phi$ due to our assumption that i and j are not Green's function connected. $\qquad \square$

With this inequality in hand, we can start estimating the terms that appear in equation (5.8). This will complete the proof of Proposition 2.16.

Proof of Proposition 2.16. The simplest term to estimate is the last one in equation (5.8).

Consider the following expression,

$$(5.17) \qquad [P(P|_{W_S})^{-1}[G^{(\mathbb{T})} - P]]_{ij} = \sum_{a,b \in W_S} |P_{ia}||(P|_{W_S})^{-1}_{ab}||[G^{(\mathbb{T})} - P]_{bj}|.$$

Let i correspond to one of $b_\alpha, b_\alpha + N$ and j correspond to one of $b_\beta, b_\beta + N$ for $b_\alpha, b_\beta \in W_S$.

As we have argued before, due to the block diagonal structure of P, P_{ia} is nonzero for at most 2 values of a. These values can be b_α or $b_\alpha + N$. Given this value of a, $(P|_{W_S})^{-1}_{ab}$ is nonzero for at most 2 values of b. This b can still be only one of b_α or $b_\alpha + N$. Finally, we can apply the bound that $|[G^{(\mathbb{T})} - P]_{bj}| \leq \epsilon'$ in general, or $\leq \phi$ if i and j are not Green's function connected.

Now, let us consider bounding the second term in equation (5.8). This is,

$$(5.18) \qquad \sum_{a,b \in W_S} |P_{ia}||[(P|_{W_S})^{-1} - (G^{(\mathbb{T})}|_{W_S})^{-1}]_{ab}||G^{(\mathbb{T})}|_{bj}.$$

We first remark that since we know that $|G^{(\mathbb{T})}_{bj} - P_{bj}| \leq \epsilon'$, this shows that $|G^{(\mathbb{T})}|$ will have the worst case bound $O(1)$ for at most 2 values of the index b (those that are in the same block as j in P). There are a further $O(\log N)$ indices of b where we would need to apply the worst case bound $|G^{(\mathbb{T})}|_{bj} \leq \epsilon'$ and $P_{bj} = 0$. Finally, we can bound the Green's function for all remaining indices that are possibilities for b as $|G^{(\mathbb{T})}|_{bj} \leq \phi$, due to our assumption that the indices are not Green's function connected.

Further, we also know that for each fixed index a, there are at most $O(\log N)$ indices where we would need to apply the worst case bound $|[(P|_{W_S})^{-1} - (G^{(\mathbb{T})}|_{W_S})^{-1}]_{ab}| \leq \epsilon'$. These are the indices that a is Green's function connected to b. Otherwise, we could apply the bound ϕ. As before, $|P_{ia}|$ is 0 for all except for 2 possible indices a.

Now, consider the case that i is not Green's function connected to j. Then, if b is one of the two indices connected to j in W_S, then we would have the improved bound $|[(P|_{W_S})^{-1} - (G^{(\mathbb{T})}|_{W_S})^{-1}]_{ab}| \leq \phi$ for these indices b and all indices a such that $|P_{ia}| \neq 0$.

Thus, we would derive the bound,

$$\sum_{a,b \in W_S} |P_{ia}||[(P|_{W_S})^{-1} - (G^{(\mathbb{T})}|_{W_S})^{-1}]_{ab}||G^{(\mathbb{T})}|_{bj}$$
$$\lesssim \phi + \log N(\epsilon')^2 + \log N|W_S|\epsilon'\phi + |W_S|^2\phi.$$

Similar techniques allow us to treat the first term in equation (5.8). $\qquad \square$

6. Relating $G^{(\mathbb{T} \cup W_S)}$ to $\tilde{G}^{(\mathbb{T})}$ of the switched graph: Proof of Proposition 2.17

As in the previous sections, we will prove Proposition 2.16 by applying the Schur complement formula and other Green's function manipulations. Since we are adding vertices back into the graph, we will have slightly different formulas for $\tilde{G}_{ij}^{(\mathbb{T})}$ from $G_{ij}^{(\mathbb{T} \cup W_S)}$ depending on whether i or j is in W_S or not. The full conclusion of Proposition 2.16 can be derived from the conclusions of Lemmas 6.1, 6.3, and 6.4.

We consider P given by $G(Ext(B_r(\{i,j\} \cup W_S, \tilde{\mathcal{G}} \backslash \mathbb{T}), Q_I, Q_O))$ where $\tilde{\mathcal{G}}$ is the graph obtained by switching the edges associated to the switching set W_S. We observed that $P^{(W_S)} = G(Ext(B_r(\{i,j\} \cup (\mathbb{T} \cup W_S), \mathcal{G} \backslash \mathbb{T}), Q_I, Q_O))$. Namely, $P^{(W_S)}$ is exactly the Green's function we were relating G to in the previous section.

In this section we will apply the Schur complement formula in the form,

$$
\begin{aligned}
\tilde{G}^{(\mathbb{T})}|_{W_S} &= (A - \tilde{B} G^{(\mathbb{T} \cup W_S)} \tilde{B}^*)^{-1}, \\
P|_{W_S} &= (A - \tilde{B} P^{(W_S)} \tilde{B}^*)^{-1},
\end{aligned}
\tag{6.1}
$$

where $\tilde{B}$ represents the adjacency matrix between the vertices of W_S to those outside of W_S.

The Schur complement formula give us

$$
\begin{aligned}
\tilde{G}^{(\mathbb{T})}|_{W_S \times W_S^c} &= -\tilde{G}^{(\mathbb{T})}|_{W_S} \tilde{B} G^{(\mathbb{T} \cup W_S)} \\
P|_{W_S \times W_S^c} &= -P|_{W_S} \tilde{B} P^{(W_S)},
\end{aligned}
\tag{6.2}
$$

and an application of the resolvent identity would give us,

$$
\begin{aligned}
\tilde{G}^{(\mathbb{T})}|_{W_S^c \times W_S^c} &= G^{(\mathbb{T} \cup W_S)} + G^{(\mathbb{T} \cup W_S)} \tilde{B}^* \tilde{G}^{(\mathbb{T})}|_{W_S} \tilde{B} G^{(\mathbb{T} \cup W_S)} \\
P &= P^{(W_S)} + P^{(W_S)} \tilde{B}^* P|_{W_S} \tilde{B} P^{(W_S)}.
\end{aligned}
\tag{6.3}
$$

We see first that it is crucial to understand the difference of $\tilde{G}^{(\mathbb{T})}$ and P on the set W_S at the first step. This can be done by the resolvent identity,

$$
\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S} = \tilde{G}^{(\mathbb{T})}|_{W_S} \tilde{B} (P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}) \tilde{B}^* P|_{W_S}.
\tag{6.4}
$$

We can estimate the maximum difference of the right hand side.

Lemma 6.1. *For any two indices i and j corresponding to vertices in W_S, we have,*

$$
|[\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}]_{ij}| \lesssim \epsilon'.
\tag{6.5}
$$

Furthermore, if i and j correspond to vertices that are not Green's function connected, then we have the improved bound,

$$
|[\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}]_{ij}| \lesssim \phi.
\tag{6.6}
$$

Proof. Let $\Gamma := \max_{i,j \in W_S} |[\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}]_{ij}|$. Using this definition inside (6.4), we would derive the relation,

$$
\begin{aligned}
|[\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}]_{ij}| &= ([\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}] \tilde{B} (P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}) \tilde{B}^* P|_{W_S})_{ij} \\
&\quad + (P|_{W_S} \tilde{B} (P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}) \tilde{B}^* P|_{W_S})_{ij}.
\end{aligned}
\tag{6.7}
$$

Let i correspond to one of the indices b_α or $b_\alpha + N$ and let j correspond to one of the indices b_β or $b_\beta + N$, where b_α and b_β are vertices that belong to W_S

The first quantity on the right hand size can be bounded by $(\log N)^{8\mathfrak{a}} \epsilon' \Gamma \ll \Gamma$. We write the second term on the right hand side as,

$$
\begin{aligned}
&(P|_{W_S} \tilde{B} (P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}) \tilde{B}^* P|_{W_S})_{ij} \\
&= \sum_{x,y,z,w} (P|_{W_S})_{ix} \tilde{B}_{xy} (P^{(W_S)} - G^{(\mathbb{T} \cup W_S)})_{yz} (\tilde{B}^*)_{zw} (P|_{W_S})_{wj}.
\end{aligned}
\tag{6.8}
$$

For the second quantity on the right hand size, we remark that $[P|_{W_S}]_{ix} = 1$ only if x corresponds to one of the indices b_α or $b_\alpha + N$. This is due to the fact that b_α and b_β for $\alpha \neq \beta$ are still far apart even after b_α is connected to l_α for all α in W_S. Thus, $P|_{W_S}$ is still a block-diagonal matrix. Similarly, w must be one of b_β or $b_\beta + N$ if $(P|_{W_S})$ were to be non-zero. Since B and B' are the adjacency matrices for d-regular graphs, this would further imply that there are only finitely many values that y could take and finitely many values that z could take. By our choice of the construction, we would know that these y vertices and these z vertices would all have a local treelike neighborhood.

Since we have the bound $|P_{cd}^{(W_S)} - G_{cd}^{(\mathbb{T} \cup W_S)}| \leq 2\epsilon'(z, w)$ for any generic pair c, d that has a tree-like neighborhood, we would apply this bound everywhere and observe that,

$$\tag{6.9} \sum_{x,y,z,w} (P|_{W_S})_{ix} \tilde{B}_{xy} (P^{(W_S)} - G^{(\mathbb{T} \cup W_S)})_{yz} (\tilde{B}^*)_{zw} (P|_{W_S})_{wj} \lesssim C\epsilon'.$$

From this point, a standard maximum principle argument completes the the proof of the bound on $P - G$ for general i and j.

If i and j are not Green's function connected, the only observation we need is that we can improve the bound on $|P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}|$ in equation (6.9) to ϕ. The first term in equation (6.7) can be bounded by $\epsilon'^2 \ll \phi$. □

We can apply these estimates in order to understand the differences of $\tilde{G}^{(\mathbb{T})}$ and P on the rectangular sub-matrix $W_S \times W_S^c$. The following lemma encompasses this information. However, we mention that some of these estimates can improve if we instead consider some Green's function connected information on the specific vertices $\{c_\alpha, b_\alpha\}$ that take part in the switching. We introduce a definition to encompass this information.

Definition 6.2. We define a new set U of switched vertices b_α, c_α that satisfy some specific properties. The vertices b_α, c_α are included in the set U if the following properties are satisfied:

(1) For $\beta \neq \alpha$, the set of vertices $\{b_\beta, c_\beta\}$ is not Green's function connected to $\{b_\alpha, c_\alpha\}$.
(2) The $\mathfrak{R}$ neighborhood around c_α in $\mathcal{G} \backslash \mathbb{T}$ is tree-like.

Lemma 6.3. *Let i and j be two vertices: i belonging to W_S and j belonging to W_S^c. Then, we have the following comparison bound,*

$$\tag{6.10} |\tilde{G}_{ij}^{(\mathbb{T})} - P_{ij}| \lesssim \epsilon'.$$

Recall our notation $\{a_\alpha, b_\alpha, c_\alpha\}$ for the types of vertices that participate in the switching. Assume that j is of the form c_α for c_α a vertex in U and that $i = b_\beta$ for $\beta \neq \alpha$.

$$\tag{6.11} |\tilde{G}_{b_\beta c_\alpha}^{(\mathbb{T})} - P_{b_\beta c_\alpha}| \lesssim \phi + (\epsilon')^2.$$

Proof. General Case:
We start with the identity,

$$\tag{6.12} \tilde{G}^{(\mathbb{T})}|_{W_S \times W_S^c} - P|_{W_S \times W_S^c} = [P|_{W_S} - \tilde{G}^{(\mathbb{T})}|_{W_S}] \tilde{B} G^{(\mathbb{T} \cup W_S)} + P|_{W_S} \tilde{B} [P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}].$$

Let us first deal with the second term $(P|_{W_S} \tilde{B} [P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}])_{ij}$. We have,

$$\tag{6.13} (P|_{W_S} \tilde{B} [P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}])_{ij} = \sum_{x,y} (P|_{W_S})_{ix} \tilde{B}_{xy} [P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}]_{yj}.$$

Now, $(P|_{W_S})_{ix}$ can only be non-zero only when x is one of the indices corresponding to the same vertex as i. Since y is connected to x, there are finitely many choices for y and all of these y vertices have tree-like neighborhoods. We can finally apply the bound that $[P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}]_{yj}$ is less than $2\epsilon'$.

We can deal with the first term similarly, by first writing,

$$\tag{6.14} [P|_{W_S} - \tilde{G}^{(\mathbb{T})}|_{W_S}] \tilde{B} G^{(\mathbb{T} \cup W_S)} = [P|_{W_S} - \tilde{G}^{(\mathbb{T})}|_{W_S}] \tilde{B} P^{(W_S)} + [P|_{W_S} - \tilde{G}^{(\mathbb{T})}|_{W_S}] \tilde{B} [G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}].$$

The second term is quadratic in the error $(\epsilon')^2$ and can compensate for any factor of the form $(\log N)^{4\mathfrak{a}}$ that might appear. The first term above can be dealt with the exact same way as the previous bound we considered in this proof.

Specific Case of $i = c_\alpha$ and $j = b_\beta$ with $\beta \neq \alpha$:

We can apply the same identity detailed in (6.12). We do not need to consider the contribution of terms with two factors of $G - P$ (as these will give an error of the form $(\epsilon')^2$.) There are two terms remaining using only a single factor of the form $G - P$ and the analysis of both of these terms is very similar.

Let us consider the term,

$$(6.15) \qquad (P|_{W_S})_{b_\beta x} \tilde{B}_{xy} [P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}]_{y c_\alpha}.$$

We abuse notation slightly as we let b_β, c_α refer to any of the two indices that correspond to the vertex b_β, c_α rather than the vertex itself. Since $P|_{W_S}$ is diagonal, x again has to be an index corresponding to b_β. In addition, y has to correspond to a vertex adjacent to b_β , since $\tilde{B}$ is an adjacency matrix. Finally, in the last term, we use the fact that $y = b_\beta$ and c_α are no longer Green's function connected to bound $[P^{(W_S)} - G^{(\mathbb{T} \cup W_S)}]_{y c_\alpha}$ by ϕ. This shows that the above term is bounded by a constant multiple of ϕ, as desired. $\qquad\square$

At this point, we can finally continue and prove bounds of the difference when the indices i and j are both in W_S^c. Furthermore, we can prove superior bounds on some Green's function entries if we take into account the Green's function connectivity properties.

Lemma 6.4. *Let i and j be vertices of W_S^c in $\tilde{\mathcal{G}} \backslash \mathbb{T}$. We have the following general estimate on Green's function bounds,*

$$(6.16) \qquad |\tilde{G}_{ij}^{(\mathbb{T})} - P_{ij}| \lesssim \epsilon'.$$

Recall the notation $c_\alpha, b_\alpha, a_\alpha$ for the vertices that participate in the switching. Let c_α be a vertex in U and let j be a vertex that is not Green's function connected to $\{c_\alpha, b_\alpha\}$ in $\mathcal{G} \backslash \mathbb{T}$. Then, we have the bound,

$$(6.17) \qquad |\tilde{G}_{ij}^{(\mathbb{T})} - P_{ij}| \lesssim \phi.$$

Proof. General Case:

We see we can write the difference between $\tilde{G}^{(\mathbb{T})}$ and P on $W_S^c \times W_S^c$ as,

$$
\begin{aligned}
&G^{(\mathbb{T} \cup W_S)} - P^{(W_S)} + (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}) \tilde{B}^* P|_{W_S} \tilde{B} P^{(W_S)} \\
&+ P^{(W_S)} \tilde{B}^* (\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}) \tilde{B} P^{(W_S)} + P^{(W_S)} \tilde{B}^* P|_{W_S} \tilde{B} (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}) \\
&+ (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}) \tilde{B}^* (\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}) \tilde{B} P^{(W_S)} \\
(6.18) \quad &+ (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}) \tilde{B}^* P|_{W_S} \tilde{B} (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}) \\
&+ (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}) \tilde{B}^* (\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}) \tilde{B} (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}) \\
&+ P^{(W_S)} \tilde{B}^* (\tilde{G}^{(\mathbb{T})}|_{W_S} - P|_{W_S}) \tilde{B} (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)})
\end{aligned}
$$

The terms that involve two differences of $G - P$ would have an error of at most quadratic order in ϵ'^2. This is already much smaller than ϵ', even with prefactors of the form $(\log N)^{4\mathfrak{a}}$. To derive the worst case bound that is of order ϵ', it suffices to understand the term $[(G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}) \tilde{B}^* P|_{W_S} \tilde{B} P^{(W_S)}]_{ij}$ (since the other terms are similar). We will write this term as,

$$(6.19) \qquad \sum_{x,y,z,w} (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)})_{ix} (\tilde{B}^*)_{xy} (P|_{W_S})_{yz} (\tilde{B})_{zw} (P^{(W_S)})_{wj}.$$

We remark first that since $P^{(W_S)}$ is a block diagonal matrix w must belong to the same diagonal block as j. Furthermore, $\tilde{B}$ is an adjacency matrix connecting w to some element z in W_S. In the general case, this must imply that z is the unique element of W_S that is adjacent to the block of W_S^c containing j. There are then at most $2d$ choices of w (these are indices corresponding to vertices adjacent to z in W_S). As $P|_{W_S}$ is a block diagonal matrix matrix, with each block corresponding to the two indices mapped to a vertex, there

are only 2 choices for y. Finally, there are at most $2d$ choices of x since $\tilde{B}$ is an adjacency matrix. Thus, we have a finite number of choice for x, y, z and w that give nonzero values. We can bound $(G^{(\mathbb{T} \cup W_S)} - P^{(W_S)})_{ix}$ by ϵ' for every x in this finite set. In general, we see this term is bounded by $C\epsilon'$. The analysis of terms involving only a single factor of $G - P$ is similar. Other terms involving two $G - P$'s will be squared and have decay at least $(\epsilon')^2$.

Specific case of $i = c_\alpha$ and j being Green's function separated

Now, we give the superior analysis in the case that i is the vertex c_α for some α. These superior estimates are necessary when we reintroduce the tree $\mathbb{T}$ into the switched graph $\tilde{\mathcal{G}} \backslash \mathbb{T}$.

Let us first consider the case that j is not Green's function connected to the vertex c_α and that the distance of j to a_α is greater than r in $\tilde{\mathcal{G}} \backslash \mathbb{T}$. We use the same expansion detailed in equation (6.18). As before, we mention that those terms that involve two factors of $G - P$ will incur a factor of $(\epsilon')^2 \ll \phi$. This means that we will only have to deal with the contributions from those terms involving a single $G - P$ factor. It would suffice as before to deal with one such term; the analysis from the contribution of the other terms would be very similar.

Let us consider the term,

$$(6.20) \qquad \sum_{x,y,z,w} (G^{(\mathbb{T} \cup W_S)} - P^{(W_S)})_{c_\alpha x} (\tilde{B}^*)_{xy} (P|_{W_S})_{yz} (\tilde{B})_{zw} (P^{(W_S)})_{wj}.$$

Applying the same logic as before , we know that there is a single choice for z, the single vertex in W_S that is adjacent to the block of W_S^c containing j. This vertex $z = b_\beta$ cannot be b_α, since otherwise j would be Green's function connected to b_α. Thus, $\beta \neq \alpha.y$ must also be an index corresponding to this vertex b_β and x is an index corresponding to a vertex adjacent to b_β. Finally, we can use the fact that c_α and b_β are not Green's function connected to bound $|G^{(\mathbb{T} \cup W_S)} - P^{(W_S)}|$ by ϕ. This shows that the above term considered is necessarily less than some multiple of ϕ, as desired.

$\square$

7. Reintroduction of the tree and relating $\tilde{G}^{(\mathbb{T})}$ to $\tilde{G}$:Proof of Proposition 2.18

As we have done previously, our basic step is to apply the Schur complement formula and resolvent estimates in order to treat the removal of the subset $\mathbb{T}$. We make one important remark before proceeding. At the first step, it is easier to compare $\tilde{G}$ with extensions involving the parameters Q_I and Q_O for the original graph $\mathcal{G}$. Lemmas 7.1 and 7.3 do exactly this. It is only later in Lemma 7.4 where we consider the difference between Q_I and $\tilde{Q}_I$. A simple perturbation bound as from Theorem 1.13 will complete the proof.

We will decompose the adjacency matrix as,

$$(7.1) \qquad \tilde{G}^{-1} = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix},$$

where A will be the adjacency matrix of the neighborhood $\mathbb{T}$ and B contains the adjacency information of the boundary of $\mathbb{T}$ and the vertices they were switched with.

As we have done previously, we will relate the element $\tilde{G}_{ij}$ to the term P_{ij}, where P is defined to be $G(\text{Ext}(B_r(\{i,j\} \cup \mathbb{T}, \tilde{\mathcal{G}}), Q_I, Q_O))$. We see that P^{-1} has the adjacency matrix,

$$(7.2) \qquad P^{-1} = \begin{bmatrix} A & B \\ B^* & D' \end{bmatrix}.$$

The inverse of the matrix D' (will be $G(\text{Ext}(B_r(\{i,j\} \cup \mathbb{T}, \tilde{\mathcal{G}}^{(\mathbb{T})}), Q_I, Q_O))$, which we have analyzed in the previous section. Here, $B_r(\mathbb{T})$ should be understood as the r neighborhood of $\mathbb{T}$ in $\mathcal{G}$ restricted to $\mathbb{T}^c$.

For vertices i and j inside $\mathbb{T}$, we can apply the Schur complement formula and derive,

$$(7.3) \qquad \begin{aligned} \tilde{G}|_{(\mathbb{T})} &= (A + B\tilde{G}^{(\mathbb{T})}B^*)^{-1} \\ P|_{(\mathbb{T})} &= (A + BP^{(\mathbb{T})}B^*)^{-1}. \end{aligned}$$

We can now use the resolvent formula to compute the difference,

$$(7.4) \qquad \tilde{G}|_{(\mathbb{T})} - P|_{(\mathbb{T})} = P|_{(\mathbb{T})}B(\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})B^*P + (\tilde{G}|_{(\mathbb{T})} - P|_{(\mathbb{T})})B(\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})B^*.$$

Our formal statement on the value of the difference is as follows.

Lemma 7.1. *Let i and j be two points found in our neighborhood $\mathbb{T}$. Let d_i denote the distance from i to the root and d_j denote the distance between j and the root. We have the following estimates,*

$$(7.5) \qquad |\tilde{G}|_{(\mathbb{T})} - P|_{(\mathbb{T})}| \lesssim \epsilon'(\log N)^2 \Big(\frac{1}{\sqrt{d-1}}\Big)^{2\ell - d_i - d_j} + (1 + |d_j + d_i|)\Big(\frac{1}{d-1}\Big)^{\frac{|d_j - d_i|}{2}} \epsilon'.$$

Remark 7.2. The second error term on the right hand side of (7.5) has the issue that it does not have a decay of the power $\big(\frac{1}{d-1}\big)^{\ell}$. By contrast, the first error term on the right hand side of (7.5) has the power of $(d-1)^{-2\ell}$, which can suppress powers of log as long as we ensure that d_i and d_j are small relative to ℓ.

Proof. The computations of the previous section show that entries of $(\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})$ are all bounded by $C\epsilon'$ and a few other terms have superior bounds.

We have for (7.4), that

$$(PB(\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})B^*P)_{ij} \leq \frac{1}{d-1}\sum_{l_\alpha, l_\beta} P_{i, l_\alpha}(\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})_{\tilde{c}_\alpha, \tilde{c}_\beta} P_{l_\beta, j}$$

$$(7.6) \qquad\qquad\qquad \leq \frac{1}{d-1}\sum_{\alpha} P_{i, l_\alpha} P_{l_\alpha, j}(\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})_{\tilde{c}_\alpha, \tilde{c}_\alpha}$$

$$+ \frac{1}{d-1}\sum_{\alpha \neq \beta} P_{i, l_\alpha} P_{l_\beta, j}(\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})_{\tilde{c}_\alpha, \tilde{c}_\beta}.$$

α enumerates the vertices on the boundary of $(\mathbb{T})$. In a slight abuse of notation, l_α and c_α actually represent the possible indices corresponding to the vertices on the boundary rather than the vertices themselves; we hope that this usage of notation is clear in context. The sum is split into two parts that will have different estimates based on whether α is equal to β or not. When $\alpha \neq \beta$ we will be able to get superior estimates when we use the separation of Green's function distance.

Let us now treat the second term on which $\alpha \neq \beta$. This can be divided even further based on whether we are considering an index from the set U from Definition 6.2. For α corresponding to a vertex in U, we can use the fact that $|(\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})_{\tilde{c}_\alpha, \tilde{c}_\beta}| \leq \phi$. There are $O((\log N)^2)$ that have both of their entries from outside U. Furthermore, we can use the bound $|P_{i, l_\alpha}| \leq \big(\frac{1}{\sqrt{d-1}}\big)^{\ell - d_i}$. Applying this, we see that we have the following bound on our cross terms,

$$(7.7) \qquad \frac{\phi}{d-1}\sum_{\alpha \neq \beta} |P_{i, l_\alpha}||P_{l_\beta, j}| + \frac{\epsilon'(\log N)^2}{d-1}\Big(\frac{1}{\sqrt{d-1}}\Big)^{\ell - d_i}\Big(\frac{1}{\sqrt{d-1}}\Big)^{\ell - d_j}$$

For the terms that are not cross terms (with $\alpha = \beta$), we merely apply the bound that $|\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})}| \leq \epsilon'$. We see that our full bound on the left hand side of equation (7.6) is the following,

$$(7.8) \qquad \frac{\epsilon'}{d-1}\sum_{\alpha} |P_{i, l_\alpha}||P_{l_\alpha, j}| + \frac{\phi}{d-1}\sum_{\alpha \neq \beta} |P_{i, l_\alpha}||P_{l_\beta, j}| + \frac{\epsilon'(\log N)^2}{d-1}\Big(\frac{1}{\sqrt{d-1}}\Big)^{\ell - d_i}\Big(\frac{1}{\sqrt{d-1}}\Big)^{\ell - d_j}$$

Lemma 4.5 assures us $\sum_\alpha |P_{i, l_\alpha}| \leq (\log N)^{\mathfrak{j}(\mathfrak{d})\mathfrak{a}}$. Thus, we see that

$$\phi \sum_{\alpha \neq \beta} |P_{i, l_\alpha}||P_{l_\beta, j}|\frac{(d-1)^{(2\ell - d_i - d_j)/2}}{(d-1)^{(2\ell - d_i - d_j)/2}} \leq \frac{\phi(\log N)^{2\mathfrak{j}(d)\mathfrak{a} + 4\mathfrak{a}}}{(d-1)^{(2\ell - d_i - d_j)/2}}.$$

Now, since $\phi \ll \epsilon(\log N)^{-K\mathfrak{a}/2}$, we see that for sufficiently large, but still finite, values of d we can show that the factor on the numerator is still less than ϵ.

We finally need to find a bound on $\sum_\alpha |P_{i, l_\alpha}||P_{l_\alpha, j}|$ that depends on the distance between i and j in $\mathbb{T}$. From the proof of Lemma 3.5, we see that this decay factor is $\lesssim (1 + \mathrm{dist}(i, j))\big(\frac{1}{\sqrt{d-1}}\big)^{\mathrm{dist}(i,j)}$. Just to give

a brief summary of the method here, we use the factor that $|P_{i,l_\alpha}|$ and $|P_{l_\alpha,j}|$ can decompose as a product based on the edges of the path connecting i to l_α and l_α to j respectively. Let $\hat{r}$ be the first common point of these paths.

The product $|P_{i,l_\alpha}||P_{l_\alpha,j}|$ roughly decomposes as $K_{i,\hat{r}}K_{\hat{r},j}|P_{\hat{r},l_\alpha}|^2$, where $K_{i,\hat{r}}$ and $K_{\hat{r},j}$ are two quantities that depend on the path between i to $\hat{r}$ and j to $\hat{r}$ respectively. A naive bound on $K_{i,\hat{r}}$ is $\left(\frac{1}{\sqrt{d-1}}\right)^{\text{dist}(i,\hat{r})}$, while a naive bound on $K_{\hat{r},j}$ is $\left(\frac{1}{\sqrt{d-1}}\right)^{\text{dist}(\hat{r},j)}$. The product of these two terms gives the decay $\left(\frac{1}{\sqrt{d-1}}\right)^{\text{dist}(i,j)}$.

If we fix $\hat{r}$ and sum over all l_α that would be possible descendants or $\hat{r}$, the computations in Lemma 3.5 show that $\sum_{l_\alpha}|P_{\hat{r}l_\alpha}|^2$ is $O(1)$. Finally, there are $1 + \text{dist}(i,j)$ choices for the first common point $\hat{r}$. This is the basic strategy that bound $\sum_\alpha |P_{i,l_\alpha}||P_{l_\alpha,j}|$ and finishes the proof of our bound. $\qquad \square$

To complete this section, we also include estimates on the differences $|\tilde{G}_{ij} - P_{ij}|$ when at least one of i and j are outside of $\mathbb{T}$. First consider the case that i is in $\mathbb{T}$ and j is in $\mathbb{T}^c$. Applying the resolvent identity, we have $\tilde{G}|_{\mathbb{T}\times\mathbb{T}^c} = \tilde{G}|_{\mathbb{T}}B\tilde{G}^{(\mathbb{T})}$.

Lemma 7.3. *Let i be a vertex in $\mathbb{T}$ and j be a vertex in $\mathbb{T}^c$. We have the following estimate on $\tilde{G}_{ij} - P_{ij}$.*

$$\tag{7.9} |\tilde{G}_{ij} - P_{ij}| \lesssim (\log N)\epsilon'\left(\frac{1}{\sqrt{d-1}}\right)^{\ell-d_i}.$$

Proof. Applying the resolvent identity and taking the difference, we have,

$$\tag{7.10} \tilde{G}|_{\mathbb{T}\times\mathbb{T}^c} - P|_{\mathbb{T}\times\mathbb{T}^c} = (G|_{\mathbb{T}} - P|_{\mathbb{T}})BP^{(\mathbb{T})} + P|_{\mathbb{T}}B(G^{(\mathbb{T})} - P^{(\mathbb{T})}) + (G|_{\mathbb{T}} - P|_{\mathbb{T}})B(G^{(\mathbb{T})} - P^{(\mathbb{T})}).$$

As before, the term with two factors of the form $(G - P)$ would give an error of the form $(\epsilon')^2$, which is far smaller than any main term error we might have to consider.

Now, let us analyze the first term; writing out the indices, it will be,

$$\tag{7.11} \sum_\alpha (G|_{\mathbb{T}} - P|_{\mathbb{T}})_{i,l_\alpha} P^{(\mathbb{T})}_{\tilde{c}_\alpha,j}.$$

We first remark that $P^{(\mathbb{T})}$ is a block-diagonal matrix and that j would belong to a block containing at most $O(1)$ elements $\tilde{c}_\alpha$. Thus, $P^{(\mathbb{T})}_{\tilde{c}_\alpha,j}$ would be zero for all except for $O(1)$ many elements. Now, we can apply the estimate from equation (7.5) in order to bound the first error term $(G|_{\mathbb{T}} - P|_{\mathbb{T}})$.

Now, let us analyze the second error term. Again, in terms of coordinates, we have,

$$\tag{7.12} \sum_\alpha (P|_{(\mathbb{T})})_{il_\alpha}(G^{(\mathbb{T})} - P^{(\mathbb{T})})_{\tilde{c}_\alpha,j}.$$

Again, we have to apply the logic of connectedness. We first remark that j cannot be Green's function connected to more than $O(\log N)$ vertices of the form c_α by the construction of our switching event. For all c_α to which j is not Green's function connected, we apply the improved bound $|G^{(\mathbb{T})} - P^{(\mathbb{T})}| \leq (\epsilon')^2 + \phi$. For the other $O(\log N)$ terms, we merely use the error bound ϵ' instead.

Our total error can be bounded by,

$$\tag{7.13} \sum_\alpha |(P|_{(\mathbb{T})})_{il_\alpha}|((\epsilon')^2 + \phi) + C\log N\epsilon'\left(\frac{1}{\sqrt{d-1}}\right)^{\ell-d_i},$$

where d_i is the distance of the vertex i from the center of the tree $\mathbb{T}$. From Lemma A.2, we can bound $\sum_\alpha |(P|_{(\mathbb{T})})_{il_\alpha}|(d-1)^{\ell/2} \leq (\log N)^{\mathfrak{j}(d)\mathfrak{a}+\mathfrak{a}/2}$. Even after multiplying this constant by ϕ, we obtain a value that is less than ϵ. Therefore, the error coming from this term is still less than $C\log N\left(\frac{1}{\sqrt{d-1}}\right)^{\ell-d_i}$. $\qquad \square$

7.1. **Changing $Q_{I/O}$ to $\tilde{Q}_{I/O}$.**

Our stability estimates are finally enough to understand the shift of $\mathcal{G}$ to $\tilde{\mathcal{G}}$ for Q_I and Q_O. The proofs of these two quantities are similar, so we will only prove the results for the shift of Q_I.

Lemma 7.4. *We have that,*

$$(7.14) \qquad |Q_{I/O}(\mathcal{G}) - \tilde{Q}_{I/O}| \lesssim \frac{1}{N^{1-\mathfrak{c}}} + (2d-1)^{2\ell} \frac{Im[m_{\mathcal{T}}^d(z,w)] + \epsilon' + \frac{\epsilon}{S_g^1}}{N\eta}.$$

Proof. Clearly, once we remove the neighborhood $\mathbb{T}$ and switching set W_S, the graphs $\mathcal{G}$ and $\tilde{\mathcal{G}}$ are the same. Thus, our goal is to relate terms of the form $G_{jj}^{(i)}$ to those of the form $G_{jj}^{(\{i\}\cup\mathbb{T}\cup W_S)}$ via resolvent estimates. We will use $i\mathbb{T}W_S$ as a shorthand for set $\{i\} \cup \mathbb{T} \cup W_S$. Let $\chi(\cdot)$ be the characteristic function indicating whether the vertex in question is of at least distance $\frac{\mathfrak{R}}{4}$ away from any of the vertices $a_\alpha, b_\alpha, c_\alpha$ that take part in the switching and, furthermore, j has a radius $\frac{\mathfrak{R}}{4}$ tree-like neighborhood. We note that there are $\ll N^{\mathfrak{c}}$ vertices j such that $\chi(j) \neq 1$.

Namely, our goal is to prove the following two inequalities,

$$(7.15) \qquad \begin{aligned} \frac{1}{Nd}\sum_{j\to i}\chi(j)|G_{jj}^{(i)} - G_{jj}^{(i\mathbb{T}W_S)}| &\leq (2d-1)^{2\ell}\frac{Im[m_{\mathcal{T}}^d(z,w)] + \epsilon' + \frac{\epsilon}{S_g^1}}{N\eta} \\ \frac{1}{Nd}\sum_{j\to i}\chi(j)|\tilde{G}_{jj}^{(i)} - G_{jj}^{(i\mathbb{T}W_S)}| &\leq (2d-1)^{2\ell}\frac{Im[m_{\mathcal{T}}^d(z,w)] + \epsilon' + \frac{\epsilon}{S_g^1}}{N\eta}, \end{aligned}$$

as the proof is similar, we only give the details for the first inequality.

For those values j such that $\chi(j) = 0$, we use the fact that $G_{jj}^{(i)}$ and $\tilde{G}_{jj}^{(i)}$ are both $O(1)$. Thus,

$$(7.16) \qquad \frac{1}{Nd}\sum_{j\to i}(1-\chi(j))|G_{jj}^{(i)} - \tilde{G}_{jj}^{(i)}| \lesssim N^{\mathfrak{c}-1},$$

since the sum is non-zero for $\ll N^{\mathfrak{c}}$ many vertices.

By the resolvent identity, we have,

$$(7.17) \qquad G_{jj}^{(i)} = G_{jj}^{(i\mathbb{T}W_S)} + [G^{(i\mathbb{T}W_S)}B^*G|_{i\mathbb{T}W_S}BG^{(i\mathbb{T}W_S)}]_{jj},$$

Here, B is the adjacency matrix for the graph $\mathcal{G}$ between $i\mathbb{T}W_S$ and the complement.

We would also have,

$$(7.18) \qquad \tilde{G}_{jj}^{(i)} = G_{jj}^{(i\mathbb{T}W_S)} + [G^{(i\mathbb{T}W_S)}\tilde{B}^*\tilde{G}^{(i)}|_{i\mathbb{T}W_S}\tilde{B}G^{(i\mathbb{T}W_S)}]_{jj}.$$

Here, $\tilde{B}$ is the switched adjacency matrix for the matrix $\tilde{\mathcal{G}}$. We will consider the matrix product that appears in the end (the GB^*GBG) to be purely an error term.

We see that we have,

$$(7.19) \qquad \begin{aligned} &\frac{1}{Nd}\sum_{(j\to i)\in E}\chi(j)|G_{jj}^{(i)} - G_{jj}^{(i\mathbb{T}W_S)}| \\ &\leq \frac{1}{Nd}\sum_{(j\to i)\in E}\sum_{x,y\in\delta(\mathbb{T}W_S)}\sum_{a,b\in\mathbb{T}W_S}|G_{jx}^{(i\mathbb{T}W_S)}|B_{xa}|[G^{(i)}|_{\mathbb{T}W_S}]_{ab}|B_{by}|G_{yj}^{(i\mathbb{T}W_S)}|. \end{aligned}$$

$\delta(\mathbb{T}W_S)$ consists of the boundary vertices in $\mathcal{G}\backslash\{i\}$ that border a vertex of $\mathbb{T}\cup W_S$. Now, for given x and y, there are at most $O(1)$ choices for a and b for which one would not have $B_{xa} \neq 0$, and $B_{by} \neq 0$. Now, we argue that $|[G^{(i)}|_{\mathbb{T}W_S}]_{ab}|$ is $O(1)$ for all pairs of vertices a and b in $\mathbb{T}\cup W_S$.

An application of the resolvent identity would give,

$$(7.20) \qquad G_{ab} = G_{ab}^{(i)} + G|_{a\times\{i,i+N\}}(G|_{\{i,i+N\}})^{-1}G|_{\{i,i+N\}\times b}.$$

By applying our induction hypothesis from Definition 2.7 on the terms of G, we would know that G_{ab} and the entries of $G_{a\times\{i,i+N\}}$ and $G_{\{i,i+N\}\times b}$ are all $O(1)$ quantities. One can also apply the induction hypothesis and simple perturbation theory to show that the entries of $(G|_{\{i,i+N\}})^{-1}$ are themselves $O(1)$ quantities. By rearrangement, this means that

$$(7.21) \qquad \frac{1}{Nd} \sum_{(j \to i) \in E} \chi(j) |G_{jj}^{(i)} - G_{jj}^{(i\mathbb{T}W_S)}| \lesssim \frac{1}{Nd} \sum_{(j \to i) \in E} \sum_{x,y \in \delta(\mathbb{T}W_S)} \chi(j) |G_{jx}^{(i\mathbb{T}W_S)}||G_{yj}^{(i\mathbb{T}W_S)}|.$$

Now, we apply the Cauchy-Schwarz inequality to bound this by,

$$(7.22) \qquad \frac{1}{Nd} \sum_{x,y \in \delta(\mathbb{T}W_S)} \Big[\sum_{(j \to i) \in E} \chi(j) |G_{jx}^{(i\mathbb{T}W_S)}|^2 \sum_{(j \to i) \in E} \chi(j) |G_{yj}^{(i\mathbb{T}W_S)}|^2 \Big]^{1/2}.$$

We apply the resolvent identity again to relate $G^{(i\mathbb{T}W_S)}$ to $G^{(\mathbb{T}W_S)}$, which we do have estimates to control. Namely, we again see that,

$$(7.23) \qquad G_{jx}^{(i\mathbb{T}W_S)} = G_{jx}^{(\mathbb{T}W_S)} - G^{(\mathbb{T}W_S)}|_{j \times \{i,i+N\}} (G^{(\mathbb{T}W_S)}|_{\{i,i+N\}})^{-1} G^{(\mathbb{T}W_S)}|_{\{i,i+N\} \times x}.$$

We know that the terms of $G^{(TW_S)}|_{j \times \{i,i+N\}}$ are of $O(1)$ by Proposition 2.16. Recall that since $\chi(j)$ is 1, j must have a tree-like neighborhood of radius $\frac{\mathfrak{R}}{4}$ in the graph $\mathcal{G} \setminus (\mathbb{T} \cup W_S)$; as i is a vertex adjacent to j. Thus, we can apply estimates from perturbation theory to ensure that the terms of $(G^{(\mathbb{T}W_S)}|_{\{i,i+N\}})^{-1}$ are $O(1)$. From these two facts, we now can assert that,

$$(7.24) \qquad \chi(j) |G_{jx}^{(i\mathbb{T}W_S)}|^2 \lesssim |G_{jx}^{(\mathbb{T}W_S)}|^2 + |G_{ix}^{(\mathbb{T}W_S)}|^2,$$

Now, by applying the Ward identity, we see that,

$$(7.25) \qquad \sum_{(j \to i)} |G_{jx}^{(i\mathbb{T}W_S)}|^2 \lesssim \frac{\text{Im}[G_{jj}^{(\mathbb{T}W_S)}]}{\eta} \lesssim \frac{\text{Im}[m_{\mathcal{T}}^d(z,w)] + \epsilon' + \frac{\epsilon}{S_g^1}}{\eta}.$$

We used the following triangle inequality to estimate $G_{jj}^{(\mathbb{T}W_S)}$,

$$(7.26) \qquad \begin{aligned} |G_{jj}^{(\mathbb{T}W_S)} - m_{\mathcal{T}}^d(z,w)| \leq & |G_{jj}^{(\mathbb{T}W_S)} - G_{jj}(Ext(B_r(j,\mathcal{G}), Q_O, Q_I))| \\ & + |G_{jj}(Ext(B_r(j,\mathcal{G}), Q_O, Q_I)) - G_{jj}(Ext(B_r(j,\mathcal{G}), m_\infty, m_\infty))|. \end{aligned}$$

The term $G_{jj}(Ext(B_r(\{j\},\mathcal{G}), m_\infty, m_\infty))$ is the same as $m_{\mathcal{T}}^d(z,w)$. Now, we can estimate the first term on the right hand side of the above inequality by ϵ' by using Proposition 2.15. The second term on the right hand side of the above inequality is bounded by $\frac{\epsilon}{S_g^1}$ by our inductive hypothesis on the distance between $Q_{I/O}$ and m_∞ from Definition 2.7 .

We can return to the expression in (7.22) to bound the quantity inside the square root by the term above. Now, the number of x vertices on the boundary of $\mathbb{T}W_S$ is less than $(2d-1)^\ell$; the same is true of y. Thus, the outer sum in equation (7.22) will give a factor of no more than $(2d-1)^{2\ell}$. This will give us the desired inequality in equation (7.14). $\qquad \square$

With this estimate on the difference between Q_I/Q_O and $\tilde{Q}_I/\tilde{Q}_O$, we can finally return to our proof of Proposition 2.18 is simple.

Proof of Proposition 2.18. From Lemmas 7.1and 7.3, we have a comparison between $\tilde{G}$ and a P whose extension is given by parameters Q_I and Q_O. The last Lemma 7.4 suggests that the difference between $Q_{I/O}$ and $\tilde{Q}_{I/O}$ are small enough that we can apply the perturbation estimates from Theorem 1.13. These estimates are sufficient for the proof of Proposition 2.18. $\qquad \square$

8. Concentration Estimates

The main issue with our stability estimate in equation (7.5) is the presence of the term whose decay is $\left(\frac{1}{d-1}\right)^{\frac{|d_j-d_i|}{2}}$. As long as d_j and d_i are relatively small, this term will not decay as N and ℓ increase. Recall that this error term came from estimating the following term,

$$\frac{1}{d-1}\sum_\alpha P_{i,l_\alpha}P_{l_\alpha,j}(\tilde{G}^{(\mathbb{T})}-P^{(\mathbb{T})})_{c_\alpha,c_\alpha}.$$

While the triangle inequality will not show an improvement, the main point is that this new graph came from random switching. Thus, we can try to apply a concentration argument to show that that the quantity that appears above has superior bounds. Loosely speaking, the improvement will improve the constant $\sum_\alpha |P_{i,l_\alpha}||P_{l_\alpha,j}|$ to $\sqrt{\sum_\alpha |P_{i,l_\alpha}P_{l_\alpha,j}|^2}$. Under the l_2 norm on $|P_{i,l_\alpha}||P_{l_\alpha,j}|$, the branching factor of the tree no longer cancels out the effect of the decrease. This is the main improvement that we want to seek. To derive this improvement, we use the fact that the c_α's are randomly chosen, independent from each other. Using this fact, one would expect better concentration results for most switching neighborhoods.

The main result of this section is the following Theorem,

Theorem 8.1. *Let $\mathcal{G}$ be a graph from $\Omega^1_{S,\mathfrak{q},o}$ with ℓ neighborhood $\mathbb{T}$. Consider the following weight families. These weight families will allow us to understand Q_I.*

  (1) *Let v be a vertex in $\mathbb{T}$ and let $\tilde{v}$ be an index corresponding to v (so it is either v or $v+N$). Recall our indexing of boundary vertices of $\mathbb{T}$ as l_α. We define the in-weights $w^{in}_\alpha := P_{xl_\alpha}P_{l_\alpha x}$. There are $\lesssim |\mathbb{T}|$ of these types of indices depending on the choice of the index x.*
  (2) *If $\mathbb{T}$ is a truncated d-regular tree, fix a vertex k that is connected to o; assume for now that the edge is oriented $k\to o$. Define A to be a set of indices as follows. Consider a vertex l_α on the boundary of $\mathbb{T}$; if the path from o to v_α does not pass through k, then we include the index α in A. We then define the weight $w^{in}_\alpha := P^{(k)}_{o,l_\alpha}P^{(k)}_{l_\alpha,o}$. $P^{(k)}$ is a shorthand for $G_{o,l_\alpha}(Ext(B_r(\mathbb{T},\mathcal{G}^{(i)}),Q_I,Q_O))$. If instead the edge were oriented $o\to k$, we would instead choose the weights $P^{(k)}_{o+N,l_\alpha}P^{(k)}_{l_\alpha,o+N}$. There are $2d$ such weights of this type.*

There exists an event $S_C(\mathcal{G})\subset S_G(\mathcal{G})$ such that $1-\mathbb{P}(S_C(\mathcal{G}))=1-\mathbb{P}(S_G(\mathcal{G}))-(|T|+d)\exp[-(\log N)^2]$ such that we have the following estimate,

$$\left|\sum_\alpha w_\alpha(\tilde{G}^{(\mathbb{T})}_{\tilde{a}_\alpha+N,\tilde{a}_\alpha+N}-G_{\tilde{a}_\alpha+N,\tilde{a}_\alpha+N}(Ext_i(B_r(\tilde{a}_\alpha,\tilde{\mathcal{G}}^{(\mathbb{T})}),Q_I,Q_O)))-\sum_\alpha w_\alpha(Q_I-Y_{i,r}(Q_I,Q_O))\right|$$

$$\lesssim (\log N)\epsilon'\sqrt{\sum_\alpha |w_\alpha|^2}+\left((2d-1)^{2\ell}\frac{Im[m^d_\mathcal{T}]+\epsilon'+\frac{\epsilon}{S^1_g}}{N\eta}+N^{\mathfrak{c}-1}\right)+\frac{1}{N^{1-\mathfrak{c}}}.$$

Remark 8.2. We have a similar concentration statement to understand the concentration around Q_O. We do not write out the whole statement for simplicity, but mention the appropriate changes to the weights and computed function. In the first type of weight family, we change the weights to $w^{\text{out}}_\alpha := P_{x,l_\alpha+N}P_{l_\alpha+N,x}$. The function which we try to compute using concentration estimates will be of the form $\sum_\alpha w_\alpha \tilde{G}_{\tilde{a}_\alpha,\tilde{a}_\alpha}$.

The proof of the above statement will be divided into smaller parts. Since the random variables $\tilde{a}_\alpha$ are chosen independently of each other given the graph, $\tilde{G}^{(\mathbb{T})}_{\tilde{a}_\alpha+N,\tilde{a}_\alpha+N}$ are random variables that are almost independent of the weights w_α and of each other . However, the fact that they are not truly independent prevents us from directly applying standard concentration estimates. (The value of the Green's function $\tilde{G}^{(\mathbb{T})}_{\tilde{a}_\alpha+N,\tilde{a}_\alpha+N}$ does not depend on only the local geometry, but on the full graph structure of the switched graph. Thus, changing the value of $\tilde{a}_\alpha$ for a single α, while keeping all others the same, will have subtle effects on $\tilde{G}^{(\mathbb{T})}_{\tilde{a}_\beta,\tilde{a}_\beta}$ for $\beta\neq\alpha$. The distribution of $\tilde{G}^{(\mathbb{T})}_{\tilde{a}_\alpha,\tilde{a}_\alpha}$ will also depend on l_α as well.) Our first step is to derive concentration estimates for a closely related family of independent random variables.

8.1. **Concentration estimates for a related family of independent random variables.** For a vertex $v \in \mathcal{G}$, we use $(v\mathbb{T})$ as a short way to express the union $\{v\} \cup \mathbb{T}$. Also let $\chi(\cdot)$ is the indicator function of the event that the vertex is question is of distance at least $\mathfrak{R}/4$ from the vertices that bound $\mathbb{T}$ and that the vertex in question has a radius $\mathfrak{R}/4$ tree-like neighborhood. As a shorthand for what will follow, we will define D_α as a shorthand for the difference,

$$(8.1) \qquad D_\alpha^{\mathrm{in}} := \chi(c_\alpha)(G^{(b_\alpha \mathbb{T})} - P^{(b_\alpha \mathbb{T})})_{c_\alpha+N, c_\alpha+N},$$

where recall that $P^{(b_\alpha \mathbb{T})}$ was defined to be the Green's function of the extension $P_{ij}^{(b_\alpha \mathbb{T})} := G(Ext(B_r(\{i,j\}), \mathcal{G}\backslash(b_\alpha\mathbb{T})), Q_I, Q_O))$, while $P_{ij}^{(\mathbb{T})}$ is simply $G_{ij}(Ext(B_r(\{i,j\}), \mathcal{G}\backslash\mathbb{T}), Q_I, Q_O))$. Note that we specifically chose the unswitched graph here.

Lemma 8.3. *Consider all of the families of weights w_α as given in Theorem 8.1 Let $S_C(\mathcal{G}) \subset S_G(\mathcal{G})$ be the set of switching events such that the following event holds for all of the weight families.*

$$(8.2) \qquad \begin{aligned} &\Big| \sum_\alpha w_\alpha D_\alpha^{in} - \sum_\alpha w_\alpha(Q_I - Y_{i,r}(Q_I, Q_O)) \Big| \lesssim \\ &(\log N)\epsilon' \sqrt{\sum_\alpha |w_\alpha|^2} + \Big((2d-1)^{2\ell} \frac{Im[m_{\mathcal{T}}^d] + \epsilon' + \frac{\epsilon}{S_g^1}}{N\eta} + N^{\mathfrak{c}-1} \Big). \end{aligned}$$

Then, $\mathbb{P}(S_C(\mathcal{G}) = \mathbb{P}(S_G(\mathcal{G})) - (|T|+d)\exp[-(\log N)^2]$.

Proof. We have shown in Proposition 2.15 that each of the differences satisfy $|(G^{(\mathbb{T})} - P^{(\mathbb{T})})_{c_\alpha+N, c_\alpha+N}| \leq \epsilon'$. We can apply the resolvent identity to understand $|G_{c_\alpha+N, c_\alpha+N}^{(b_\alpha\mathbb{T})} - P_{c_\alpha+N, c_\alpha+N}^{(b_\alpha\mathbb{T})}|$. First of all, we see that, for edges $(i \to j)$,

$$(8.3) \qquad \chi(j)G_{j+N, j+N}^{(i\mathbb{T})} = \chi(j)(G_{j+N, j+N}^{(\mathbb{T})} - G^{(\mathbb{T})}|_{j+N\times\{i, i+N\}}(G^{(\mathbb{T})}|_{\{i, i+N\}})^{-1}G^{(\mathbb{T})}|_{\{i, i+N\}\times j+N}).$$

In addition, we have that,

$$(8.4) \qquad \chi(j)P_{j+N, j+N}^{(i\mathbb{T})} = \chi(j)(P_{j+N, j+N}^{(\mathbb{T})} - P^{(\mathbb{T})}|_{j+N\times\{i, i+N\}}(P^{(\mathbb{T})}|_{\{i, i+N\}})^{-1}P^{(\mathbb{T})}|_{\{i, i+N\}\times j+N}).$$

We have the estimate that $|G^{(\mathbb{T})} - P^{(\mathbb{T})}| \lesssim \epsilon'$ from Proposition 4.3. We also know that entries of $(P^{(\mathbb{T})}|_{\{i, i+N\}})^{-1}$ are $O(1)$ quantities when $\chi(j) \neq 0$, since, in this case, i has at least a radius $\mathfrak{R}$ treelike-neighborhood in $\mathcal{G}\backslash T$. This allows us to apply a simple perturbation argument to assert that,

$$(8.5) \qquad \begin{aligned} \chi(j)\Big| &G^{(\mathbb{T})}|_{j+N\times\{i, i+N\}}(G^{(\mathbb{T})}|_{\{i, i+N\}})^{-1}G^{(\mathbb{T})}|_{\{i, i+N\}\times j+N} \\ &- P^{(\mathbb{T})}|_{j+N\times\{i, i+N\}}(P^{(\mathbb{T})}|_{\{i, i+N\}})^{-1}P^{(\mathbb{T})}|_{\{i, i+N\}\times j+N} \Big| \lesssim \epsilon'. \end{aligned}$$

These facts ensure that,

$$(8.6) \qquad \chi(j)|G_{j+N, j+N}^{(i\mathbb{T})} - P_{j+N, j+N}^{(i\mathbb{T})}| \leq \epsilon',$$

uniformly. Furthermore, once we have fixed the graph $\mathcal{G}^{(\mathbb{T})}$, the switched edges are chosen independently of each other. This means that the variables D_α^{in} are independent and identically distributed.

By Hoeffding's inequality, we have the following probability bound,

$$(8.7) \qquad \mathbb{P}\Big(\big|w_\alpha(D_\alpha^{\mathrm{in}} - \mathbb{E}[D_\alpha^{\mathrm{in}}])\big| \geq t\sum_\alpha |w_\alpha|^2\epsilon' \Big) \leq \exp[-t^2/C],$$

for some constant C. Here, the probability $\mathbb{P}$ and the expectation $\mathbb{E}$ is only over the randomness of the choice of switched sets, only, not on the randomness of the random graph.

Our goal now is to relate $\mathbb{E}[D_\alpha^{\text{in}}]$ to the quantities Q_I. We first define the set $\mathcal{E}^{(\mathbb{T})}$ to be the set of all edges in the complement of $\mathbb{T}$ in $\mathcal{G}$. By definition,

$$(8.8) \qquad \mathbb{E}[D_\alpha^{\text{in}}] = \frac{1}{|\mathcal{E}^{(\mathbb{T})}|} \sum_{(i \to j)} \chi(j)(G_{j+N,j+N}^{(i\mathbb{T})} - P_{j+N,j+N}^{(i\mathbb{T})}).$$

We have earlier shown that all of these quantities are less than ϵ' uniformly. In addition, we have that $|\mathcal{E}^{(\mathbb{T})}| \geq Nd - O((2d-1)^\ell) \geq Nd - O(N^{\mathfrak{c}})$. By adjusting the normalization factor, we have that

$$(8.9) \qquad \mathbb{E}[D_\alpha^{\text{in}}] = \frac{1}{Nd} \sum_{(i \to j)} \chi(j)(\tilde{G}_{j+N,j+N}^{(i\mathbb{T})} - P_{j+N,j+N}^{(i\mathbb{T})}) + O\left(\frac{N^{\mathfrak{c}}}{Nd}\right).$$

Now, we relate $G_{j+N,j+N}^{(i)}$ to $G_{j+N,j+N}^{(i\mathbb{T})}$ via the Schur complement formula. We can apply the same argument as found in the proof of equation (7.15) to show that,

$$(8.10) \qquad \frac{1}{Nd} \sum_{(i \to j)} \chi(j)|G_{j+N,j+N}^{(i)} - G_{j+N,j+N}^{(i\mathbb{T})}| \lesssim (2d-1)^{2\ell} \frac{\text{Im}[m_{\mathcal{T}}^d] + \epsilon' + \frac{\epsilon}{S_g^1}}{N\eta},$$

Furthermore, $\chi(j) = 1$ implies that the vertex j has a radius $\mathfrak{R}/4$ tree-like neighborhood ensures. Thus, we have that,

$$(8.11) \qquad \frac{1}{Nd} \sum_{(i,j)} \chi(j) P_{j+N,j+N}^{(i\mathbb{T})} = \frac{1}{Nd} \sum_{(i \to j)} \chi(j) Y_{1,r}(Q_I, Q_O).$$

For the remaining quantities with $\chi(j) = 0$, we know that $P_{j+N,j+N}^{(i)}$ is $O(1)$ from Lemma 3.5. Using Proposition 2.15 would further imply that the entries $G_{j+N,j+N}^{(i)}$ are also $O(1)$. This bounds the sum of terms $\frac{1}{Nd} \sum_{i \to j}(1 - \chi(j)) G_{j+N,j+N}^{(i)}$ by $N^{\mathfrak{c}-1}$.

Combining these facts, we see that,

$$(8.12) \qquad \sum_\alpha w_\alpha \mathbb{E}[D_\alpha^{in}] = \sum_\alpha w_\alpha(Q_I(z,w) - Y_{i,r}(Q_i, Q_O)) + O\left((2d-1)^{2\ell} \frac{\text{Im}[m_{\mathcal{T}}^d] + \epsilon' + \frac{\epsilon}{S_g^1}}{N\eta} + N^{\mathfrak{c}-1}\right).$$

In the final line, we used the fact that $\sum_\alpha w_\alpha \lesssim 1$. $\qquad \square$

To use our concentration estimates for the unswitched graph onto the switched graph, we apply the following lemma, which will show that the values of $\tilde{G}^{(\mathbb{T})})_{c_\alpha+N,c_\alpha+N}$ are close to those of $G_{c_\alpha+N,c_\alpha+N}^{(b_\alpha\mathbb{T})}$ as well as similar results for the index c_α.

Lemma 8.4. *Assume that c_α has distance at least $\mathfrak{R}/4$ from all of the other vertices $\{a_\beta, b_\beta, c_\beta\}, \beta \neq \alpha$ that participate in the switching and that the pair c_α, b_α is not Green's function connected to $\{c_\beta, b_\beta\}$ for $\beta \neq \alpha$. Then, we have that,*

$$(8.13) \qquad |G_{c_\alpha,c_\alpha}^{(b_\alpha\mathbb{T})} - \tilde{G}_{c_\alpha,c_\alpha}^{(\mathbb{T})}| \lesssim (2d-1)^\ell \phi^2$$

Proof. The desired estimate is a consequence of the following intermediate comparison estimates,

$$(8.14) \qquad \begin{aligned} |G_{c_\alpha,c_\alpha}^{(\mathbb{T}W_S)} - G_{c_\alpha,c_\alpha}^{(\mathbb{T}b_\alpha)}| &\lesssim (2d-1)^\ell (\phi)^2 \\ |G_{c_\alpha,c_\alpha}^{(\mathbb{T}W_S)} - \tilde{G}_{c_\alpha,c_\alpha}^{(\mathbb{T})}| &\lesssim (2d-1)^\ell (\phi^2) \end{aligned}$$

The proof of both of these estimates are similar, so it will suffice to just prove the first one.

By the resolvent identity, we have,

$$(8.15) \qquad |G_{c_\alpha,c_\alpha}^{(\mathbb{T}W_S)} - G_{c_\alpha,c_\alpha}^{(b_\alpha\mathbb{T})}| \leq \left| \sum_{x,y \in W_s \backslash b_\alpha} G_{c_\alpha,x}^{(b_\alpha\mathbb{T})} (G^{(b_\alpha\mathbb{T})}|_{W_s\backslash b_\alpha})_{xy}^{-1} G_{y,c_\alpha}^{(b_\alpha\mathbb{T})} \right|$$

Now, the proof of Proposition 2.16 would also show that the elements $G_{c_\alpha,x}^{(b_\alpha\mathbb{T})}$ and $G_{y,c_\alpha}^{(b_\alpha\mathbb{T})}$ are both $\lesssim \phi$. Furthermore, Lemma 5.5 would show that $|(G^{(b_\alpha\mathbb{T})}|_{W_S\backslash b_\alpha})_{xy}^{-1}|$ is bounded by ϵ' if x and y are not indices that

correspond to the same vertex, while we can bound this by $O(1)$ if they do correspond to the same index. Using these facts, we see that the bound on the right hand side of (8.15) is $(2d-1)^\ell \phi^2 + (2d-1)^{2\ell} \epsilon' \phi^2$ which is dominated by the first term.

For the term involving the switched graph, the Schur complement formula gives us,

$$(8.16) \qquad |\tilde{G}^{(\mathbb{T})}_{c_\alpha, c_\alpha} - G^{(\mathbb{T}W_S)}_{c_\alpha, c_\alpha}| \leq \Big| \sum_{x,y \in \delta(W_S)} (G^{(\mathbb{T}W_S)}B)_{c_\alpha, x}(\tilde{G}^{(\mathbb{T})})_{xy}(B'G^{(\mathbb{T}W_S)})_{y, c_\alpha} \Big|,$$

where $\delta(W_S)$ correspond to the vertices that neighbor those of W_S in $\mathcal{G}\backslash W_S$. We have shown earlier in Proposition 2.16 that we can bound the entries of $(G^{(\mathbb{T}W_S)}B)_{c_\alpha, x}$ and $(B'G^{(\mathbb{T}W_S)})_{y, c_\alpha}$ by ϕ. Furthermore, we would also know that $\tilde{G}^{(\mathbb{T})}_{xy}$ is $O(1)$ if x and y correspond to the same vertex or is bounded by ϵ' otherwise by Lemma 6.3. The same counting gives us the desired bound. $\qquad \square$

With these two lemmas in hand, we can return to a proof of Theorem 8.1.

Proof of Theorem 8.1. From Lemma 8.3, we know that $\sum_\alpha w_\alpha D^{in}_\alpha$ is close to the desired estimate $\sum_\alpha w_\alpha (Q_I - Y_{i,r}(Q_I, Q_O))$ on the event $S_C \mathcal{G}$. It suffices to understand the differences,

$$(8.17) \qquad \sum_\alpha |w_\alpha||D^{in}_\alpha - (\tilde{G}^{(\mathbb{T})}_{\tilde{a}_\alpha + N, \tilde{a}_\alpha + N} - G_{\tilde{a}_\alpha + N, \tilde{a}_\alpha + N}(Ext(B_r(\tilde{a}_\alpha, \tilde{\mathcal{G}}^{(\mathbb{T})}), Q_1, Q_2))|.$$

We know that for all terms in which the pair c_α, b_α is not Green's function connected to c_β, b_β for $\beta \neq \alpha$ that we know $|\tilde{G}^{(\mathbb{T})}_{a_\alpha + N, a_\alpha + N} - G^{(b_\alpha \mathbb{T})}_{a_\alpha + N, a_\alpha + N}| \lesssim (2d-1)^\ell \phi^2$ via Lemma 8.4. Furthermore, we would also know that since $\tilde{a}_\alpha$ would have a radius $\mathfrak{R}/4$ tree-like neighborhood that,

$$(8.18) \qquad G_{\tilde{a}_\alpha + N, \tilde{a}_\alpha + N}(Ext_i(B_r(\tilde{a}_\alpha, \tilde{\mathcal{G}}\backslash\mathbb{T}), Q_I, Q_O)) = G_{\tilde{a}_\alpha + N, \tilde{a}_\alpha + N}(Ext_i(B_r(\tilde{a}_\alpha, \mathcal{G}\backslash\mathbb{T}), Q_I, Q_O)).$$

This deals with all except for $O(1)$ many indices α. For the remaining indices, we may use the trivial bound ϵ'.

Thus, we see that,

$$\sum_\alpha w_\alpha (\tilde{G}^{(\mathbb{T})}_{\tilde{a}_\alpha + N, \tilde{a}_\alpha + N} - G_{\tilde{a}_\alpha + N, \tilde{a}_\alpha + N}(Ext_i(B_r(\tilde{a}_\alpha, \tilde{\mathcal{G}}^{(\mathbb{T})}), Q_1, Q_2))) - \sum_\alpha w_\alpha (Q_I - Y_{i,r}(Q_I, Q_O))$$

$$= \sum_\alpha w_\alpha D^{in}_\alpha + O(\sum_\alpha |w_\alpha|(2d-1)^\ell \phi^2 + \max_\alpha |w_\alpha| \epsilon')$$

Manifestly, $\max_\alpha |w_\alpha| \epsilon'$ is still less than $(\log N)\epsilon' \sqrt{\sum_\alpha |w_\alpha|^2}$. Furthermore, we can use the fact that $\sum_\alpha |w_\alpha| \lesssim 1$ and the value of ϕ from Definition 5.2 to assert that we can bound $\sum_\alpha |w_\alpha|(2d-1)^\ell \phi^2 \ll N^{1-\mathfrak{c}}$. Since $\max_\alpha |w_\alpha|$ decays as $(d-1)^{-\ell/2}$, we have sufficient control of this completes the proof of Theorem 8.1. $\qquad \square$

8.2. **Improved Bounds on $\tilde{G}$.**

Given the graph $\mathcal{G}$ in $\Omega^1_{S, \mathsf{q}, o}$, we choose a switching event $S \in \mathbb{S}_{\mathcal{G}, o}$ that belongs in $S_C(\mathcal{G})$ as in Theorem 8.1. With these concentration estimates in hand, we are now in position to prove improved bounds on our Green's functions in order to prove Theorem 2.11.

Proof of Theorem 2.11. We consider the event $S_C(\mathcal{G})$. The proof of Theorem 2.11 involves proving the three estimates (2.13), (2.14), and (2.15).

We start with the equation (2.15). This is a rather simple consequence of the estimates of Section 7. From Lemma 7.3, equation (7.9), we know that, $|G_{oi} - P_{oi}| \lesssim \frac{(\log N)^2 \epsilon'}{(d-1)^{\ell/2}}$. Here,

$$P_{oi} = G_{oi}(Ext(B_r(i\mathbb{T}, \tilde{\mathcal{G}}), Q_I, Q_O)),$$

In the above expression, we have to replace Q_I and Q_O with $\tilde{Q}_I$ and $\tilde{Q}_O$ as well as change the domain to $B_r(\{i, o\}$ instead of $B_r(i\mathbb{T})$.

First of all, we have bounds on the difference $Q_I - \tilde{Q}_I$ as well as $Q_O - \tilde{Q}_O$ from Lemma 7.4. Combining this bound with a slight modification of the perturbation bound from Lemma 3.5.

Namely, we have that,

$$|G_{oi}(Ext(B_r(i\mathbb{T}, \tilde{\mathcal{G}}), Q_I, Q_O)) - G_{oi}(Ext(B_r(i\mathbb{T}, \tilde{\mathcal{G}}), \tilde{Q}_I, \tilde{Q}_O))|$$

(8.19)
$$\lesssim \Big(\frac{1}{d-1}\Big)^{\mathrm{dist}(i,j)} (1 + \log N)^3 [|Q_I - \tilde{Q}_I| + |Q_O - \tilde{Q}_O|]$$

$$\lesssim \Big(\frac{1}{d-1}\Big)^{\mathrm{dist}(i,j)} (\log N)^3 \Big[\frac{1}{N^{1-\mathfrak{c}}} + (2d-1)^{2\ell} \frac{\mathrm{Im}[m_{\mathcal{T}}^d(z,w)] + \epsilon' + \frac{\epsilon}{S_g^1}}{N\eta}\Big].$$

Finally, we can use Lemma 3.7 to change the domain used in the Ext operator from $B_r(i\mathbb{T}, \mathcal{G})$ to $B_r(\{i, o\})$. This introduces the terms involving $|\tilde{Q}_I + \tilde{Q}_O - 2m_\infty|$, etc. on the right hand side of equation (2.15). Finally, we can combine all of these estimates using a triangle inequality to finish the proof of (2.15).

We remark here that, for the purposes of an induction argument, it is important to ensure that the term $\frac{(\log N)^2 \epsilon'}{(d-1)^{\ell/2}}$ would be less than the original error term ϵ . This guarantees that a switching would give a provable improvement to the error bounds. It is very important that we had the constant bound $\mathfrak{j}(d) < \frac{1}{2}$; this constant bound is what allows us to compensate the $(\log N)$ factors that appear in ϵ' with $(d-1)^{\ell/2}$ in the denominator.

The derivation equation (2.14) from equation (7.5) is quite similar to the strategy we used to prove equation (2.15). On a high level, we know that G_{oi} should be close to $G_{oi}(Ext(B_r(i\mathbb{T}, \tilde{\mathcal{G}}), Q_I, Q_O))$ due to the estimates from Section 7, specifically equation (7.5). We still need to replace the appearances of Q_I and Q_O in this bound with the appropriate $\tilde{Q}_I$ and $\tilde{Q}_O$. These rely on the same perturbation estimates we have used earlier.

However, we now need to find an improvement in the second term of the right hand side of equation (7.5); the first term is sufficiently small. Tracing through the proof of equation (7.5), This term comes from an attempt to bound the term on the second line of equation (7.6). Namely, the sum of the on-diagonal terms.

(8.20)
$$\frac{1}{d-1} \sum_\alpha P_{i,l_\alpha} P_{l_\alpha,j} (\tilde{G}^{(\mathbb{T})} - P^{(\mathbb{T})})_{\tilde{c}_\alpha, \tilde{c}_\alpha}$$

We can replace $P_{\tilde{c}_\alpha, \tilde{c}_\alpha}^{(\mathbb{T})}$ is defined with respect to the domain $B_r(c_\alpha \mathbb{T}, \mathcal{G}^{(\mathbb{T})})$. This can be replaced with the extension with respect to the domain $B_r(c_\alpha, \mathcal{G}^{(\mathbb{T})})$ with an error term of the same order of magnitude as the right hand side of (2.14). Now, since $\sum_\alpha |P_{i,l_\alpha}||P_{l_\alpha,i}| \lesssim 1$, it causes not issue to apply a triangle inequality.

After the appropriate replacement of the the domain of the extension, the term in (8.20) is controlled by the estimates of Theorem 8.1. This completes the desired bound of equation (2.14). The proof of (2.13) requires very similar manipulations as the previous two, but uses the second family of weights w_α from Theorem 8.1 rather than the first. $\qquad\square$

APPENDIX A. COLLECTED ESTIMATES ON GREEN'S FUNCTION OF INFINITE GRAPHS

A.1. **Bounds on G_{ij} for tree-like and almost tree-like graphs.** Recall our inductive procedure to compute for a directed tree, all Green's function values (via the Schur complement formula), starting from the solution m_∞ of (1.30). The success of this procedure requires that the sum of $|G_{ij}|^2$ over all vertices j of graph-distance k from i, be bounded uniformly in k. While the naive bound $G_{ij} \le (\sqrt{d-1})^{-\mathrm{dist}(i,j)}$ is not strong enough to compensate the branching factor of the tree, our next bound on m_∞, helps control the latter branching factor.

Lemma A.1. *For some $C_0 = C_0(d)$, $d \ge 3$ finite, any $z \in \mathbb{C}_+$, $w \in \mathbb{C}$ and any solution $m_\infty : \mathbb{C}_+ \to \mathbb{C}_+$ of* (1.30), *setting*

(A.1)
$$X := |m_\infty|^2, \qquad Y := \frac{d}{d-1} |m_\infty^{uod}|^2 = \frac{d}{d-1} \frac{|w|^2 |m_\infty|^2}{\big|z + \frac{d}{d-1} m_\infty\big|^2},$$

we have that for some finite $C_0 = C_0(d)$,

$$(\text{A.2}) \qquad S_g^1 = 1 - X - Y \geq \frac{1}{2}(\Im(z) \wedge 1) > 0, \qquad |m_\infty^{sd}| \leq C_0 \qquad |m_{\mathcal{T}}^d| \leq C_0.$$

Proof. Taking $(1/\cdot)$ of both sides of (1.30), results with

$$(\text{A.3}) \qquad \frac{1}{m_\infty} = \frac{|w|^2}{z + \frac{d}{d-1}m_\infty} - z - m_\infty.$$

The imaginary part of this identity is precisely

$$-\frac{\Im(m_\infty)}{|m_\infty|^2} = -\frac{|w|^2\big(\Im(z) + \frac{d}{d-1}\Im(m_\infty)\big)}{|z + \frac{d}{d-1}m_\infty|^2} - \Im(z) - \Im(m_\infty).$$

Set $\Gamma := \frac{\Im(z)}{\Im(m_\infty)} > 0$ and multiply the preceding by $-\frac{|m_\infty|^2}{\Im(m_\infty)}$, to arrive after some algebra at

$$(\text{A.4}) \qquad S_g^1 = 1 - X - Y = \Gamma\Big(X + \frac{d}{d-1}Y\Big).$$

With $\Gamma > 0$, necessarily $X + Y < 1$ and in particular $\Im(m_\infty) \leq \sqrt{X} \leq 1$. The RHS of (A.4) is thus at least $\Im(z)(1 - S_g^1)$ and the LHS of (A.2) follows.

Next, the real part of the identity (A.3) amounts to

$$\frac{\Re(m_\infty)}{|m_\infty|^2} = \frac{|w|^2\big(\Re(z) + \frac{d}{d-1}\Re(m_\infty)\big)}{|z + \frac{d}{d-1}m_\infty|^2} - \Re(z) - \Re(m_\infty).$$

With $X, Y \in [0, 1]$, it follows after some algebra that

$$\frac{\Re(m_\infty)}{\Re(z)} = \frac{\frac{d-1}{d}Y - X}{1 - Y + X} \geq -\frac{1}{2}.$$

Since $\Im(m_\infty) > 0$ and $\Im(z) > 0$, this implies in turn that also $\Re(m_\infty/z) \geq -1/2$. With $|m_\infty| \leq 1$, we thus deduce from (1.31) that for any $d \geq 3$,

$$|m_\infty^{sd}| \leq \frac{|1 + m_\infty/z|}{|\frac{d-1}{d} + m_\infty/z|} \leq \sup_{\Re(S) \geq -1/2}\left\{\frac{|1 + S|}{|\frac{d-1}{d} + S|}\right\} := C_0(d)$$

is finite, as claimed.

Now, for $m_{\mathcal{T}}^d$, we see that we have,

$$(\text{A.5}) \qquad \begin{aligned} \frac{1}{m_{\mathcal{T}}^d} &= \frac{|w|^2}{z + \frac{d}{d-1}m_\infty} - \Big(z + \frac{d}{d-1}m_\infty\Big) = \frac{1}{m_\infty} - \frac{m_\infty}{d-1} \\ &= \frac{(d-1) - m_\infty^2}{(d-1)m_\infty} \end{aligned}$$

Thus,

$$(\text{A.6}) \qquad |m_{\mathcal{T}}^d| = \frac{(d-1)|m_\infty|}{|d - 1 - m_\infty^2|} \leq \frac{d-1}{d-2}.$$

$\square$

Using Lemma A.1, we next show that $\sum_{y \in \partial B_k(\hat{r})} |G_{\hat{r},y}|^2$ is bounded, uniformly in k, for any one of the directed trees of interest to us. This computation from the root is relevant for the coefficient of various first order terms that will appear in our perturbation series.

Lemma A.2. *For the trees $\mathcal{T}_1$ and $\mathcal{T}_2$ from Definition 1.10 and any $k \geq 1$,*

(A.7)
$$A_{1,k} := \sum_{v \in \partial B_k(\hat{r}), o_v \in \{0, \aleph\}} |(G_{\mathcal{T}_1})_{\hat{r}+\aleph, v+o_v}|^2 \lesssim 1,$$
$$A_{2,k} := \sum_{v \in \partial B_k(\hat{r}), o_v \in \{0, \aleph\}} |(G_{\mathcal{T}_2})_{\hat{r}, v+o_v}|^2 \lesssim 1,$$

where the implicit constants do not depend on k. The corresponding quantities for the Green's function $G_{\mathcal{T}}$ of the 'pure' d-regular directed tree, are also bounded uniformly in k.

Proof. Suppose $l \in \mathcal{T}_1$ and $\hat{r} = v_0, v_1, \ldots, v_k = l$ is the (undirected) path in $\mathcal{T}_1$ from r to l. Note that upon removal of the edge from r to v_1, the vertex l is in a copy $\mathcal{T}_1^{(\hat{r})}$ or $\mathcal{T}_2^{(\hat{r})}$, respectively, of $\mathcal{T}_1$ or $\mathcal{T}_2$ rooted at v_1. Thus, similarly to the derivation of (1.27)-(1.28), but now utilizing (1.25), we get in case of an out-edge from $\hat{r}$ leading to v_1, that

(A.8)
$$(G_{\mathcal{T}_1})_{\hat{r}+\aleph, l+o_l} = \frac{-1}{\sqrt{d-1}}(G_{\mathcal{T}_1})_{\hat{r}+\aleph, \hat{r}}(G_{\mathcal{T}_1^{(\hat{r})}})_{v_1+\aleph, l+o_l},$$
$$(G_{\mathcal{T}_2})_{\hat{r}, l+o_l} = \frac{-1}{\sqrt{d-1}}(G_{\mathcal{T}_2})_{\hat{r}, \hat{r}}(G_{\mathcal{T}_1^{(\hat{r})}})_{v_1+\aleph, l+o_l},$$

whereas if an in-edge of $\hat{r}$ is connected to v_1, we get instead that

(A.9)
$$(G_{\mathcal{T}_1})_{\hat{r}+\aleph, l+o_l} = \frac{-1}{\sqrt{d-1}}(G_{\mathcal{T}_1})_{\hat{r}+\aleph, \hat{r}+\aleph}(G_{\mathcal{T}_2^{(\hat{r})}})_{v_1, l+o_l},$$
$$(G_{\mathcal{T}_2})_{\hat{r}, l+o_l} = \frac{-1}{\sqrt{d-1}}(G_{\mathcal{T}_2})_{\hat{r}, \hat{r}+\aleph}(G_{\mathcal{T}_2^{(\hat{r})}})_{v_1, l+o_l}.$$

We take $|\cdot|^2$ of both sides, sum over o_l and all vertices $l \in \partial B_k(\hat{r})$ and substitute our expressions (1.30)–(1.31) for $(G_{\mathcal{T}_1})_{\hat{r}+\aleph, \hat{r}}$, $(G_{\mathcal{T}_1})_{\hat{r}+\aleph, \hat{r}+\aleph}$, $(G_{\mathcal{T}_2})_{\hat{r}, \hat{r}+\aleph}$ and $(G_{\mathcal{T}_2})_{\hat{r}, \hat{r}}$, to arrive at the following recursion relations

(A.10)
$$A_{1,k} = YA_{1,k-1} + XA_{2,k-1},$$
$$A_{2,k} = XA_{1,k-1} + YA_{2,k-1},$$

in terms of (X, Y) of (A.1). Solving the recursion (A.10) leads to

(A.11)
$$\begin{bmatrix} A_{1,k} \\ A_{2,k} \end{bmatrix} = \begin{bmatrix} Y & X \\ X & Y \end{bmatrix}^k \begin{bmatrix} A_{1,0} \\ A_{2,0} \end{bmatrix}.$$

The central matrix has eigenvalues $Y \pm X$. From (A.2) we know that $|Y \pm X| < 1$, which upon transforming into the basis of eigenvectors, suffices for the uniform in k boundedness of $A_{1,k}$ and $A_{2,k}$.

In case of the tree $\mathcal{T}$, we derive the corresponding recursion via the Schur complement formula after the removal of the root. Indeed, we have that,

(A.12)
$$\sum_{l \in \partial B_k(\hat{r}), o_l} |(G_{\mathcal{T}})_{r+o_r, l+o_l}|^2 = \frac{d}{d-1}|(G_{\mathcal{T}})_{\hat{r}+o_{\hat{r}}, \hat{r}}|^2 \sum_{l \in \partial B_{k-1}(\hat{r}), o_l} |(G_{\mathcal{T}_1})_{\hat{r}+\aleph, l+o_l}|^2$$
$$+ \frac{d}{d-1}|(G_{\mathcal{T}})_{\hat{r}+o_{\hat{r}}, \hat{r}+\aleph}|^2 \sum_{l \in \partial B_{k-1}(\hat{r}), o_l} |(G_{\mathcal{T}_2})_{\hat{r}, l+o_l}|^2.$$

We conclude by observing that both terms are bounded, uniformly in k, per (A.7). $\qquad\square$

Proceeding to generalize the computation in Lemma A.2 to other higher order terms in perturbation expansions, we first prove a general factorization property of the Green's functions on directed trees.

Lemma A.3. *Denote by p the path in $\mathcal{T}$, $\mathcal{T}_1$, or $\mathcal{T}_2$, between two indices x and y (each being either u or $u+\aleph$ for some vertex u). For $\mathcal{T}_1$ and $\mathcal{T}_2$, specify the root as $\hat{r}$ and denote by $x \wedge y$ the common ancestor of x and y of the largest depth. Let E_v denote the types of the pair of edges in p adjacent to each vertex $v \in p$ (i.e., whether they are in-edges or out-edges). For some explicit function $K(E) \leq \frac{1}{\sqrt{d-1}}$ depending only on*

z, w, d and the edge types and some universal constant C, we have the following bound on the corresponding Green's function

$$(\text{A.13}) \qquad |G_{x,y}| \leq C \prod_{\substack{v \in p \\ v \neq x, y, x \wedge y}} K(E_v),$$

uniformly over $w \in \mathbb{D}$, a fixed compact region in $\mathbb{C}$ (where $v \neq x \wedge y$ means that v is not the ancestor of the points x and y on the path to the root in $\mathcal{T}_1$ and $\mathcal{T}_2$.)

Remark A.4. While this shows that $G_{x,y}$ has 'nice' factorization properties, one has to be careful here. Indeed, (A.13) is more complicated due to the dependence of $G_{x,y}$ on the specific orientations along the path connecting x and y.

Proof. In principle, this factorization is a consequence of the Schur complement formula. However, as mentioned, one has to be aware of the orientations along p while performing a case by case analysis. First, if either the index x or y corresponds to the root vertex, then (A.13) is fairly straightforward. Indeed, as we saw in the recursions (A.8)-(A.9) leading to Lemma A.2, in this case each of the $K(E_v)$ is the absolute value of one of $(G_{\mathcal{T}_1})_{v+\aleph, v}, (G_{\mathcal{T}_1})_{v+\aleph, v+\aleph}, (G_{\mathcal{T}_2})_{v,v}$ or $(G_{\mathcal{T}_2})_{v,v+\aleph}$ times $\frac{1}{\sqrt{d-1}}$. Further, for any intermediate vertex $v \in p$, $v \neq x, y$, the specific Green's function entry to choose among these four options depends only on the orientation of the edges along p immediately before it and immediately after it. Each of the factors corresponding to our pair of fixed indices x and y is covered by a universal finite constant C_0 as long as $|w|$ is uniformly bounded.

Since $\mathcal{T}$ is regular, in that case we shall assume with no loss of generality that x is the root. The more delicate case is when neither of x or y are at the root in either $\mathcal{T}_1$ or $\mathcal{T}_2$. Due to duality, it suffices to proceed with the detailed analysis only for $\mathcal{T}_1$, and we assume first that $x \wedge y \neq \hat{r}$ and denote by $\ell \geq 1$ the length of the path $p_{x \wedge y}$ from $\hat{r}$ to $x \wedge y$. Denoting by c the child of $\hat{r}$ which is on this path, upon removing the edge from $\hat{r}$ to c, we get from (1.26) that

$$(\text{A.14}) \qquad G_{x,y} = G_{x,y}^{(\hat{r})} + \frac{1}{d-1} G_{x,c+\tilde{o}}^{(\hat{r})} (G_{\mathcal{T}_1})_{c+o,c+o} G_{c+\tilde{o},y}^{(\hat{r})}$$

o and $\tilde{o}$ are either $\aleph$ or 0 depending on the orientation of the edge connecting c to $\hat{r}$. The quantity $(G_{\mathcal{T}_1})_{c+o,c+o}$ is either m_∞ or m_∞^{sd} and in either case bounded in absolute value by the universal, finite C_0. Note that the tree $\mathcal{T}_1^{(\hat{r})}$ after the removal of r is either a $\mathcal{T}_1$ rooted at c or a $\mathcal{T}_2$ rooted at c. Thus, following (A.8)-(A.9) yields the product bound of (A.13) for both $G_{x,c+\tilde{o}}^{(\hat{r})}$ and $G_{c+\tilde{o},y}^{(\hat{r})}$. Combining the latter two bounds gives the terms $K(E_v)$ for all vertices $v \in p$ except possibly x, y and $x \wedge y$, as well as the product of $K(E_u)^2 \leq \frac{1}{d-1}$ over all $u \in p_{x \wedge y}$, $u \neq \hat{r}$. Consequently,

$$|G_{x,y}| \leq |G_{x,y}^{(\hat{r})}| + C_0^3 (d-1)^{-\ell} \prod_{\substack{v \in p \\ v \neq x, y, x \wedge y}} K(E_v).$$

By the preceding reasoning, when starting with $G_{x,y}^{(\hat{r})}$ instead of $G_{x,y}$, we get the same bound on the second term on the RHS of (A.14), except for reducing ℓ by one. Thus, setting $\widehat{C}_i = C_0^3 \sum_{k \geq i} (d-1)^{-k}$, $i \geq 0$, and iterating till all that remains from $p_{x \wedge y}$ is the common ancestor $x \wedge y$, we conclude that

$$(\text{A.15}) \qquad |G_{x,y}| \leq |(G_{\mathcal{T}_{(\star)}})_{x,y}| + \widehat{C}_2 \prod_{\substack{v \in p \\ v \neq x, y, x \wedge y}} K(E_v),$$

where $\mathcal{T}_{(\star)}$ is either a $\mathcal{T}_1$ tree or a $\mathcal{T}_2$ tree, now rooted at $x \wedge y$. Turning to deal with $|(G_{\mathcal{T}_{(\star)}})_{x,y}|$, we appeal once more to (1.26), where since now $\hat{r} = x \wedge y$, the first term on the RHS is zero at x, y. Assuming WLOG that $\mathcal{T}_{(\star)}$ is a $\mathcal{T}_1$ tree and with $c_x \neq c_y$ denoting the children of $x \wedge y$ along the path p_x and p_y leading in that tree to x and y respectively, we find that

$$(\text{A.16}) \qquad (G_{\mathcal{T}_{(\star)}})_{x,y} = \frac{1}{d-1} G_{x,c_x+\tilde{o}_1}^{(\star)} (G_{\mathcal{T}_1})_{r+o_1, r+o_2} G_{c_y+\tilde{o}_2, y}^{(\star)},$$

where $G^{(\star)}$ denotes the Green function of $\mathcal{T}_{(\star)} \backslash (x \wedge y)$, and $\tilde{o}_i$, o_i are 0 or $\aleph$, depending on the types of edges connecting c_x and c_y to $x \wedge y$. We use as before the factorization properties of the disjoint sub-trees of $\mathcal{T}_{(\star)} \backslash (x \wedge y)$ containing x and y, to see that

$$|G^{(\star)}_{x,c_x+\tilde{o}_1}| \leq C_0 \prod_{\substack{v \in p_x \\ v \neq x, x \wedge y}} K(E_v), \qquad |G^{(\star)}_{c_y+\tilde{o}_2,y}| \leq C_0 \prod_{\substack{v \in p_y \\ v \neq y, x \wedge y}} K(E_v).$$

Plugging the latter bounds and our usual bound $|(G_{\mathcal{T}_1})_{\hat{r}+o_1, r+o_2}| \leq C_0$ into (A.16) then combining with (A.15), we finally arrive at (A.13) (with $C = \widehat{C}_1$ a universal constant). $\qquad \square$

Building on the preceding factorization property and bounding the expressions $\{K(E_v)\}$, we now generalize Lemma A.2, starting at any $x \in \mathcal{T}_1$ (and not only at its root).

Lemma A.5. *Let $\mathcal{K}_{xy}$ denote the* RHS *of (A.13). Then, for the tree $\mathcal{T}_1$ with root r, uniformly in $k \geq 1$, $w \in \mathbb{D}$ compact and $x \in \mathcal{T}_1$,*

$$(A.17) \qquad \sum_{y \in B_k(x)} |\mathcal{K}_{xy}|^2 \lesssim k.$$

Proof. Let p_x denote the path from x to the root $\hat{r}$. Let v_l denote the vertex on p_x that is of distance l from x. We let $C(v_l)$ denote the children of v_l that are of distance less than $k - l$ from v_l aside from those that can only reached using an edge of p_x. The product expression (A.13) gives us,

$$(A.18) \qquad \sum_{y \in B_k(x)} |\mathcal{K}_{xy}|^2 = \sum_{l=0}^{k} \left(\frac{1}{d-1}\right)^{l-1} \sum_{z \in C(v_l)} |\mathcal{K}_{v_l z}|^2.$$

Now, our main observation is that $|\mathcal{K}_{v_l z}|^2$ can be roughly understood as $|P_{v_l z}|^2$ in a tree that uses v_l as a root. More formally, let $ch_1^l, \ldots, ch_{2d-2}^l$ denote the direct children of v_l in $C(v_l)$ and let $C(ch_m^l)$ denote the children of ch_m^l in $C(v_l)$, while we let $T_{ch_m^l}$ denote the tree with root ch_m^l. We see we have,

$$(A.19) \qquad \sum_{z \in C(v_l)} |\mathcal{K}_{v_l z}|^2 = \sum_{m=1}^{2d-1} \sum_{z \in C(ch_m^l)} |G_{ch_m^l, z}|^2.$$

Each of the internal sums on the right hand side can be explicitly computed and shown to be $\lesssim k - l$ by Lemma A.2. The outer sum will just give a constant factor outside. Finally, we can evaluate the full sum in equation (A.18). The summability of the exponential series will show that this will be bounded by some universal constant times k, as desired. $\qquad \square$

The following lemma allows us to generalize our bounds on the decay of Green's function values with distance to the case of almost tree-like graphs with at most one cycle. The most important part is the relation of the Green's function of the graph to the Green's function of the simply connected cover, which is the infinite tree.

Lemma A.6. *Consider a digraph $\mathcal{G}$ with excess at most 1; let i and j be vertices in $\mathcal{G}$. Let p_1 be the shortest path connecting i and j in $\mathcal{G}$ and let p_2 be the second shortest non-backtracking path connecting i and j in $\mathcal{G}$. Furthermore, let sp_1 be a sequence of symbols fw or rv indicating whether each edge in the path p_1 were traversed in the forward or reverse direction, and let sp_2 be the same object for the path p_2. Given these sequences sp_1 and sp_2, consider the 'pure' infinite tree $\mathcal{T}$ and vertices v_1 and v_2 in $\mathcal{T}$ that are reachable from the root $\hat{r}$ using a sequences of edges that match sp_1 and sp_2, respectively. For example, if sp_1 involved only forward edges, then v_1 should be a vertex that is reachable from $\hat{r}$ using only forward edges.*

We can derive the following estimate on the Green's function extension,

$$|G(Ext(\mathcal{G}, m_\infty, m_\infty))_{i+o_i, j+o_j}| \lesssim |(G_\mathcal{T})_{\hat{r}+o_i, v_1+o_j}| + |(G_\mathcal{T})_{\hat{r}+o_i, v_2+o_j}|$$

$$(A.20) \qquad\qquad\qquad\qquad \lesssim \left[\frac{\max\{|m_\infty|, |m_\infty^{uod}|\}}{\sqrt{d-1}}\right]^{dist(i,j)}.$$

Here, o_i and o_j can be 0 or $\aleph$ to distinguish the two indices that correspond to the same vertex.

Remark A.7. In the top line of the preceding bound, we distinguish the specific edges that are used in the path connecting i and j in order to be able to derive sharp bounds when large neighborhoods are considered.

Proof. The extension of $Ext(\mathcal{G}, m_\infty, m_\infty)$ involves adding an appropriate infinite tree to every boundary vertex of the digraph $\mathcal{G}$ that does not yet have full degree. Since we have assumed that the digraph $\mathcal{G}$ has excess at most 1, this would imply that there is at most one cycle in the extension graph.

Now, we can always relate the Green's function of a graph to its infinite cover as follows. First, we remark that the Green's function is the unique solution to the following system of equations,

$$
\text{(A.21)}
\begin{aligned}
\delta_{ik} + z G_{ik} &= w G_{\aleph+i,k} + \frac{1}{\sqrt{d-1}} \sum_{(i \to j) \in E} G_{\aleph+j,k} \\
z G_{i,\aleph+k} &= w G_{\aleph+i,\aleph+k} + \frac{1}{\sqrt{d-1}} \sum_{(i \to j) \in E} G_{\aleph+j,\aleph+k} \\
z G_{\aleph+i,k} &= \overline{w}\, G_{ik} + \frac{1}{\sqrt{d-1}} \sum_{(j \to i) \in E} G_{j,k} \\
\delta_{ik} + z G_{\aleph+i,\aleph+k} &= \overline{w}\, G_{i,\aleph+k} + \frac{1}{\sqrt{d-1}} \sum_{(j \to i) \in E} G_{\aleph+j,\aleph+k}.
\end{aligned}
$$

We argue that if we have a new graph $\hat{H}$, with Green's function $\hat{G}$, that is a lift of the graph H, with covering map π. Then, we would have the following relationship between the Green's function G of the original graph and the Green's function of the covering graph.

$$
\text{(A.22)}
\begin{aligned}
G_{ij} &= \sum_{y:\pi(y)=j} G_{xy} \\
G_{\aleph+i,j} &= \sum_{y:\pi(y)=j} G_{\aleph+x,y} \\
G_{i,j+\aleph} &= \sum_{y:\pi(y)=j} G_{x,\aleph+y} \\
G_{i+\aleph,j+\aleph} &= \sum_{y:\pi(y)=j} G_{x+\aleph,y+\aleph}.
\end{aligned}
$$

Here, x is a vertex in $\hat{H}$ that maps to i under the map π; the value on the right hand side does not depend on the choice of x due to translation invariance. It is implicitly implied that the above results hold as long as the quantities that appear on the right hand side are convergent. In order to check the above relation, we substitute the proposed equalities in equation (A.22) into the defining equations of equation (A.21) and check that equality holds. This mainly involves realizing that if we choose $\hat{i}$ so that $\pi(\hat{i}) = i$, then we can parameterize the vertices j such that $i \to j$ is an edge by some lifted vertex $\hat{j}$ such that $\pi(\hat{j}) = j$ and $\hat{i} \to \hat{j}$ is a edge in $\hat{H}$. This confirms the first two equations in equation (A.21). We can do the same thing for edges $j \to i$ as well, and this verifies the last two equations of (A.21).

We can use the above formula to compute the Green's function of $Ext(\mathcal{G}, m_\infty, m_\infty)$ by setting $H = Ext(\mathcal{G}, m_\infty, m_\infty)$. This graph has only one cycle due to our assumption on the excess; thus, the covering graph $\hat{H}$ is an infinite tree. The preimage of any vertex $i \in H$ will correspond to the set of non-backtracking paths in $\hat{H}$. $\sum_{y:\pi(y)=j}$ now merely correspond to a 'sum' over non-backtracking paths.

For points i and j that are of distance k in H, the non-backtracking paths all have lengths at least k and the different non-backtracking paths have length that differ by $|C|$ where $|C|$ is the length of the unique simple cycle in H. Recall that p_1 was the shortest path between i and j while p_2 was the second-shortest non-backtracking path between i and j. We have the following structural properties on the paths p_j. We can divide p_2 as $Q_i \cup C \cup Q_o$ where C is the cycle and further shortest backtracking paths p_j are $Q_i \cup C \cup \ldots \cup C \cup Q_o$

where there are $j-1$ copies of C. The path p_1 must be either included completely in Q_o or equal to the union $Q_i \cup Q_o$.

We also have the following property, let v be a fixed lift of i in $\hat{H}$ and w_1 be the lift of j in $\hat{H}$ that is closest to v. Then, if $\hat{p}_1$ is the path between v and w_1, and $\hat{sp}_1$ is the set of labels (either f or b) of the edges of $\hat{p}_1$, then $\hat{sp}$ is the same sequence as sp_1. In addition, let w_2 is the lift of j in $\hat{H}$ that is the second closest to v. If $\hat{p}_2$ is the path between v and w_2, and $\hat{sp}_2$ is the set of labels of this path, then $\hat{sp}_2$ is the same as sp_2.

Thus, we see that

$$|G_{ij}(Ext(\mathcal{G}, m_\infty, m_\infty)| \leq \sum_{n=1}^{\infty} |(G_\mathcal{T})_{v,w_n}| \leq |(G_\mathcal{T})_{v,w_1}| + |(G_\mathcal{T})_{v,w_2}| \frac{1}{1 - \frac{|m_\infty|\vee|m_\infty^{uod}|}{\sqrt{d-1}}}.$$

Here, w_n is the nth closest lift of j to v. To derive the final inequality, we used the decomposition of G found in Lemma A.5. The paths p_j have $(j-1)|C|$ more K factors than p_2. We can thus apply a geometric sum for these terms. $\square$

As a simple corollary of this estimate, we can readily show that we have the same type of estimate for the first order term in perturbation series as in Lemma A.2. Later, we will show more general estimates.

Corollary A.8. *Let $\mathcal{G}$ be a graph with excess at most 1. Fix a vertex i in $\mathcal{G}$ and let R_k denote all of the vertices in $\mathcal{G}$ of distance exactly k from i in $\mathcal{G}$. Then, we have the following estimate,*

$$(A.23) \qquad \sum_{j \in R_k} |G_{ij}(Ext(\mathcal{G}, m_\infty, m_\infty)|^2 \lesssim 1.$$

Proof. A graph with excess 0 would correspond to an infinite tree after the expansion, so this estimate is already known.

We can understand graphs of excess 1 by, again, referring to the infinite cover of the tree. Recall that we have,

$$(A.24) \qquad G_{ij}(Ext(\mathcal{G}, m_\infty, m_\infty)) = \sum_{l=1}^{\infty} (G_\mathcal{T})_{vw_l},$$

where v is a fixed lift of i and w_l is a set of lifts of w in order from closest to most distant. Furthermore if j is of distance k from i then w_1 is of distance k from v and w_2 is of distance $\geq k + |C|$ from v.

The Cauchy-Schwartz inequality will show that,

$$(A.25) \qquad |\sum_{l=1}^{\infty} (G_\mathcal{T})_{vw_l}|^2 \leq \sum_{l=1}^{\infty} |(G_\mathcal{T})_{vw_l}|^2 e^{l\theta} \sum_{l=1}^{\infty} e^{-l\theta} \lesssim |(G_\mathcal{T})_{vw_1}|^2 + |(G_\mathcal{T})_{vw_2}|^2.$$

Setting $\theta = -\frac{1}{2} \log \frac{|m_\infty|\vee|m_\infty^{uod}|}{\sqrt{d-1}}$, makes sure that the infinite sum $\sum_{l=1}^{\infty} e^{-l\theta}$ is finite. Furthermore, $|(G_\mathcal{T})_{vw_l}|^2 e^{l\theta}$ will still have exponential decay since the factor e^θ is chosen to be smaller than the ratio $\frac{(G_\mathcal{T})_{vw_{l-1}}}{(G_\mathcal{T})_{vw_l}}$.

By applying this inequality to all vertices j that are in R_k, we see that we can bound,

$$(A.26) \qquad \sum_{j \in R_k} |G_{ij}(Ext(\mathcal{G}, m_\infty, m_\infty))|^2 \lesssim \sum_{w \in S_k \cup S_{k+|C|}} |(G_\mathcal{T})_{v,w}|^2 \lesssim 1.$$

Here, S_k denotes the elements that are of distance k from v in $\mathcal{T}$. $\square$

A.2. **Comparison of Stieltjes transforms.**

Lemma A.9. *We show the following equality between the Stieltjes transform proposed in equations (1.31) and those coming from free probability computations as in [3]:*

$$m_\mathcal{T}^d(z,w) = \sqrt{d-1} m_\star(\sqrt{d-1}z, \sqrt{d-1}w).$$

Proof. First, notice that if we remove the scaling of edges in the adjacency matrix by $\frac{1}{\sqrt{d-1}}$, then the corresponding equations for m_∞ and $m_{\mathcal{T}}^d$ would be,

(A.27)
$$m_\infty = \frac{z + dm_\infty}{|w|^2 - (z + dm_\infty)(z + (d-1)m_\infty)},$$
$$m_{\mathcal{T}}^d = \frac{z + dm_\infty}{|w|^2 - (z + dm_\infty)^2}.$$

Consider $\mathcal{T}_d$, our infinite d-regular digraph tree; from this, pick a subgraph S_{d-1} in $\mathcal{T}_d$ that is isomorphic to the infinite $d-1$ regular tree $\mathcal{T}_{d-1}$. For every vertex v that is part of S_{d-1}, there will be one edge, i_v, into it and one edge, o_v, out of it that will be part of the original graph $\mathcal{T}_d$, but will not be part of the subgraph S_{d-1}. If we remove the edge o_v, the connected component remaining that is totally disjoint from S_{d-1} will be isomorphic to $\mathcal{T}_1$ (the $2d-1$-ary tree). If we remove i_v, the connected component remaining that is totally disjoint from S_{d-1} will be isomorphic to $\mathcal{T}_2$. For $|z| > (d-1) + |w|^2$, we compute the Green's function of the Hermitized (unscaled) adjacency matrices of $\mathcal{T}_d$ by counting paths and carefully tracking where such paths uses edges of S_{d-1}. (We will naturally consider w and w^* to correspond to a weighted self-loops at the appropriate vertex.)

Consider trying to compute $(G_{\mathcal{T}_d})_{\hat{r}\hat{r}}$ at the root $\hat{r}$. For its first step, it could either take an edge that is part of S_{d-1} or else it could take the outedge $o_{\hat{r}}$. If the path takes the outedge $o_{\hat{r}}$, the path will perform a walk in $\mathcal{T}_1$ before returning to the root by traversing $o_{\hat{r}}$ in the reverse direction. The weighted contribution of paths that take this option would be $\frac{m_\infty(z)}{z}$, corresponding to the weighted contribution of walks that return to the root in $\mathcal{T}_1$. Before the walk takes an edge that is part of S_{d-1}, it could repeat this process indefinitely. Thus, the total weighted contribution of paths that could occur before the walk's first step in S_{d-1} is,

$$\sum_{k=0}^\infty \left(\frac{m_\infty}{z}\right)^k = \frac{1}{1 - \frac{m_\infty}{z}}.$$

If the path takes an edge that is part of S_{d-1} to get to the vertex p, we can repeat a similar argument. Either the next step after this is also a part of S_{d-1} or it could take the inedge i_p. (Note that this is still true if we took the weighted edge corresponding to the self loop.) We can count the contribution of all paths that use the inedge i_p as $\frac{1}{1 - \frac{m_\infty}{z}}$. If we repeat this argument as we traverse edges of S_{d-1} we see that $m_{\mathcal{T}}^d$ can be expressed as follows. Let $\mathcal{S}_k$ denote the total weight of all paths in S_{d-1} that use k edges.

(A.28)
$$m_{\mathcal{T}}^d = \frac{1}{z} \sum_{k=0}^\infty \frac{\mathcal{S}_k}{z^k} \left(\frac{1}{1 - \frac{m_\infty(z)}{z}}\right)^{k+1}.$$

Thus,

(A.29)
$$m_{\mathcal{T}}^d = m_{\mathcal{T}}^{d-1}(z - m_\infty(z)).$$

If we return to the equations in (A.27), we find that,

(A.30) $\quad \dfrac{1}{m_{\mathcal{T}}^d} = \dfrac{|w|^2}{z + dm_\infty} - (z + dm_\infty) = \dfrac{|w|^2}{z + dm_\infty(z)} - (z + (d-1)m_\infty) - m_\infty(z) = -m_\infty(z) + \dfrac{1}{m_\infty}.$

Solving the quadratic gives,

(A.31)
$$m_\infty = \frac{-1 + \sqrt{1 + 4m_d^2}}{2m_d}.$$

Returning to equation (A.29), we see that,

(A.32)
$$m_{\mathcal{T}}^d(z) = m_{\mathcal{T}}^{d-1}\left(z - \frac{2m_{\mathcal{T}}^d}{1 + \sqrt{1 + 4(m_{\mathcal{T}}^d)^2}}\right).$$

This is the same recursion as in equation (4.6) of [3]. Notice that though we only confirmed the identity for $|z| \geq (d-1) + |w|^2$, the identity would hold for all values of z by analytic continuation. To fully confirm the equality of Green's functions, it suffices to show the equality of Green's functions when $d = 1$.

When $d = 1$, we have to solve the quadratic,

$$\tag{A.33} m_\infty^2 + m_\infty \left[z + \frac{1}{z} - \frac{|w|^2}{z} \right] + 1 = 0.$$

Setting the quantity $A = z + \frac{1}{z} - \frac{|w|^2}{z}$, we see that we have,

$$\tag{A.34} m_\infty = \frac{-A - \sqrt{A^2 - 4}}{2},$$

and, using (A.30), we see that,

$$
\begin{aligned}
m_T^1 &= -\frac{1}{\sqrt{A^2 - 4}} = -\frac{1}{\sqrt{\left(z + \frac{1}{z} - \frac{|w|^2}{z} \right)^2 - 4}} \\
&= -\frac{z}{\sqrt{(z^2 + 1 - |w|^2)^2 - 4z^2}} \\
&= -\frac{z}{\sqrt{(z^2 + 2z + 1 - |w|^2)(z^2 - 2z + 1 - |w|^2)}} \\
&= -\frac{z}{\sqrt{((z+1)^2 - |w|^2)((z-1)^2 - |w|^2)}} \\
&= -\frac{z}{\sqrt{(z + 1 - |w|)(z + 1 + |w|)(z - 1 - |w|)(z - 1 + |w|)}} \\
&= -\frac{z}{\sqrt{(z^2 - (1 - |w|)^2)(z^2 - (1 + |w|)^2)}}.
\end{aligned}
\tag{A.35}
$$

This Green's function can be checked to correspond to the density of equation (4.1) in [3]. $\square$

References

[1] G. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Stud. Adv. Math.* Cambridge Univ.Press , Cambridge, UK, 2010.

[2] A. Basak, N. Cook, and O. Zeitouni. Circular law for the sum of random permutation matrices. *Electron. Jour. Probab.*, 23(33):1–51, 2018.

[3] A. Basak and A. Dembo. Limiting spectral distribution of sums of unitary and orthogonal matrices. *Electron. Commun. Probab.*, 18(69):1–19, 2013.

[4] A. Basak, E. Paquette, and O. Zeitouni. Regularization of non-normal matrices by Gaussian noise – the banded Toeplitz and twisted Toeplitz cases. *Forum Math. Sigma*, 7(e3):72, 2019.

[5] A. Basak, E. Paquette, and O. Zeitouni. Spectrum of random perturbations of Toeplitz matrices with finite symbols. *Trans. Amer. Math. Soc.*, 373(7):4999–5023, 2020.

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

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

[8] C. Bordenave and D. Chafai. Around the circular law. *Probability Surveys*, 9:1–89, 2012.

[9] N. Cook. The circular law for random regular digraphs. *Ann. Inst. H. Poincaré Probab. Statist.*, 55(4):2111–2167, 2019.

[10] S. Coste. The spectral gap of sparse random digraphs. *Ann. Inst. H. Poincaré Probab. Statist.*, 57(2):644–684, 2021.

[11] S. Coste, G. Lambert, and Y. Zhu. The characteristic polynomial of sums of random permutations and regular digraphs. *Int. Math. Res. Notices*, 2024:2461–2510, 2024.

[12] L. Erdos and H.-T. Yau. *Dynamical Approach to Random Matrix Theory*. AMS and Courant Institute of Mathematical Sciences, 2017.

[13] V. Girko. The circular law. *Teor. Veroyatnost. i Primenen*, 29(4):669–679, 1984.

[14] A. Guoinnet, M. Krishnapur, and O. Zeitouni. The single ring theorem. *Ann. of Math.*, 174(2):1189–1217, 2011.

[15] A. Guoinnet, P. Wood, and O. Zeitouni. Convergence of spectral measure of non normal matrices. *Proc. Amer. Math. Soc.*, 142(2):667–679, 2014.

[16] U. Haagerup and F. Larsen. Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras. *J. Funct.Anal.*, 176:331–367, 2000.

[17] J. Huang. Invertibility of adjacency matrices for random d-regular graphs. *Duke Math. Jour.*, 170(18):3977 –4032, 2021.

[18] J. Huang and H.-T. Yau. Spectrum of random d-regular graphs up to the edge. *Comm. Pure Appl. Math*, 77:1635–1723, 2024.

[19] S. Janson. Random regular graphs:asymptotic distributions and contiguity. *Combin. Probab. Comput.*, 4(4):369–405, 1995.

[20] A. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegarmann, and P. Youssef. The smallest singular value of a shifted d-regular random square matrix. *Prob.Thr. Rel. Fields*, 173:1301–1347, 2019.

[21] A. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef. Circular law for sparse random regular digraphs. *J. Euro. Math. Soc.*, 23(2):467–501, 2020.

[22] A. Litvak and K. Tikhomirov. Singularity of sparse Bernoulli matrices. *Duke Math. Jour.*, 171(5):1135–1233, 2022.

[23] C. Male. Traffic distributions and independence: Permutation invariant random matrices and the three notions of independence. *Memoirs of the AMS*, 267:1300, 2020.

[24] F. L. Metz, I. Neri, and T. Rogers. Spectral theory of sparse non-Hermitian random matrices. *J. Phys. A: Math. Theo.*, 52(43):434003, 2019.

[25] H. H. Nguyen and A. Pan. A note on the singularity probability of random directed d-regular graphs. *Eur. J. Combinatorics*, 122:104039, 2024.

[26] M. Rudelson and R. Vershynin. Invertibility of random matrices: unitary and orthogonal perturbations. *Bull. Amer. Math. Soc.*, 27(2):293–338, 2014.

[27] A. Sah, J. Sahasrabudhe, and M. Sawhney. The limiting spectral law for sparse iid matrices. *Forum Math. Pi (or arXiv:2310.17635)*, 13, 2025.

[28] J. Sjöstrand and M. Vogel. General Toeplitz matrices subject to Gaussian perturbations. *Ann. Henri Poincaré*, 22:49–81, 2021.

[29] P. Sniady. Random regularization of Brown spectral measure. *J. Funct. Anal.*, 193:291–313, 2002.

[30] T. Tao and V. Vu. Random matrices: universality of ESDs and the circular law,. *Ann. Probab.*, 38(5):2023–2065, 2010.

ARKA ADHIKARI

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND-COLLEGE PARK, COLLEGE PARK, MD, USA
    Email address: `arkaa@umd.edu`

AMIR DEMBO

DEPARTMENT OF MATHEMATICS AND STATISTICS, STANFORD UNIVERSITY, STANFORD, CA, USA
    Email address: `adembo@stanford.edu`