Annals of Mathematics **189** (2019), 753–835
https://doi.org/10.4007/annals.2019.189.3.3

Quantum ergodicity on graphs: From spectral to spatial delocalization

By Nalini Anantharaman and Mostafa Sabri

Abstract

We prove a quantum-ergodicity theorem on large graphs, for eigenfunctions of Schrödinger operators in a very general setting. We consider a sequence of finite graphs endowed with discrete Schrödinger operators, assumed to have a local weak limit. We assume that our graphs have few short loops, in other words that the limit model is a random rooted *tree* endowed with a random discrete Schrödinger operator. We show that an absolutely continuous spectrum for the infinite model, reinforced by a good control of the moments of the Green function, imply "quantum ergodicity," a form of spatial delocalization for eigenfunctions of the finite graphs approximating the tree. This roughly says that the eigenfunctions become equidistributed in phase space. Our result applies, in particular, to graphs converging to the Anderson model on a regular tree, in the regime of extended states studied by Klein and Aizenman-Warzel.

1. **Introduction**

1.1. *The problem.* Consider a very large, but finite, graph $G = (V, E)$. Are the eigenfunctions of its adjacency matrix *localized*, or *delocalized*? These words are used in a variety of contexts, with several different meanings.

For discrete Schrödinger operators on infinite graphs (e.g., for the celebrated Anderson model describing the metal-insulator transition), localization can be understood in a spectral, spatial or dynamical sense. Given an interval $I \subset \mathbb{R}$, one can consider

- *spectral localization*: pure point spectrum in I,
- *exponential localization*: the corresponding eigenfunctions decay exponentially,
- *dynamical localization*: an initial state with energy in I that is localized in a bounded domain essentially stays in this domain as time goes on.

Keywords: Quantum ergodicity, large graphs, delocalization
AMS Classification: Primary: 58J51; Secondary: 60B20, 81Q10.

On the other hand, delocalization may be understood at different levels:

- *spectral delocalization*: purely absolutely continuous spectrum in I,
- *ballistic transport*: wave packets with energies in I spread on the lattice at a specific (ideally, linear) rate as time goes on.

In this paper we want to discuss a notion of *spatial delocalization*. Since the wavefunctions corresponding to the absolutely continuous spectrum are not square summable, a natural interpretation of spatial delocalization is to consider a sequence of growing "boxes" or finite graphs (G_N) approximating the infinite system in some sense, and to ask if the eigenfunctions on (G_N) become delocalized as $N \to \infty$. Can they concentrate on small regions or, on the contrary are they uniformly distributed over (G_N)? Large, finite graphs are also a subject of interest on their own. Actually, an infinite system is often an idealized version of a large finite one.

Localization/delocalization of eigenfunctions is believed to bear some relation with *spectral statistics*: localization is supposedly associated with Poissonian spectral statistics, whereas delocalization should be associated with Random Matrix statistics (GOE/GUE). In the field of quantum chaos, the former notion is often associated with *integrable dynamics* and the latter with *chaotic dynamics* [16], [17], [18]. However, specific examples show that the relation is not so straightforward [38], [39], [34]. Understanding how far one can push these ideas is one amongst many reasons for studying models of large graphs [31], [41], [42].

Recently, the question of delocalization of eigenfunctions of large matrices or large graphs has been a subject of intense activity. Let us mention several ways of testing delocalization that have been used. Let M_N be a large symmetric matrix of size $N \times N$, and let $(\psi_j)_{j=1}^N$ be an orthonormal basis of eigenfunctions. The eigenfunction ψ_j defines a probability measure $\sum_{x=1}^N |\psi_j(x)|^2 \delta_x$. The goal is to compare this probability measure with the uniform measure, which puts mass $1/N$ on each point:

- ℓ^∞ *norms*: Can we have a pointwise upper bound on $|\psi_j(x)|$, in other words, is $\|\psi_j\|_\infty$ small, and how small compared with $1/\sqrt{N}$?
- ℓ^p *norms*: Can we compare $\|\psi_j\|_p$ with $N^{1/p-1/2}$? In [24], a state ψ_j is called *non-ergodic* (and *multi-fractal*) if $\|\psi_j\|_p$ behaves like $N^{f(p)}$ with $f(p) \neq 1/p - 1/2$. Related criteria appear in [2].
- *Scarring*: Can we have full concentration ($\sum_{x\in\Lambda} |\psi_j(x)|^2 \geq 1-\epsilon$) or partial concentration ($\sum_{x\in\Lambda} |\psi_j(x)|^2 \geq \epsilon$) with Λ a set of "small" cardinality? We borrow the term "scarring" from the term used in the theory of quantum chaos [38].
- *Quantum ergodicity*: Given a function $a : \{1, \dots, N\} \longrightarrow \mathbb{C}$, can we compare $\sum_x a(x)|\psi_j(x)|^2$ with $\frac{1}{N}\sum_x a(x)$? This criterion, borrowed again from

quantum chaos, was applied to discrete regular graphs in [7], [5]. Quantum ergodicity means that the two averages are close for *most* j. If they are close for *all* j, one speaks of *quantum unique ergodicity.*

As was demonstrated in a recent series of papers, adding some randomness may allow the problem to be settled completely. For instance, *almost sure* optimal ℓ^∞-bounds and quantum unique ergodicity for various models of *random* matrices and *random* graphs, such as Wigner matrices, sparse Erdös-Rényi graphs, random regular graphs of slowly increasing or bounded degrees were obtained in [28], [29], [20], [27], [13], [11], [12]. The invariance of the probability distribution under certain elementary transformations plays an important role. The completely different point of view that we adopt is to consider deterministic graphs and to prove delocalization as resulting directly from the geometry of the graphs. Up to now, in this deterministic setting, only eigenfunctions of the adjacency matrix of regular graphs have been treated, taking advantage of the completely explicit Fourier analysis on regular trees. The papers [7], [21], [5] give various proofs of quantum ergodicity; the paper [22] proves the absence of scarring on sets of cardinality $N^{1-\epsilon}$ and also contains (although not stated) a logarithmic upper bound on the ℓ^∞ norms.

The aim of this paper is to prove a *quantum ergodicity theorem* for eigenfunctions of discrete Schrödinger operators on quite general large graphs. As we will see, a particularly interesting point of our result is that it gives a direct relation between *spectral delocalization* of infinite systems and *spatial delocalization* of large finite systems. Our result may be summarized as follows (with proper additional assumptions to be described later):

> "*If a large finite system is close* (*in the Benjamini-Schramm topology*) *to an infinite system having a purely absolutely continuous spectrum in an interval* I, *then the eigenfunctions* (*with eigenvalues lying in* I) *of the finite system satisfy quantum ergodicity.*"

1.2. *The results.* Consider a sequence of connected graphs without self-loops and multiple edges $(G_N)_{N\in\mathbb{N}}$. We assume each vertex has at least three neighbors. It will be convenient to write G_N as a quotient of a tree $\widetilde{G_N}$ by a group of automorphisms Γ_N, that is, $G_N = \Gamma_N\backslash\widetilde{G_N}$, where Γ_N acts freely on the vertices of $\widetilde{G_N}$; i.e., given $v \in \widetilde{G_N}$, $\gamma_1 v = \gamma_2 v$ implies $\gamma_1 = \gamma_2$. In other words, $\widetilde{G_N}$ is the "universal cover" of G_N. We will work under the assumption that the degree of $\widetilde{G_N}$ is everywhere smaller than some fixed D.

We denote by $\widetilde{V_N}$ and $\widetilde{E_N}$ the set of vertices and edges of $\widetilde{G_N}$, respectively. We denote by V_N and E_N the vertices and edges of G_N, respectively. We assume $|V_N| = N$ and work in the limit $N \longrightarrow \infty$.

Define the adjacency operator $\widetilde{\mathcal{A}}_N : \mathbb{C}^{\widetilde{G_N}} \to \mathbb{C}^{\widetilde{G_N}}$ by

$$(\widetilde{\mathcal{A}}_N f)(v) = \sum_{w \sim v} f(w),$$

where $v \sim w$ means v and w are nearest neighbors. The operator $\widetilde{\mathcal{A}}_N$ is bounded on $\ell^2(\widetilde{G_N})$. It also preserves the space of Γ_N-invariant functions on $\widetilde{V_N}$, in other words it defines an operator on $\ell^2(V_N)$, which we denote by $\mathcal{A}_N$. (We will drop the index N and write $\widetilde{\mathcal{A}}, \mathcal{A}$ when no confusion may arise.) Consider a bounded function $\widetilde{W_N} : \widetilde{V_N} \longrightarrow \mathbb{R}$ such that $\widetilde{W_N}(\gamma \cdot v) = \widetilde{W_N}(v)$ for all $\gamma \in \Gamma_N$. The operator of multiplication by $\widetilde{W_N}$ is bounded on $\ell^2(\widetilde{G_N})$; it also preserves the space of Γ_N-invariant functions on $\widetilde{V_N}$, thus it defines an operator on $\ell^2(V_N)$, which we denote by W_N. We define the discrete Schrödinger operators $\widetilde{H}_N = \widetilde{\mathcal{A}}_N + \widetilde{W_N}$ and $H_N = \mathcal{A}_N + W_N$. The central object of our study are the eigenfunctions of H_N, and their behaviour (localized/delocalized) as $N \longrightarrow +\infty$. The fact that Γ_N acts freely implies that H_N is symmetric (self-adjoint) on $\ell^2(V_N)$.

For comfort, we will always work under the assumption that W_N takes values in some fixed interval $[-A, A]$. This implies that the spectrum of all operators we will encounter is contained in some fixed interval $I_0 = [-A - D, A + D]$.

We define the Laplacian $P_N : \mathbb{C}^{V_N} \to \mathbb{C}^{V_N}$ by

$$(P_N f)(x) = \frac{1}{d_N(x)} \sum_{y \sim x} f(y), \tag{1.1}$$

where $d_N(x)$ stands for the number of neighbors of x. If we introduce the positive measure on V_N assigning to x the weight $d_N(x)$, then P_N is self-adjoint on $\ell^2(V_N, d_N)$.

We shall assume the following conditions on our sequence of graphs:

(EXP) The sequence (G_N) forms an expander family. By this we mean that the Laplacian P_N has a uniform spectral gap in $\ell^2(V_N, d_N)$. More precisely, the eigenvalue 1 of P_N is simple, and the spectrum of P_N is contained in $[-1 + \beta, 1 - \beta] \cup \{1\}$, where $\beta > 0$ is independent of N.

Note that 1 is always an eigenvalue, corresponding to constant functions. Our assumption implies, in particular, that each G_N is connected and non-bipartite. It is well known that a uniform spectral gap for P_N is equivalent to a Cheeger constant bounded away from 0; see, for instance, [25, §3].

Our second assumption is that (G_N) has few short loops:

(BST) For all $r > 0$,

$$\lim_{N \to \infty} \frac{|\{x \in V_N : \rho_{G_N}(x) < r\}|}{N} = 0,$$

where $\rho_{G_N}(x)$ is the *injectivity radius* at x, i.e., the largest ρ such that the ball $B_{G_N}(x,\rho)$ is a tree.

The general theory of Benjamini-Schramm convergence (or local weak convergence), briefly recalled in Appendix A, allows us to assign a limit object to the sequence (G_N, W_N), which is a probability distribution carried on *trees*. More precisely, up to passing to a subsequence, assumption (BST) above is equivalent to the following assumption:

(BSCT) The sequence (G_N, W_N) has a local weak limit $\mathbb{P}$ that is concentrated on the set of (isomorphism classes of) colored rooted *trees*, denoted $\mathscr{T}_*^{D,A}$.

Assumption (BSCT) says that (G_N, W_N) converges in a distributional sense to a random system of rooted trees $\{[\mathcal{T}, o]\}$, endowed with a map $\mathcal{W} : \mathcal{T} \longrightarrow \mathbb{R}$. More precisely, the empirical measure of (G_N, W_N), defined by choosing a root $x \in V_N$ uniformly at random, converges weakly to a probability measure $\mathbb{P}$ concentrated on trees.

If $[\mathcal{T}, o, \mathcal{W}] \in \mathscr{T}_*^{D,A}$ and $\mathcal{A}$ is the adjacency matrix of $\mathcal{T}$, we denote by $\mathcal{H} = \mathcal{A} + \mathcal{W}$ the limiting random Schrödinger operator, which is self-adjoint on $\ell^2(\mathcal{T})$.

Call $(\lambda_j^{(N)})_{j=1}^N$ the eigenvalues of H_N on $\ell^2(V_N)$. Assumption (BSCT) implies the convergence of the empirical law of eigenvalues: for any continuous $\chi : \mathbb{R} \longrightarrow \mathbb{R}$, we have

$$\frac{1}{N}\sum_{j=1}^{N} \chi(\lambda_j^{(N)}) \underset{N\longrightarrow+\infty}{\longrightarrow} \mathbb{E}\left(\langle \delta_o, \chi(\mathcal{H})\delta_o\rangle\right) =: \rho(\chi); \tag{1.2}$$

see Remark A.3. Here $\mathbb{E}$ is the expectation with respect to $\mathbb{P}$, that is,

$$\mathbb{E}(f) = \int_{\mathscr{T}_*^{D,A}} f([\mathcal{T}, o, \mathcal{W}])\, \mathrm{d}\,\mathbb{P}([\mathcal{T}, o, \mathcal{W}]).$$

The measure ρ is called the *integrated density of states* in the theory of random Schrödinger operators.

We need some notation for our last assumption. Let $[\mathcal{T}, o, \mathcal{W}] \in \mathscr{T}_*^{D,A}$. Given $x, y \in \mathcal{T}$, and $\gamma \in \mathbb{C} \setminus \mathbb{R}$, we introduce the Green function

$$\mathcal{G}^{\gamma}(x,y) = \langle \delta_x, (\mathcal{H} - \gamma)^{-1}\delta_y\rangle_{\ell^2(\mathcal{T})}.$$

Given $v, w \in \mathcal{T}$ with $v \sim w$, we denote by $\mathcal{T}^{(v|w)}$ the tree obtained by removing from the tree $\mathcal{T}$ the branch emanating from v that passes through w. We define the restriction $\mathcal{H}^{(v|w)}(u,u') = \mathcal{H}(u,u')$ if $u, u' \in \mathcal{T}^{(v|w)}$ and zero otherwise. The corresponding Green function is denoted by $\mathcal{G}^{(v|w)}(\cdot,\cdot;\gamma)$. We then put $\hat{\zeta}_w^{\gamma}(v) := -\mathcal{G}^{(v|w)}(v,v;\gamma)$.

(Green) There is a non-empty open set I_1 such that for all $s > 0$, we have

$$\sup_{\lambda\in I_1,\eta_0\in(0,1)} \mathbb{E}\left(\sum_{y:y\sim o} |\operatorname{Im}\hat{\zeta}_o^{\lambda+i\eta_0}(y)|^{-s}\right) < \infty.$$

To understand (Green), define the (rooted) spectral measure of $[\mathcal{T}, o, \mathcal{W}] \in \mathscr{T}_*^{D,A}$ by

$$\mu_o(J) = \langle \delta_o, \chi_J(\mathcal{H})\delta_o\rangle \qquad \text{for Borel } J \subseteq \mathbb{R}. \tag{1.3}$$

Assumption (Green) implies that $\sup_{\lambda\in I_1,\eta_0>0} \mathbb{E}(|\mathcal{G}^\gamma(o,o)|^2) < \infty$; see Remark A.4. As shown in [32], this implies that for $\mathbb{P}$-a.e. $[\mathcal{T}, o, \mathcal{W}] \in \mathscr{T}_*^{D,A}$, the spectral measure μ_o is absolutely continuous in I_1, with density $\frac{1}{\pi}\operatorname{Im}\mathcal{G}^{\lambda+i0}(o,o)$. Hence, (Green) implies that $\mathbb{P}$-a.e. operator $\mathcal{H}$ has a purely absolutely continuous spectrum in I_1. This is a natural assumption since our aim is to prove delocalization properties of eigenfunctions.

Now let $(\psi_j^{(N)})_{j=1}^N$ be an orthonormal basis of $\ell^2(V_N)$ consisting of eigenfunctions of H_N. Pick $j \in \{1, \ldots, N\}$. The problem of quantum ergodicity is to understand if the probability measure $\sum_{x\in V_N} |\psi_j^{(N)}(x)|^2\delta_x$ on V_N is "localized" (essentially carried by $o(N)$ vertices) or "delocalized" (ideally, close to the uniform measure on V_N, or maybe, to some other natural measure on V_N, comparable to the uniform measure). More generally, we want to know if the correlations $\overline{\psi_j^{(N)}(x)}\psi_j^{(N)}(y)$, for x and $y \in V_N$ at some fixed distance, approach some limiting object. From a mathematical point of view, the question was addressed in [7], [21] for eigenfunctions of the adjacency matrix of large deterministic *regular* graphs, and for the adjacency matrix of *random* regular graphs or Erdös-Rényi graphs in the recent works [27], [13], [11], [12]. The main motivation of our paper is to extend the results of [7] to disordered systems, that is, to non-regular graphs, possibly with a potential on the vertices or weights on the edges. This necessarily requires a different method from that of [7], which was specific to regular graphs. New methods to prove quantum ergodicity were already explored in [5]. We insist on the fact that, contrary to [27], [13], [11], [12], [30], our sequence of graphs and potentials are deterministic. The results may, in particular, be applied to random graphs and/or random potentials, provided one knows that Assumptions (EXP), (BSCT) and (Green) hold true for some realizations. We discuss the relation with existing work more extensively in Section 1.5.

Let us state the main abstract result; its concrete meaning will be explored afterwards. For $x, y \in \widetilde{V}_N$, and $\gamma \in \mathbb{C}\setminus\mathbb{R}$, we introduce the lifted Green function

$$\tilde{g}_N^\gamma(x,y) = \langle \delta_x, (\widetilde{H}_N - \gamma)^{-1}\delta_y\rangle_{\ell^2(\widetilde{V}_N)}. \tag{1.4}$$

Recall that we write G_N as a quotient $\Gamma_N\backslash\widetilde{G}_N$, where $\widetilde{G}_N$ is a tree. We denote by $\mathcal{D}_N$ a fundamental domain of the action of Γ_N on the vertices of $\widetilde{G}_N$. Thus

$\mathcal{D}_N$ contains N vertices of $\widetilde{G}_N$, each of them projecting to a distinct vertex of G_N.

Let I_1 be the open set of Assumption (Green), and let us fix an interval I (or finite union of intervals) such that $\bar{I} \subset I_1$.

THEOREM 1.1. *Assume that the graphs G_N and the potentials W_N satisfy* (BSCT), (EXP) *and* (Green). *Call $(\lambda_j^{(N)})_{j=1}^N$ the eigenvalues of the Schrödinger operator H_N on $\ell^2(V_N)$, and let $(\psi_j^{(N)})_{j=1}^N$ be a corresponding orthonormal eigenbasis.*

For each N, let $a = a_N$ be a function on V_N with $\sup_N \sup_{x\in V_N} |a_N(x)| \leq 1$. For $\gamma \in \mathbb{C} \setminus \mathbb{R}$, define $\langle a\rangle_\gamma = \sum_{x\in V_N} a(x)\Phi_\gamma^N(\tilde{x},\tilde{x})$, where $\Phi_\gamma^N(\tilde{x},\tilde{x}) = \frac{\operatorname{Im}\tilde{g}_N^\gamma(\tilde{x},\tilde{x})}{\sum_{\tilde{x}\in\mathcal{D}_N}\operatorname{Im}\tilde{g}_N^\gamma(\tilde{x},\tilde{x})}$. Then

$$\lim_{\eta_0\downarrow 0}\lim_{N\to+\infty}\frac{1}{N}\sum_{\lambda_j^{(N)}\in I}\left|\sum_{x\in V_N} a(x)|\psi_j^{(N)}(x)|^2 - \langle a\rangle_{\lambda_j^{(N)}+i\eta_0}\right| = 0.$$

Here, $\tilde{x}$ is a lift of $x \in V_N$ in the universal cover $\widetilde{V}_N$.

COROLLARY 1.2. *Under the same assumptions, for any $\epsilon > 0$, we have*

$$\frac{1}{N}\#\left\{\lambda_j^{(N)} \in I : \left|\sum_{x\in V_N} a(x)|\psi_j^{(N)}(x)|^2 - \langle a\rangle_{\lambda_j^{(N)}+i\eta_0}\right| > \epsilon\right\} \underset{N\to+\infty,\,\eta_0\downarrow 0}{\longrightarrow} 0.$$

More generally, we have the following result on eigenfunction correlators, which says that $\overline{\psi_j(x)}\psi_j(y)$ "approaches" the function $\Phi^N_{\lambda_j+i0}(\tilde{x},\tilde{y})$ defined in (1.5). For technical reasons we have to assume the (ψ_j) are real-valued. More precisely, we need $\overline{\psi_j(x)}\psi_j(y)$ to be real for any $j = 1,\ldots,N$ and $x, y \in V_N$ with $x \sim y$.

THEOREM 1.3. *Assume* (G_N, W_N) *satisfies* (BSCT), (EXP) *and* (Green). *Call $(\lambda_j^{(N)})_{j=1}^N$ the eigenvalues of H_N on $\ell^2(V_N)$, and let $(\psi_j^{(N)})_{j=1}^N$ be a corresponding orthonormal eigenbasis. Assume the $(\psi_j)_{j=1}^N$ are real-valued.*

Fix $R \in \mathbb{N}$. For each N, let $\mathbf{K} = \mathbf{K}_N$ be an operator on $\ell^2(V_N)$ whose kernel $K = K_N : V_N \times V_N \longrightarrow \mathbb{C}$ is such that $K(x,y) = 0$ for $d(x,y) > R$. (*In other words, K is supported at distance $\leq R$ from the diagonal.*) *Assume that $\sup_N \sup_{x,y\in V_N} |K_N(x,y)| \leq 1$.*

For $\gamma \in \mathbb{C}\setminus\mathbb{R}$, define

(1.5)

$$\langle \mathbf{K}\rangle_\gamma = \sum_{\tilde{x}\in\mathcal{D}_N,\tilde{y}\in\widetilde{V}_N} K(\tilde{x},\tilde{y})\Phi_\gamma^N(\tilde{x},\tilde{y}), \quad \textit{where} \quad \Phi_\gamma^N(\tilde{x},\tilde{y}) = \frac{\operatorname{Im}\tilde{g}_N^\gamma(\tilde{x},\tilde{y})}{\sum_{\tilde{x}\in\mathcal{D}_N}\operatorname{Im}\tilde{g}_N^\gamma(\tilde{x},\tilde{x})}.$$

Then

$$\lim_{\eta_0\downarrow 0}\lim_{N\to+\infty}\frac{1}{N}\sum_{\lambda_j^{(N)}\in I}\left|\langle\psi_j^{(N)},\mathbf{K}\psi_j^{(N)}\rangle_{\ell^2(V_N)}-\langle\mathbf{K}\rangle_{\lambda_j^{(N)}+i\eta_0}\right|=0.$$

The "kernel" above is the matrix of $\mathbf{K}$ in the basis (δ_x), i.e., $K(x,y)=\langle\delta_x,\mathbf{K}\delta_y\rangle_{\ell^2(V_N)}$. To define (1.5) properly, we lift K to $\widetilde{V}_N\times\widetilde{V}_N$ by letting

$$K(\tilde{x},\tilde{y})=K(x,y)\mathbb{1}_{\mathrm{dist}_{\widetilde{G}_N}(\tilde{x},\tilde{y})\le R} \tag{1.6}$$

if $x,y\in V_N=\Gamma_N\backslash\widetilde{V}_N$ are the projections of $\tilde{x},\tilde{y}\in\widetilde{V}_N$.

If we know, in addition, that $\rho(\partial I_1)=0$, where ρ is the integrated density of states measure (1.2), then our main theorems hold with I replaced by I_1; see the end of Section 10. Note that if (Green) holds on $\overline{I_1}$, then $\rho(\partial I_1)=0$.

Although we tend to skip it from the notation, the "observables" $\mathbf{K}$ and a necessarily depend on N. On the other hand, they do not depend on j, the index of the eigenfunction. (They are actually allowed to depend on $\lambda_j^{(N)}$ in the proof, but this dependence cannot be wild — it has to be at least continuous.) We interpret Corollary 1.2 as follows: for a given observable a, the average $\sum_{x\in V_N}a(x)|\psi_j^{(N)}(x)|^2$ is close to $\langle a\rangle_{\lambda_j^{(N)}+i\eta_0}$ for most indices j. It follows similarly from Theorem 1.3 that $\sum_{x,y\in V_N}K(x,y)\overline{\psi_j^{(N)}(x)}\psi_j^{(N)}(y)$ is close to $\langle\mathbf{K}\rangle_{\lambda_j^{(N)}+i\eta_0}$ for most j. One of the subtleties of the result is that the indices j for which this holds may a priori depend on the observables a, $\mathbf{K}$. If we wanted to have a common set of indices j that do the job for all observables (whose number is exponential in N), we would need to have an exponential rate of convergence in Theorems 1.1 and 1.3. As is seen in the case of regular graphs and $W=0$ [5], our proof gives a rate that is at best a negative power of a parameter related to the girth, which is typically of order $\log N$. So, the result is far from showing that $|\psi_j^{(N)}(x)|^2$ is close to the uniform measure in total variation.

Note the presence of the extra parameter η_0, in comparison with the case of regular graphs [7], [5]. This is due to the fact that, generally speaking, the quantities $\langle a\rangle_{\lambda_j^{(N)}+i\eta_0}$ and $\langle\mathbf{K}\rangle_{\lambda_j^{(N)}+i\eta_0}$ are not necessarily bounded as $\eta_0\downarrow 0$ for fixed N. They will however stay bounded in the limits $N\to+\infty$ followed by $\eta_0\downarrow 0$ (as a result of (A.14) and (Green)).

1.3. *Understanding the weighted averages.* In order to clarify the relevance of Theorems 1.1 and 1.3, we now investigate the meaning of the quantities $\langle a\rangle_{\lambda+i\eta_0}$ and $\langle\mathbf{K}\rangle_{\lambda_j+i\eta_0}$. Let us start with Theorem 1.1. A good illustration is to choose $a_N=\mathbb{1}_{\Lambda_N}$, the characteristic function of a set $\Lambda_N\subset V_N$ of size $\approx\alpha N$ for some $\alpha\in(0,1)$, say $\alpha=\frac{1}{2}$.

In the special case where (G_N) is regular and $H_N = \mathcal{A}_N$, and also for the anisotropic model treated in [5], the Green function $\tilde{g}_N^{\gamma}(\tilde{x},\tilde{y})$ does not depend on N, as it coincides with the limiting Green function $\mathcal{G}^{\gamma}(\tilde{x},\tilde{y})$. Moreover, $\mathcal{G}^{\gamma}(\tilde{x},\tilde{x}) = \mathcal{G}^{\gamma}(o,o)$ for all $\tilde{x} \in \mathcal{D}_N$. It follows that $\langle \mathbb{1}_{\Lambda_N}\rangle_{\lambda_j+i\eta_0} = \sum_{x\in\Lambda_N} \frac{\mathcal{G}^{\lambda_j+i\eta_0}(o,o)}{N\mathcal{G}^{\lambda_j+i\eta_0}(o,o)} = \alpha$. So Corollary 1.2 implies that $\|\mathbb{1}_{\Lambda_N}\psi_j^{(N)}\|^2 \approx \alpha$ for most $\psi_j^{(N)}$. This shows that most $\psi_j^{(N)}$ are uniformly distributed, in the sense that if we consider any $\Lambda_N \subset V_N$ containing half the vertices, we find half the mass of $\|\psi_j^{(N)}\|^2$. As we show in the next subsection, such an iterpretation is also valid for the Anderson model.

For general models, we cannot assert that $\langle \mathbb{1}_{\Lambda_N}\rangle_{\lambda+i\eta_0} = \alpha$. Still, we prove in Section A.3 that there exists $c_\alpha > 0$ such that for any $\Lambda_N \subset V_N$ with $|\Lambda_N| \geq \alpha N$, we have

$$\inf_{\eta_0\in(0,1)} \liminf_{N\longrightarrow\infty} \inf_{\lambda\in I_1} \langle \mathbb{1}_{\Lambda_N}\rangle_{\lambda+i\eta_0} \geq 2c_\alpha. \tag{1.7}$$

Combined with Corollary 1.2, this implies

COROLLARY 1.4. *For any* $\alpha \in (0,1)$, *there exists* $c_\alpha > 0$ *such that for any* $\Lambda_N \subset V_N$ *with* $|\Lambda_N| \geq \alpha N$, *we have*

$$\frac{1}{N}\#\left\{\lambda_j^{(N)} \in I : \left\|\mathbb{1}_{\Lambda_N}\psi_j^{(N)}\right\|^2 < c_\alpha\right\} \underset{N\longrightarrow+\infty}{\longrightarrow} 0.$$

Hence, while in the simple case we had $\|\mathbb{1}_{\Lambda_N}\psi_j^{(N)}\|^2 \approx \alpha$ for most $\psi_j^{(N)}$, in the general case, we can still assert that $\|\mathbb{1}_{\Lambda_N}\psi_j^{(N)}\|^2 \geq c_\alpha > 0$ for most $\psi_j^{(N)}$. This indicates that our theorem can truly be interpreted as a delocalization theorem. The bad indices j (for which $\|\mathbb{1}_{\Lambda_N}\psi_j^{(N)}\|^2 < c_\alpha$) will a priori depend on Λ_N.

We now turn to the general averages $\langle \mathbf{K}\rangle_{\gamma_j}$. Recall that $\Phi_\gamma^N(\tilde{x},\tilde{y}) = \frac{\operatorname{Im}\tilde{g}_N^{\gamma}(\tilde{x},\tilde{y})}{\sum_{\tilde{x}\in\mathcal{D}_N}\operatorname{Im}\tilde{g}_N^{\gamma}(\tilde{x},\tilde{x})}$. We will show in Section A.3 that under assumption (BSCT), we have uniformly in $\lambda \in I_0$,

$$\frac{1}{N}\sum_{x\in V_N}\operatorname{Im}\tilde{g}_N^{\lambda+i\eta_0}(x,x) \underset{N\longrightarrow+\infty}{\longrightarrow} \mathbb{E}\left(\operatorname{Im}\mathcal{G}^{\lambda+i\eta_0}(o,o)\right). \tag{1.8}$$

This already shows that $\Phi_\gamma^N(\tilde{x},\tilde{y})$ is of order $1/N$, since the denominator in its expression is of order N. We strengthen this observation by proving that for any continuous $F:\mathbb{R}\to\mathbb{R}$, we have uniformly in $\lambda \in I_0$,

$$\begin{aligned}&\frac{1}{N}\sum_{x\in V_N}\sum_{y,d(y,x)=k} F\left(N\Phi^N_{\lambda+i\eta_0}(\tilde{x},\tilde{y})\right)\\ &\underset{N\longrightarrow+\infty}{\longrightarrow} \mathbb{E}\left(\sum_{v,d(v,o)=k} F\left(\frac{\operatorname{Im}\mathcal{G}^{\lambda+i\eta_0}(o,v)}{\mathbb{E}\left(\operatorname{Im}\mathcal{G}^{\lambda+i\eta_0}(o,o)\right)}\right)\right).\end{aligned} \tag{1.9}$$

This says that the empirical distribution of $\left(N\Phi_\gamma^N(\tilde{x},\tilde{y})\right)$ (when x is chosen uniformly at random in V_N and y is then chosen uniformly among the points at distance k from x) converges to the law of $\left(\frac{\operatorname{Im}\mathcal{G}^\gamma(o,v)}{\mathbb{E}(\operatorname{Im}\mathcal{G}^\gamma(o,o))}\right)$ (v being chosen uniformly among the points at distance k from the root o). This is a second way of saying that $\Phi_\gamma^N(\tilde{x},\tilde{y})$ is of order $1/N$: when multiplied by N, it has a non-trivial limiting distribution.

1.4. *Case of the Anderson model.* It is important to check that the models covered by the assumptions of our main theorems are not reduced to the case of the Laplacian on regular graphs, already treated in [7], [21], [5]. Here we consider the important case of the Anderson model on regular graphs, i.e., the Laplacian with a random potential. We will show that if the strength of the disorder is small enough, then the assumptions of Theorems 1.1 and 1.3 are satisfied for almost every realization of the potential.

Let $\mathbb{T}_q$ be the $(q+1)$-regular tree. Let ν be a probability measure on $\mathbb{R}$, supported on a compact interval $[-A,A]$, and for every $\epsilon>0$ let ν_ϵ be the image of ν under the homothety $x\mapsto \epsilon x$. (Now ν_ϵ is supported on $[-\epsilon A,\epsilon A]$.) Let $\Omega=\mathbb{R}^{\mathbb{T}_q}$, and define $\mathbf{P}_\epsilon$ on Ω by $\mathbf{P}_\epsilon=\otimes_{v\in\mathbb{T}_q}\nu_\epsilon$. We shall denote by $\mathbf{E}_\epsilon$ the expectation with respect to $\mathbf{P}_\epsilon$. Given $\omega=(\omega_v)\in\Omega$, define $\mathcal{W}^\omega(v)=\omega_v$ for $v\in\mathbb{T}_q$. Then the $\{\omega_v\}_{v\in\mathbb{T}_q}$ are independent and identically distributed random variables with common distribution ν_ϵ. Here $\epsilon\in\mathbb{R}$ is fixed and parametrizes the strength of the disorder.

Let $G_N=(V_N,E_N)$ be a (deterministic) sequence of $(q+1)$-regular graphs with $|V_N|=N$. This means that $\widetilde{G}_N=\mathbb{T}_q$ for all N. Let $\Omega_N=\mathbb{R}^{V_N}$ and $\mathcal{P}_N^\epsilon=\otimes_{x\in V_N}\nu_\epsilon$ on Ω_N. We denote $\widetilde{\Omega}=\prod_{N\in\mathbb{N}}\Omega_N$ and let $\mathcal{P}_\epsilon$ be any probability measure on $\widetilde{\Omega}$ having $\mathcal{P}_N^\epsilon$ as a marginal on the factor Ω_N. Given $(\omega_N)_{N\in\mathbb{N}}\in\widetilde{\Omega}$, so that $\omega_N=(\omega_x)_{x\in V_N}\in\Omega_N$, we define $W^{\omega_N}(x)=\omega_x$ for $x\in V_N$.

The results of this section are proved in a companion paper [8].

Proposition 1.5. *Suppose* (G_N) *satisfies* (BST). *Then* (BSCT) *holds for* $\mathcal{P}^\epsilon$*-a.e. realization of the potential. More precisely, for* $\mathcal{P}^\epsilon$*-a.e.* $(\omega_N)\in\widetilde{\Omega}$, *the sequence* (G_N,W^{ω_N}) *has a local weak limit* $\mathbb{P}_\epsilon$ *that is concentrated on* $\{[\mathbb{T}_q,o,\mathcal{W}^\omega]:\omega\in\Omega\}$, *where* $o\in\mathbb{T}_q$ *is fixed and arbitrary. The measure* $\mathbb{P}_\epsilon$ *acts by taking the expectation with respect to* $\mathbf{P}_\epsilon$; *that is, if* $D=q+1$, *then*

$$\int_{\mathscr{G}_*^{D,\epsilon A}} f([G,v,W])\,\mathrm{d}\,\mathbb{P}_\epsilon([G,v,W])=\int_\Omega f([\mathbb{T}_q,o,\mathcal{W}^\omega])\,\mathrm{d}\mathbf{P}_\epsilon(\omega)$$
$$=\mathbf{E}_\epsilon[f([\mathbb{T}_q,o,\mathcal{W}^\omega])].$$

We make the following assumption on the random variables:

(POT) The measure ν is Hölder continuous; i.e., there exist $C_\nu>0$ and $b\in(0,1]$ such that $\nu(I)\le C_\nu|I|^b$ for all bounded $I\subset\mathbb{R}$.

The following proposition is by no means trivial; it comes from the results of [32], [3].

PROPOSITION 1.6. *Fix* $0 < \lambda_0 < 2\sqrt{q}$. *There exists* $\epsilon(\lambda_0)$ *such that if* $|\epsilon| < \epsilon(\lambda_0)$, *then assumption* (Green) *holds for the measure* $\mathbb{P}_\epsilon$ *of Proposition* 1.5 *on* $I_1 = (-\lambda_0, \lambda_0)$.

COROLLARY 1.7. *If the graphs* G_N *form an expander family and satisfy* (BST) *and if the disorder* ϵ *is small enough, the conclusions of Theorems* 1.1 *and* 1.3 *hold true for* $\mathcal{P}_\epsilon$*-a.e. realization* $(\omega_N) \in \widetilde{\Omega}$, *with* $I_1 = (-\lambda_0, \lambda_0)$.

This gives a rich enough family of examples where the assumptions of Theorems 1.1 and 1.3 hold true. Thus the conclusions of the theorems hold for any observables a_N, K_N. If, in addition, a_N or K_N are independent of the disorder, some extra averaging takes place, and we may replace $\langle \mathbf{K} \rangle_{\lambda + i\eta_0}$ by a simpler average as follows.

THEOREM 1.8. *Assume that* (POT), (EXP) *and* (BST) *hold. Given* $(\omega_N) \in \widetilde{\Omega}$, *let* $(\psi_i^{\omega_N})_{i=1}^N$ *be an orthonormal basis of eigenfunctions of* $H_N^\omega = \mathcal{A}_N + W^{\omega_N}$ *in* $\ell^2(V_N)$, *with corresponding eigenvalues* $(\lambda_i^{\omega_N})_{i=1}^N$.

Let $K_N : V_N \times V_N \to \mathbb{C}$, $\sup_N \sup_{x,y \in V_N} |K_N(x,y)| \leq 1$, $K_N(x,y) = 0$ *if* $d(x,y) > R$, *and assume* K_N *is independent of* (ω_N). *Fix* $0 < \lambda_0 < 2\sqrt{q}$. *If* $|\epsilon| < \epsilon(\lambda_0)$, *we have for* $\mathcal{P}_\epsilon$*-a.e.* (ω_N),

$$\lim_{\eta_0 \downarrow 0} \lim_{N \to \infty} \frac{1}{N} \sum_{\lambda_i^{\omega_N} \in (-\lambda_0, \lambda_0)} \left| \langle \psi_i^{\omega_N}, K_N \psi_i^{\omega_N} \rangle - \langle K_N \rangle_{\lambda_i^{\omega_N}}^{\eta_0} \right| = 0,$$

where for $\gamma \in \mathbb{C} \setminus \mathbb{R}$,

$$(1.10)\quad \langle K \rangle_\lambda^{\eta_0} = \sum_{x,y \in V_N} K(\tilde{x}, \tilde{y}) \widetilde{\Phi}_\gamma(\tilde{x}, \tilde{y}) \quad \text{and} \quad \widetilde{\Phi}_\gamma(\tilde{x}, \tilde{y}) = \frac{1}{N} \cdot \frac{\mathbf{E}_\epsilon[\operatorname{Im} \mathcal{G}^\gamma(\tilde{x}, \tilde{y})]}{\mathbf{E}_\epsilon[\operatorname{Im} \mathcal{G}^\gamma(o,o)]}.$$

As in the previous theorems, if $R = 0$, the ψ_j are arbitrary, while if $R > 0$, we assume the ψ_j are real-valued.

For the Anderson model, $\mathbf{E}_\epsilon(\operatorname{Im} \mathcal{G}^\gamma(v,w))$ depends only on $d(v,w)$:

$$\mathbf{E}_\epsilon(\operatorname{Im} \mathcal{G}^\gamma(v,w)) = \mathbf{E}_\epsilon(\operatorname{Im} \mathcal{G}^\gamma(o,u)),$$

where u is any vertex of $\mathbb{T}_q$ such that $d(o,u) = d(v,w)$.

In the special case $R = 0$, we have $\langle a_N \rangle_\lambda^{\eta_0} = \frac{1}{N} \sum_{x \in V_N} a(x)$. So choosing $a_N = \mathbb{1}_{\Lambda_N}$, Theorem 1.8 implies the strong form of delocalization given by the uniform distribution of $\psi_j^{(N)}$ on V_N, as explained in Section 1.3.

1.5. *Relation with previous work.* Our main Theorem 1.3 holds for *deterministic* sequences of graphs and potentials. For any sequence (G_N, W_N) satisfying the assumptions of the theorem, the conclusion holds for any observable K; in particular, K may depend on the graphs. As already noted, the result only says something about the delocalization of "most" eigenfunctions, where the "good" eigenfunctions exhibiting delocalization may depend on the choice of the observable K.

In the past years, there has been tremendous interest in spectral statistics and delocalization of eigenfunctions of *random* sequences of graphs and potentials. Many papers consider *random* regular graphs, with degree going slowly to infinity [44], [26], [13], [11] or fixed [30], [12], sometimes adding a random independent and identically distributed potential [30]. In particular, the recent papers [13], [11], [12] show "quantum unique ergodicity" for the *adjacency matrix* of random regular graphs: given an observable $a_N : \{1, \dots, N\} \longrightarrow \mathbb{R}$, for most $(q+1)$-regular graphs on the vertices $\{1, \dots, N\}$, we have that $\sum_{x=1}^{N} a_N(x)|\psi_j^{(N)}(x)|^2$ is close to $\langle a_N \rangle$ for *all* indices j. This is a considerable strengthening of Corollary 1.2 (or of the similar result in [7]), which only says something for *most* indices j. This possibility to prove QUE is, of course, due to the fact that a_N has to be independent of the choice of the graph and that results holds for almost all graphs.

When "ergodicity" of eigenfunctions is tested numerically as in the physics papers [24], [23], it is natural to first pick a realization of the graph and of the potential, and then to test the eigenfunctions one by one to determine if they can be localized in small parts of the graph. It is then natural to allow the test-observables to depend on the graph and the potential (which our Theorem 1.3 does, but not the results of [13], [12]), but *also* on the index j of the eigenfunction, which neither of the rigourous mathematical results achieves. The numerical results of [23] seem to indicate that, as soon as a random disorder is turned on, the eigenfunctions will be localized in small parts of the graph. This is not in contradiction with our results: the region of localization of $\psi_j^{(N)}$ might depend on j, but our result does not allow to test this. Note also that the results of [24], [23] were recently questioned in [43], where the authors argue that N has not been taken large enough to see the delocalization take place.

The paper [10] proves a very important result, saying that if ψ_j is an "almost eigenvector" of the adjacency matrix on a random regular graph G, then for almost all G and all j, the value distribution of $\psi_j(x)$ as x runs over $\{1, \dots, N\}$ is close to a Gaussian $\mathcal{N}(0, \sigma_j^2)$ with $\sigma_j \leq 1$. Proving that $\sigma_j = 1$ is a challenge; it would amount to proving that eigenfunctions cannot be localized in small parts of the graph. Our result does not say this, again because we can only test one observable a at a time. The indices j for which Corollary 1.2 proves delocalization depend on a. If we wanted to have a common set of indices j that do the job for all observables (whose number is exponential in N), we would need to have an exponential rate of convergence in Theorems 1.1 and 1.3. Our proof gives a rate that is at best a negative power of $\log N$.

Finally we would also like to mention the paper [19], where the existence of some absolutely continuous spectrum for percolation graphs on the $(q+1)$-regular tree is proven, if the percolation parameter is close enough to 1. Since the absolutely continuous spectrum is mixed with purely discrete spectrum, one cannot expect a quantum ergodicity result that claims delocalization of most eigenfunctions, but only a "partial delocalization" result for a *positive proportion* of eigenfunctions. These are the contents of [19, Th. 9]. It would be nice to investigate what the methods of our paper would give for that model.

1.6. *Outline of the proof.* We borrowed the name "Quantum Ergodicity" from a result about Laplacian eigenfunctions on Riemannian manifolds [46], [47], [45], [48]. The proof in the setting of Laplacian eigenfunctions on manifolds is made of four steps, of unequal difficulty . These four steps are also present in our proof.

Step 0. Define the quantum variance. The goal is to show that this goes to 0 as $N \to \infty$. A novelty of our proof is that we replace the usual quantum variance (10.1) by a "non-backtracking" one (3.3), where we replace the eigenfunctions ψ_j by eigenfunctions f_j, f_j^* of a non-backtracking random walk (Section 3). These new f_j, f_j^* are thus eigenfunctions of a non-selfadjoint problem. This causes new difficulties, which however will be compensated by the fact that the non-backtracking random walk has simpler trajectories than the "simple" random walk generated by the adjacency matrix $\mathcal{A}$.

Step 1. Show that the quantum variance is controlled by the Hilbert-Schmidt norm of K. Although this is obvious for the original quantum variance, this will be much harder for the "non-backtracking quantum variance" (Section 4). This uses (BSCT) and (Green).

Step 2. Due to the fact that f_j, f_j^* satisfy an eigenfunction problem, the quantum variance is invariant under certain transformations (Section 5).

Step 3. One should see behind these transformations the emergence of a "classical dynamical system." In the setting of Laplacian eigenfunctions on manifolds, this is the geodesic flow. Here, what we get is a family of stationary Markov chains on the set of infinite non-backtracking paths (Section 6, Remark 6.1). This step has been called "classicalization" by U. Smilansky in a private conversation; this is supposed to mean the opposite of "quantization."

Step 4. Iterate the classical dynamical system, and use its ergodicity to show that the quantum variance is small (Section 9). Here, the ergodicity of our Markov chains (more precisely, the fact that the mixing rate is independent on N) comes from the (EXP) condition. Assumption (Green) is also used to control the probability transitions.

There is an additional step that does not exist in the traditional setting:

Step 5. Translate the result for the "non-backtracking quantum variance" (involving f_j, f_j^*) into a result for the original one, involving the ψ_j (Section 10). Assumptions (EXP), (BSCT) and (Green) are used here again to show that the transformation sending ψ_j to f_j, f_j^* is well behaved in the limit $N \longrightarrow +\infty$.

2. **Basic identities**

2.1. "*Quantization procedure*" *on trees and their quotients.* Let $G = G_N$, $G = (V, E)$. Most of the time we will drop the subscript N in the notation. As in Section 1.2, we regard G as a quotient: $G = \Gamma\backslash\widetilde{G}$, and we let $\pi : \widetilde{V} \to V$ denote the projection. Fix a fundamental domain $\mathcal{D} \subset \widetilde{V}$ for the action of Γ on $\widetilde{V}$. Then $|\mathcal{D}| = |V|$.

Each edge $\{x_0, x_1\} \in \widetilde{E}$ gives rise to two oriented edges $e = (x_0, x_1)$ and $\hat{e} = (x_1, x_0)$ in the reverse direction. We let o_e and t_e be the origin and terminus of e, respectively. We then let $\widetilde{B}_1$, or simply $\widetilde{B}$, be the set of all such oriented edges of $\widetilde{G}$. More generally, let $\widetilde{B}_k$ be the set of non-backtracking paths of length k in $\widetilde{G}$. By convention, $\widetilde{B}_0 := \widetilde{V}$. If $\omega = (x_0, \ldots, x_k)$ and $\omega' = (x'_0, \ldots, x'_k) \in \widetilde{B}_k$, then we write $\omega \rightsquigarrow \omega'$ if $x'_0 = x_1, \ldots, x'_{k-1} = x_k$ and $(x_0, \ldots, x_k, x'_k) \in \widetilde{B}_{k+1}$. We also denote $o_\omega = x_0$, $t_\omega = x_k$.

These notions descend to the quotient. We denote by $B_k := \Gamma\backslash\widetilde{B}_k$ the set of non-backtracking paths of length k in G. By convention, $B_0 := V$. For $k = 1$, we let $B = B_1$. The set B_k is in bijection with the subset $\mathcal{D}^{(k)} \subset \widetilde{B}_k$ of elements having their origin in $\mathcal{D}$.

Let $\mathscr{H}_k = \mathbb{C}^{B_k}$ (the complex-valued functions on B_k), $\mathscr{H} = \oplus_{k=0}^{\infty} \mathscr{H}_k$ and $\mathscr{H}_{\le k} := \oplus_{\ell=0}^{k} \mathscr{H}_\ell$. It will be convenient to identify $\mathbb{C}^{B_k}$ with the Γ-invariant elements of $\mathbb{C}^{\widetilde{B}_k}$ or with $\mathbb{C}^{\mathcal{D}^{(k)}}$. For $K \in \mathscr{H}_k$ and $(x_0, \ldots, x_k) \in \widetilde{B}_k$, we will sometimes use the short-hand notation $K(x_0; x_k)$ for $K(x_0, \ldots, x_k)$. This is justified by the fact that on $\widetilde{G}$, the endpoints $(x_0; x_k)$ determine the path $(x_0, \ldots, x_k)$ uniquely. We will also use this short-hand notation on B_k, although in that case one should keep in mind that $K(x_0; x_k)$ actually depends on the full path $(x_0, \ldots, x_k)$.

Any $K \in \mathscr{H}_k$ (regarded as a Γ-invariant element of $\mathbb{C}^{\widetilde{B}_k}$) may be used to define an operator $\widehat{K}$ on the space of finitely supported functions on $\widetilde{V}$, with kernel $\langle \delta_v, \widehat{K}\delta_w \rangle_{\ell^2(\widetilde{V})} = K(v; w)$. It also defines an operator $\widehat{K}_G$ on $\mathbb{C}^V$, with kernel

$$K_G(x, y) = \sum_{\gamma \in \Gamma} K(\tilde{x}; \gamma \cdot \tilde{y}),$$

where $\tilde{x}, \tilde{y} \in \widetilde{V}$ are representatives of $x, y \in V$. The map $K \in \mathscr{H}_k \mapsto K_G$ is a priori not one-to-one. However, if $\rho_G(x) \ge k$, then $K_G(x, \cdot)$ determines $K(\tilde{x}, \cdot)$ uniquely. To see that $K \in \mathscr{H}_k \mapsto K_G$ is surjective, consider $\mathbf{k} : V \times V \longrightarrow \mathbb{C}$ supported at distance k from the diagonal, and let $K(\tilde{x}, \tilde{y}) =$

$\mathbf{k}(\pi(\tilde{x}),\pi(\tilde{y}))\mathbb{1}_{\mathrm{dist}(\tilde{x},\tilde{y})\le k}(\sharp\{\gamma\in\Gamma,\mathrm{dist}(\tilde{x},\gamma\cdot\tilde{y})\le k\})^{-1}$. Then $K_G=\mathbf{k}$, and this coincides with the lift (1.6) except at the few points where $\rho_G(x)\le k$.

Define the non-backtracking adjacency operator $\mathcal{B}:\mathbb{C}^{\widetilde{B}}\to\mathbb{C}^{\widetilde{B}}$ by

$$(\mathcal{B}f)(x_0,x_1)=\sum_{x_2\in\mathcal{N}_{x_1}\setminus\{x_0\}}f(x_1,x_2), \tag{2.1}$$

where $\mathcal{N}_x$ means the set of neighbors of x. Then an element $K\in\mathscr{H}_k$ may also be used to define an operator $\widehat{K}_{\widetilde{B}}$ on $\ell^2(\widetilde{B})$, with kernel

$$\langle\delta_{b_1},\widehat{K}_{\widetilde{B}}\delta_{b_2}\rangle_{\ell^2(\widetilde{B})}=\begin{cases}K(o_{b_1};t_{b_2}) & \text{if } \mathcal{B}^{k-1}(b_1,b_2)\neq 0,\\ 0 & \text{otherwise.}\end{cases}$$

Thus $\langle\delta_{b_1},\widehat{K}_{\widetilde{B}}\delta_{b_2}\rangle_{\ell^2(\widetilde{B})}\neq 0$ only if there is a non-backtracking path of length k in $\widetilde{G}$, starting with the oriented edge b_1 and ending with b_2.

Finally, $K\in\mathscr{H}_k$ also defines an operator $\widehat{K}_B$ on $\mathbb{C}^B$, with matrix $K_B:B\times B\to\mathbb{C}$ given by

$$K_B(b_1,b_2)=\sum_{\gamma\in\Gamma}K(\tilde{b}_1;\gamma\cdot\tilde{b}_2),$$

where $\tilde{b}_1,\tilde{b}_2\in\widetilde{B}$ are lifts of $b_1,b_2\in B$. By linearity, this extends to $K\in\mathscr{H}_{\le k}$.

Note that if $K\in\mathscr{H}_k$, then

$$\langle\psi,K_G\phi\rangle_{\ell^2(V)}=\sum_{(x_0,\ldots,x_k)\in B_k}\overline{\psi(x_0)}K(x_0;x_k)\phi(x_k)$$

for any $\psi,\phi\in\ell^2(V)$. Similarly, if $f,g\in\ell^2(B)$, then we have

$$\langle f,K_Bg\rangle_{\ell^2(B)}=\sum_{(x_0,\ldots,x_k)\in B_k}\overline{f(x_0,x_1)}K(x_0;x_k)g(x_{k-1},x_k), \tag{2.2}$$

$$\|K_Bf\|^2_{\ell^2(B)}=\sum_{(x_0,x_1)\in B}\Big|\sum_{x_{0,1}(x_2;x_k)}K(x_0;x_k)f(x_{k-1},x_k)\Big|^2, \tag{2.3}$$

where $\sum_{x_{0,1}(x_2;x_k)}$ sums over all $(x_2;x_k)\in B_{k-2}$ such that $x_2\in\mathcal{N}_{x_1}\setminus\{x_0\}$. Alternatively, we may simply sum over $(x_2;x_k)\in B_{k-2}$ but decide that $K(x_0;x_k)=0$ if $(x_0,\ldots,x_k)\notin B_k$.

Remark 2.1. The maps $K\mapsto\widehat{K}$, $K\mapsto\widehat{K}_G$, $K\mapsto\widehat{K}_{\widetilde{B}}$ and $K\mapsto\widehat{K}_B$ associate an operator to a function on the set of paths. It is tempting to view this as a form of "quantization procedure" as those used for quantum ergodicity on manifolds.

2.2. *Green functions on trees.* Assumption (BST) says that our graphs have few short loops, in other words, that most balls of a given radius look like *trees*. One of the ingredients of our proof is that the Green function on trees satisfies certain algebraic relations, which follow from the fact that removing a vertex (or cutting an edge) from a tree suffices to disconnect it.

Here we recall some standard facts that hold for an arbitrary *tree* $T = (V(T), E(T))$, endowed with a discrete Schrödinger of the form $H = \mathcal{A} + W$ acting on $\ell^2(V(T))$, where $\mathcal{A}$ is the adjacency matrix and $W : V(T) \longrightarrow \mathbb{R}$ is a bounded function. Given $\gamma \in \mathbb{C} \setminus \mathbb{R}$ and $v, w \in T$, the Green function is denoted in this section by

$$G(v, w; \gamma) = \langle \delta_v, (H - \gamma)^{-1} \delta_w \rangle_{\ell^2(V(T))}.$$

If $v \sim w$, we denote by $T^{(v|w)}$ the tree obtained by removing from T the branch emanating from v that passes through w. We define the restriction $H^{(v|w)}(u, u') = H(u, u')$ if $u, u' \in T^{(v|w)}$, and zero otherwise. The corresponding Green function is denoted by $\tilde{g}^{(v|w)}(\cdot, \cdot; \gamma)$. We finally denote

$$\frac{-1}{2m_v^\gamma} = G(v, v; \gamma) \quad \text{and} \quad \zeta_w^\gamma(v) = -\tilde{g}^{(v|w)}(v, v; \gamma).$$

Later on, we will apply these results for $(T, W) = (\widetilde{G}_N, \widetilde{W}_N)$. In this case the (full) Green function will be denoted by $\tilde{g}_N^\gamma(x, y)$, and the restricted one by $\zeta_x^\gamma(y)$. In the case $(T, W) = (\mathcal{T}, \mathcal{W})$ (the random colored rooted trees of assumption (BSCT)), the Green function will be denoted by $\mathcal{G}^\gamma(v, w)$, and the restricted one by $\hat{\zeta}_w^\gamma(v)$. As a general rule, the objects defined on the limit $(\mathcal{T}, \mathcal{W})$ will wear a hat $\hat{\ }$ to distinguish them from similar objects defined on $(\widetilde{G}_N, \widetilde{W}_N)$; see also Remark A.3.

The Green functions on trees satisfy some classical recursive relations; the following lemma is proved, for instance, in [9]. Given $v \in V(T)$, we denote by $\mathcal{N}_v$ its set of nearest neighbors.

Lemma 2.2. *For any $v \in T$ and $\gamma \in \mathbb{C} \setminus \mathbb{R}$, we have*

$$\gamma = W(v) + \sum_{u \sim v} \zeta_v^\gamma(u) + 2m_v^\gamma, \tag{2.4a}$$

$$\gamma = W(v) + \sum_{u \in \mathcal{N}_v \setminus \{w\}} \zeta_v^\gamma(u) + \frac{1}{\zeta_w^\gamma(v)}. \tag{2.4b}$$

For any non-backtracking path $(v_0; v_k)$ in T,

$$G(v_0, v_k; \gamma) = \frac{-\prod_{j=0}^{k-1} \zeta_{v_{j+1}}^\gamma(v_j)}{2m_{v_k}^\gamma}, \tag{2.5}$$

$$G(v_0, v_k; \gamma) = \zeta_{v_1}^\gamma(v_0) G(v_1, v_k; \gamma) = \zeta_{v_{k-1}}^\gamma(v_k) G(v_0, v_{k-1}; \gamma). \tag{2.6}$$

Also, for any $w \sim v$, we have

$$\zeta_w^\gamma(v) = \frac{m_w^\gamma}{m_v^\gamma}\,\zeta_v^\gamma(w) \quad \textit{and} \quad \frac{1}{\zeta_w^\gamma(v)} - \zeta_v^\gamma(w) = 2m_v^\gamma. \tag{2.7}$$

For any $v, w \in T$, we have

$$G(v,w;\gamma) = G(w,v;\gamma). \tag{2.8}$$

Next, if $\gamma = \lambda \pm i\eta$ with $\lambda \in \mathbb{R}$, $\eta > 0$, then

$$\sum_{u \in \mathcal{N}_v \setminus \{w\}} |\operatorname{Im}\zeta_v^\gamma(u)| = \frac{|\operatorname{Im}\zeta_w^\gamma(v)|}{|\zeta_w^\gamma(v)|^2} - \eta. \tag{2.9}$$

Finally, if $\Psi_{\gamma,v}(w) = \operatorname{Im} G(v,w;\gamma)$, then for any path $(v_0,\dots,v_k)$ in T, $k \geq 1$,

$$\Psi_{\gamma,v_0}(v_k) - \zeta_{v_{k-1}}^\gamma(v_k)\Psi_{\gamma,v_0}(v_{k-1}) = \operatorname{Im}\zeta_{v_{k-1}}^\gamma(v_k)\cdot\overline{G(v_0,v_{k-1};\gamma)}. \tag{2.10}$$

Note that $|\zeta_v^{\lambda+i\eta}(u)| \leq \eta^{-1}$. It follows from (2.4b) that for any $\lambda \in [-(A+D), A+D]$ and $\eta \in (0,1)$,

$$\left|\frac{1}{\zeta_w^{\lambda+i\eta}(v)}\right| \leq c_{D,A}\eta^{-1}, \tag{2.11}$$

where $c_{D,A} = 2(A+D)+1$.

Corollary 2.3. *Given $\gamma \in \mathbb{C}\setminus\mathbb{R}$, for any $v_0, v_1 \in T$, $v_0 \sim v_1$, we have*

$$\begin{aligned}\Psi_{\gamma,v_1}(v_1) - \overline{\zeta_{v_0}^\gamma(v_1)}\Psi_{\gamma,v_1}(v_0) - \zeta_{v_0}^\gamma(v_1)\Psi_{\gamma,v_0}(v_1) + |\zeta_{v_0}^\gamma(v_1)|^2\Psi_{\gamma,v_0}(v_0)& \\ = |\operatorname{Im}\zeta_{v_0}^\gamma(v_1)|.&\end{aligned} \tag{2.12}$$

Also, for any non-backtracking path $(v_0;v_k)$ in T, $k \geq 1$, we have

$$\begin{aligned}\Psi_{\gamma,v_0}(v_k) - \overline{\zeta_{v_1}^\gamma(v_0)}\Psi_{\gamma,v_1}(v_k) - \zeta_{v_{k-1}}^\gamma(v_k)\Psi_{\gamma,v_0}(v_{k-1})& \\ +\overline{\zeta_{v_1}^\gamma(v_0)}\zeta_{v_{k-1}}^\gamma(v_k)\Psi_{\gamma,v_1}(v_{k-1}) = 0.&\end{aligned} \tag{2.13}$$

Proof. By (2.10), $\Psi_{\gamma,v_0}(v_1) - \zeta_{v_0}^\gamma(v_1)\Psi_{\gamma,v_0}(v_0) = \operatorname{Im}\zeta_{v_0}^\gamma(v_1)\overline{G(v_0,v_0;\gamma)}$. As $\Psi_{\gamma,v_1}(v_0) = \Psi_{\gamma,v_0}(v_1)$, using (2.6) we thus get

$$\overline{\zeta_{v_0}^\gamma(v_1)}\Psi_{\gamma,v_1}(v_0) - |\zeta_{v_0}^\gamma(v_1)|^2\Psi_{\gamma,v_0}(v_0) = \operatorname{Im}\zeta_{v_0}^\gamma(v_1)\cdot\overline{G(v_0,v_1;\gamma)}. \tag{2.14}$$

Next, since $G(v_1,v_1;\gamma) = \frac{G(v_0,v_1;\gamma)}{\zeta_{v_1}^\gamma(v_0)}$ and $\frac{1}{\zeta_{v_1}^\gamma(v_0)} = \zeta_{v_0}^\gamma(v_1) + 2m_{v_0}^\gamma$, we have

$$\begin{aligned}G(v_1,v_1;\gamma) &= \zeta_{v_0}^\gamma(v_1)G(v_0,v_1;\gamma) + 2m_{v_0}^\gamma G(v_0,v_1;\gamma) \\ &= \zeta_{v_0}^\gamma(v_1)G(v_0,v_1;\gamma) - \zeta_{v_0}^\gamma(v_1),\end{aligned} \tag{2.15}$$

so

$$\Psi_{\gamma,v_1}(v_1) = \operatorname{Im}\zeta_{v_0}^\gamma(v_1)[\operatorname{Re}G(v_0,v_1;\gamma) - 1] + \operatorname{Re}\zeta_{v_0}^\gamma(v_1)\Psi_{\gamma,v_0}(v_1),$$

and thus

$$\begin{aligned}\Psi_{\gamma,v_1}(v_1) &- \zeta_{v_0}^{\gamma}(v_1)\Psi_{\gamma,v_0}(v_1) \\ &= \operatorname{Im}\zeta_{v_0}^{\gamma}(v_1)[\operatorname{Re} G(v_0,v_1;\gamma)-1] - i\operatorname{Im}\zeta_{v_0}^{\gamma}(v_1)\Psi_{\gamma,v_0}(v_1) \\ &= \operatorname{Im}\zeta_{v_0}^{\gamma}(v_1)\overline{G(v_0,v_1;\gamma)} - \operatorname{Im}\zeta_{v_0}^{\gamma}(v_1).\end{aligned} \tag{2.16}$$

This completes the proof of the first claim, by (2.14). Next, we use again that $\Psi_{\gamma,v_0}(v_1) - \zeta_{v_0}^{\gamma}(v_1)\Psi_{\gamma,v_0}(v_0) = \operatorname{Im}\zeta_{v_0}^{\gamma}(v_1)\overline{G(v_0,v_0;\gamma)}$. In addition, by (2.16),

$$\begin{aligned}\overline{\zeta_{v_1}^{\gamma}(v_0)}[\Psi_{\gamma,v_1}(v_1) - \zeta_{v_0}^{\gamma}(v_1)\Psi_{v_1}(v_0)] &= \operatorname{Im}\zeta_{v_0}^{\gamma}(v_1)[\overline{\zeta_{v_1}^{\gamma}(v_0)G(v_0,v_1;\gamma)} - \overline{\zeta_{v_1}^{\gamma}(v_0)}] \\ &= \operatorname{Im}\zeta_{v_0}^{\gamma}(v_1)\overline{G(v_0,v_0;\gamma)},\end{aligned}$$

where the last equality is proved as in (2.15). This proves the second claim for $k=1$.

Now let $k \ge 2$. If we apply (2.10) with v_1 instead of v_0 and use (2.6), we get

$$\overline{\zeta_{v_1}^{\gamma}(v_0)}\Psi_{\gamma,v_1}(v_k) - \overline{\zeta_{v_1}^{\gamma}(v_0)}\zeta_{v_{k-1}}^{\gamma}(v_k)\Psi_{\gamma,v_1}(v_{k-1}) = \operatorname{Im}\zeta_{v_{k-1}}^{\gamma}(v_k)\cdot\overline{G(v_0,v_{k-1};\gamma)}.$$

The second claim for $k \ge 2$ now follows by (2.10). □

We conclude by recalling the fact that for Lebesgue-a.e. $\lambda \in \mathbb{R}$, the Green function has a finite limit on the real axis almost surely. Remember that $\mathscr{T}_*^{D,A}$ is the set of colored rooted trees and that $\mathbb{P}$ is the probability measure on $\mathscr{T}_*^{D,A}$ appearing in (BSCT).

Proposition 2.4. *There exists a Lebesgue-null set $\mathfrak{A} \subset \mathbb{R}$ such that, to each $\lambda \in \mathfrak{S} := \mathbb{R}\setminus\mathfrak{A}$, there is $\Omega_\lambda \subseteq \mathscr{T}_*^{D,A}$ with $\mathbb{P}(\Omega_\lambda) = 1$, such that if $[\mathcal{T},o,\mathcal{W}] \in \Omega_\lambda$, then the limit $G(v,w;\lambda+i0) := \lim_{\eta\downarrow 0} G(v,w;\lambda+i\eta)$ exists for any $v,w \in \mathcal{T}$.*

Proof. Fix $[\mathcal{T},o,\mathcal{W}]$. By [9, Lemma 3.3], there is a Lebesgue-null set $\mathfrak{A}_{[\mathcal{T},o,\mathcal{W}]} \subset \mathbb{R}$ such that for any $\lambda \in \mathfrak{S}_{[\mathcal{T},o,\mathcal{W}]} := \mathbb{R}\setminus\mathfrak{A}_{[\mathcal{T},o,\mathcal{W}]}$, $G(v,w;\lambda+i0)$ exists for all $v,w \in \mathcal{T}$. Let $\mathfrak{D} = \{([\mathcal{T},o,\mathcal{W}],\lambda) :$ the limit does not exist$\}$. Then

$$(\mathbb{P}\otimes\mathrm{Leb})(\mathfrak{D}) = \int_{\mathscr{T}_*^{D,A}} \mathrm{Leb}(\mathfrak{D}_{[\mathcal{T},o,\mathcal{W}]})\,\mathrm{d}\,\mathbb{P}([\mathcal{T},o,\mathcal{W}]),$$

where $\mathfrak{D}_{[\mathcal{T},o,\mathcal{W}]} = \{\lambda \in \mathbb{R} : ([\mathcal{T},o,\mathcal{W}],\lambda) \in \mathfrak{D}\}$. Since $\mathfrak{D}_{[\mathcal{T},o,\mathcal{W}]} \subseteq \mathfrak{A}_{[\mathcal{T},o,\mathcal{W}]}$, we have $\mathrm{Leb}(\mathfrak{D}_{[\mathcal{T},o,\mathcal{W}]}) = 0$ for all $[\mathcal{T},o,\mathcal{W}]$. Hence,

$$0 = (\mathbb{P}\otimes\mathrm{Leb})(\mathfrak{D}) = \int_{\mathbb{R}} \mathbb{P}(\mathfrak{D}_\lambda)\,\mathrm{d}\lambda,$$

where $\mathfrak{D}_\lambda = \{[\mathcal{T},o,\mathcal{W}] \in \mathscr{T}_*^{D,A} : ([\mathcal{T},o,\mathcal{W}],\lambda) \in \mathfrak{D}\}$. It follows that $\mathbb{P}(\mathfrak{D}_\lambda) = 0$ on a Lebesgue-full set $\mathfrak{S}$. Taking $\Omega_\lambda = \mathfrak{D}_\lambda^c$ completes the proof. □

3. The non-backtracking quantum variance

Our strategy follows the one discovered in [5]. We find a transformation turning the eigenfunctions of $\mathcal{A}+W$ on $G=\Gamma\backslash\widetilde{G}$ into eigenfunctions of a "non-backtracking" random walk. The new operator is not self-adjoint, but this difficulty is superseded by the fact that the trajectories of non-backtracking random walks (on a tree) are much simpler than those of usual random walks.

The notation is the same as in the introduction except that we drop the subscript N. Suppose (ψ_j) is an orthonormal basis of eigenfunctions for $H=\mathcal{A}+W$, say $H\psi_j=\lambda_j\psi_j$.

Fix $\eta_0\in(0,1)$, let $\gamma_j=\lambda_j+i\eta_0$, and let

$$f_j(x_0,x_1)=\zeta_{x_0}^{\gamma_j}(x_1)^{-1}\psi_j(x_1)-\psi_j(x_0),$$

where $\zeta_x^\gamma(y)=-\tilde{g}_N^{(y|x)}(y,y;\gamma)$; see notation in Section 2.2. If $\mathcal{B}$ is the non-backtracking operator (2.1), we have

$$\begin{aligned}(\mathcal{B}\zeta^{\gamma_j}f_j)(x_0,x_1)&\\ &=\sum_{x_2\in\mathcal{N}_{x_1}\backslash\{x_0\}}[\psi_j(x_2)-\zeta_{x_1}^{\gamma_j}(x_2)\psi_j(x_1)]\\ &=[\lambda_j\psi_j(x_1)-W(x_1)\psi_j(x_1)-\psi_j(x_0)]-\psi_j(x_1)\left[\gamma_j-W(x_1)-\frac{1}{\zeta_{x_0}^{\gamma_j}(x_1)}\right]\\ &=f_j(x_0,x_1)-i\eta_0\,\psi_j(x_1),\end{aligned}$$

where we used (2.4b). Hence we get

$$\mathcal{B}(\zeta^{\gamma_j}f_j)=f_j-i\eta_0\,\tau_+\psi_j, \tag{3.1}$$

where $\tau_\pm:\ell^2(V)\to\ell^2(B)$ are defined by

$$(\tau_-\psi)(x_0,x_1)=\psi(x_0)\quad\text{and}\quad(\tau_+\psi)(x_0,x_1)=\psi(x_1).$$

In [5] it was possible to set $\eta_0=0$, and (3.1) said exactly that f_j was an eigenfunction of the weighted non-backtracking operator $\mathcal{B}\zeta^{\gamma_j}$ for the eigenvalue 1. At our level of generality, we do not know if ζ^{λ_j+i0} is well defined on $\widetilde{G}_N$. We have to work with $\eta_0>0$ and let η_0 tend to 0 only at the end of the proof, after N has gone to ∞. Hence, f_j is not exactly an eigenfunction, and our formulas will contain error terms of size η_0 that we will need to estimate precisely, to show that they disappear as $N\to+\infty$, followed by $\eta_0\downarrow 0$.

Similarly, if we put

$$f_j^*(x_0,x_1)=\zeta_{x_1}^{\gamma_j}(x_0)^{-1}\psi_j(x_0)-\psi_j(x_1),$$

we note that $f_j^*=\iota f_j$, where ι is the edge reversal involution, and we get

$$\mathcal{B}^*(\iota\zeta^{\gamma_j}f_j^*)=f_j^*-i\eta_0\,\tau_-\psi_j. \tag{3.2}$$

Let I be an open interval such that $\overline{I} \subset I_1$. For $K \in \mathscr{H}_k$, we define

$$\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K) = \frac{1}{N} \sum_{\lambda_j \in I} \left| \left\langle f_j^*, K_B f_j \right\rangle \right| . \tag{3.3}$$

The dependence of this quantity on η_0 is hidden in the definition of f_j, f_j^*. The scalar product $\langle \cdot, \cdot \rangle$ is on $\ell^2(B)$ endowed with the uniform measure; cf. (2.2).

Remark 3.1. We call (3.3) "quantum variance," in analogy to the quantity bearing this name in quantum chaos. However, there are some significant differences:

- We use the functions f_j and f_j^* instead of the original ψ_j. They are (quasi)-eigenfunctions, respectively of the non-selfadjoint operators $\mathcal{B}\zeta^{\gamma_j}$ and $\mathcal{B}^*\iota\zeta^{\gamma_j}$.
- In general $\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathbb{1}) \neq 1$. Actually, in the special case of regular graphs with $W \equiv 0$, one has $\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathbb{1}) = 0$.
- We did not take the square of $\left| \left\langle f_j^*, K_B f_j \right\rangle \right|$ in the definition. This is purely for technical convenience; the square will appear later when we apply the Cauchy-Schwarz inequality.

We will need to extend (3.3) to operators K that depend on the eigenvalue λ_j in a holomorphic fashion, as spelled out in the following definition. Note that K also depends on N; this tends to be implicit in our notation. We let $\mathbb{C}^+ = \{\gamma \in \mathbb{C}, \mathrm{Im}\,\gamma > 0\}$.

Definition 3.2 (Assumptions (Hol)). We assume that $\gamma \mapsto K^\gamma = K_N^\gamma$ is a map from $\gamma \in \mathbb{C}^+$ to $\mathscr{H}_k$ such that

- For $\eta_0 > 0$, for each N and $(x_0; x_k)$, the function $\lambda \mapsto K^{\lambda + i\eta_0}(x_0; x_k)$ from $\mathbb{R} \to \mathbb{C}$ has an analytic extension K_{η_0} to the strip $\{z : |\mathrm{Im}\, z| < \eta_0/2\}$.
- Given $\eta_0 > 0$, we have
$$\sup_N \sup_{\mathrm{Re}\, z \in I_1, |\mathrm{Im}\, z| < \eta_0/2} \sup_{(x_0; x_k)} |K^z_{N,\eta_0}(x_0; x_k)| < +\infty$$
and
$$\sup_N \sup_{\mathrm{Re}\, z \in I_1, |\mathrm{Im}\, z| < \eta_0/2} \sup_{(x_0; x_k)} |\partial_z K^z_{N,\eta_0}(x_0; x_k)| < +\infty.$$
We write $|||K|||_{\eta_0}$ for the maximum of these two quantities.
- For all $s > 0$,

$$\sup_{\eta_1 \in (0,1)} \limsup_{N \to +\infty} \sup_{\lambda \in I_1} \frac{1}{N} \sum_{(x_0; x_k) \in B_k} |K_N^{\lambda + i\eta_1}(x_0; x_k)|^s < +\infty. \tag{3.4}$$

If $\gamma \mapsto K^\gamma$ is holomorphic on $\mathbb{C}^+$, then it obviously satisfies the first point of the definition with $K_{\eta_0}(z) = K^{z + i\eta_0}$. For instance, if $K^\gamma(x_0; x_k)$ has the form $\sum_{n \geq 0} a^{(n)}_{(x_0; x_k)} \gamma^n$, then we see that $\lambda \mapsto K^{\lambda + i\eta_0}(x_0; x_k)$ extends to $K_{\eta_0}(z) = \sum_{n\geq 0} a^{(n)}_{(x_0;x_k)}(z+i\eta_0)^n$. Note that although $\gamma \mapsto \overline{K^\gamma}$ is not holomorphic, its restriction to an horizontal line is still a real-analytic map $\mathbb{R} \ni \lambda \mapsto \overline{K^{\lambda+i\eta_0}(x_0;x_k)}$, as it possesses an analytic extension given by $z \mapsto \sum_{n\geq 0} \overline{a^{(n)}_{(x_0;x_k)}}(z-i\eta_0)^n$. So $\overline{K^\gamma}$ will satisfy (Hol) if K^γ does.

Conditions (Hol) are stable under the sum and composition of operators.

We extend (3.3) to this setting, by letting

$$\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K^\gamma) = \frac{1}{N}\sum_{\lambda_j\in I}\left|\left\langle f_j^*, K_B^{\lambda_j+i\eta_0} f_j\right\rangle\right|. \tag{3.5}$$

Most of the paper is devoted to showing

THEOREM 3.3. *Under the assumptions* (EXP), (BSCT), *and* (Green), *if $K^\gamma \in \mathscr{H}_k$ has the form $K^\gamma = \mathcal{F}_\gamma K$ for the operators $\mathcal{F}_\gamma$ in Corollary* 10.3, *then*

$$\lim_{\eta_0\downarrow 0}\lim_{N\to+\infty} \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K^\gamma) = 0.$$

These $\gamma \mapsto \mathcal{F}_\gamma K$ satisfy (Hol). The fact that this implies Theorem 1.3 is proven in Section 10, which may be read independently of the proof of Theorem 3.3.

4. Step 1: Bound on the non-backtracking quantum variance

Given $\gamma \in \mathbb{C}^+$, we introduce a norm on each $\mathscr{H}_k$, $k \geq 1$, defined by

$$\|K\|_\gamma^2 = \frac{1}{N}\sum_{(x_0;x_k)\in B_k} \frac{|\operatorname{Im}\zeta^\gamma_{x_1}(x_0)|}{|\zeta^\gamma_{x_1}(x_0)|^2}\cdot |K(x_0;x_k)|^2 \cdot \frac{|\operatorname{Im}\zeta^\gamma_{x_{k-1}}(x_k)|}{|\zeta^\gamma_{x_{k-1}}(x_k)|^2}. \tag{4.1}$$

We denote the associated scalar product by $\langle\cdot,\cdot\rangle_\gamma$. The reason for introducing the weight $\frac{|\operatorname{Im}\zeta^\gamma_x(y)|}{|\zeta^\gamma_x(y)|^2}$ will be apparent in Section 6. The aim of this section is to prove Theorem 4.1. Here, we assume that $I = (a,b)$, with $[a,b] \subset I_1$. This implies that there is $\eta_{a,b}$ such that $(a-2\eta, b+2\eta) \subset I_1$ for all $\eta \leq \eta_{a,b}$. We then assume that $\eta \leq \min(\eta_0/2, \eta_{a,b})$.

THEOREM 4.1. *Under assumptions* (BSCT) *and* (Green), *if $K^\gamma \in \mathscr{H}_k$ satisfies the set of assumptions* (Hol), *then for any interval $I = (a,b)$ as above,*

$$\lim_{\eta_0\downarrow 0}\limsup_{N\to+\infty} \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K^\gamma)^2 \leq D\,|I| \lim_{\eta_0\downarrow 0}\lim_{\eta\downarrow 0}\limsup_{N\to\infty}\int_{a-2\eta}^{b+2\eta} \|K^{\lambda+i(\eta^4+\eta_0)}\|^2_{\lambda+i(\eta^4+\eta_0)}\,\mathrm{d}\lambda.$$

In the scheme of Section 1.6, this corresponds to Step 1. This is more complicated than usual, due to the fact that we have replaced the orthonormal family (ψ_j) by non-orthogonal functions $(f_j), (f_j^*)$, and also because K "depends on λ_j" in (3.5).

Recall that D above is the maximal degree and we assumed $|W_N(x)| \leq A$. In particular, any eigenvalue λ_j lies in $I_0 := [-(A+D), A+D]$. For $\lambda \in \mathbb{R}$ and $\eta_0 \in (0,1)$, let

$$\alpha_{\lambda+i\eta_0}(x_0, x_1) = \frac{|\operatorname{Im} \zeta_{x_1}^{\lambda+i\eta_0}(x_0)|^{1/2}}{\zeta_{x_1}^{\lambda+i\eta_0}(x_0)}.$$

Denoting $\gamma_j = \lambda_j + i\eta_0$, we have (by a double application of the Cauchy-Schwarz inequality)

$$\begin{aligned} \operatorname{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K^\gamma) &\leq \frac{1}{N} \sum_{\lambda_j \in I} \left\| \overline{\alpha_{\gamma_j}}^{-1} f_j^* \right\| \left\| \alpha_{\gamma_j} K_B^{\gamma_j} f_j \right\| \\ &\leq \frac{1}{N} \Big(\sum_{\lambda_j \in I} \left\| \overline{\alpha_{\gamma_j}}^{-1} f_j^* \right\|^2 \Big)^{1/2} \Big(\sum_{\lambda_j \in I} \left\| \alpha_{\gamma_j} K_B^{\gamma_j} f_j \right\|^2 \Big)^{1/2}. \end{aligned} \tag{4.2}$$

We check at the end of the section that

$$\lim_{\eta_0 \downarrow 0} \limsup_{N \to +\infty} \frac{1}{N} \sum_{\lambda_j \in I} \left\| \overline{\alpha_{\gamma_j}}^{-1} f_j^* \right\|^2 \leq D \cdot |I|. \tag{4.3}$$

We now introduce an approximation χ of $\mathbb{1}_I$ by an entire function, by the standard convolution procedure: Fix $0 < \eta \leq \eta_0/2$. Let $\phi(x) = \frac{1}{\pi^{1/2}} e^{-x^2}$, and denote $\phi_\epsilon(x) = \epsilon^{-1}\phi(x/\epsilon)$. Let χ be the convolution $\chi = \phi_{\eta^{3/2}} * \mathbb{1}_I$ on $\mathbb{R}$. Then χ extends to an entire function on $\mathbb{C}$ given by

$$\chi(z) = \frac{1}{\eta^{3/2}\pi^{1/2}} \int_I e^{-(z-y)^2/\eta^3}\, \mathrm{d}y. \tag{4.4}$$

Note that $0 \leq \chi(x) \leq 1$ for $x \in \mathbb{R}$, and $|\chi(z)| \leq e^{\eta^5}$ for $|\operatorname{Im} z| \leq \eta^4$. We assume η is small enough so that $\chi \geq \frac{1}{3}\mathbb{1}_I$ and $|\chi(z)| \leq e^{-1/\eta}$ on

$$\{z \in \mathbb{C} : |\operatorname{Im} z| \leq \eta^4,\ d(\operatorname{Re} z, I) \geq 2\eta\}.$$

We finally note that $|\frac{\partial \chi}{\partial t_2}(t_1 + it_2)| \leq C\eta^{-3}e^{\eta^5}$ for any $z = t_1 + it_2$ with $t_1 \in I_0$ and $|t_2| \leq \eta^4$.

By (4.2) and (4.3), we have

$$\limsup_{N\to\infty} \operatorname{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K^\gamma)^2 \leq \limsup_{N\to\infty} \frac{3D\,|I|}{N} \sum_{j=1}^{N} \chi(\lambda_j)\, \|\alpha_{\gamma_j} K_B^{\gamma_j} f_j\|^2. \tag{4.5}$$

Now by (2.3), we have

$$\begin{aligned} \|\alpha_{\gamma_j} K_B^{\gamma_j} f_j\|^2 = &\sum_{(x_0,x_1)\in B} \sum_{(x_2;x_k),(y_2;y_k)} |\alpha_{\gamma_j}(x_0,x_1)|^2 K^{\gamma_j}(x_0;x_k)\overline{K^{\gamma_j}(x_0;y_k)} \\ &\cdot [\zeta_{x_{k-1}}^{\gamma_j}(x_k)^{-1}\psi_j(x_k) - \psi_j(x_{k-1})]\overline{[\zeta_{y_{k-1}}^{\gamma_j}(y_k)^{-1}\psi_j(y_k) - \psi_j(y_{k-1})]}, \end{aligned}$$

where $(x_0;x_k) = (x_0, x_1, x_2, \ldots, x_k)$, $(x_0;y_k) = (x_0, x_1, y_2, \ldots, y_k)$ and with the convention that $K^{\gamma_j}(x_0;x_k) = 0$ if the path $(x_0, x_1, x_2, \ldots, x_k)$ backtracks.

The function $\lambda \mapsto |\alpha_{\lambda+i\eta_0}(x_0,x_1)|^2 = \frac{-\operatorname{Im}\zeta_{x_1}^{\lambda+i\eta_0}(x_0)}{|\zeta_{x_1}^{\lambda+i\eta_0}(x_0)|^2}$ extends analytically to the rectangle

$$\mathscr{R} = \{z \in \mathbb{C} : \operatorname{Re} z \in [-(A+D+\eta),(A+D+\eta)], \operatorname{Im} z \in [-\eta^4,\eta^4]\}$$

through the formula $\frac{\zeta_{x_1}^{z-i\eta_0}(x_0)-\zeta_{x_1}^{z+i\eta_0}(x_0)}{2i\,\zeta_{x_1}^{z+i\eta_0}(x_0)\zeta_{x_1}^{z-i\eta_0}(x_0)}$. We denote this by $\alpha_{\eta_0}^z(x_0,x_1)$ (which is not the same as $|\alpha_{z+i\eta_0}(x_0,x_1)|^2$). The same is true for the other ζ terms. We denote the extension of $\lambda \mapsto K^{\lambda+i\eta_0}(x_0;x_k)\overline{K^{\lambda+i\eta_0}(x_0;y_k)}$ by $K_{\eta_0}^z(x_0;x_k,y_k)$. Again, if $(x_0;y_k)=(x_0;x_k)$, this is not the same as $|K^{z+i\eta_0}(x_0;x_k)|^2$. However, see Lemma 4.4 to compare both.

Given $x,y \in V$ and $z \in \mathbb{C}\setminus\mathbb{R}$, let

$$g^z(x,y) = \langle \delta_x, (H-z)^{-1}\delta_y\rangle_{\ell^2(V)} = \sum_{j=1}^N \frac{\psi_j(x)\overline{\psi_j(y)}}{\lambda_j - z}$$

be the Green function of H on the finite graph G. Then by Cauchy's integral formula,

$$\begin{aligned}
&\frac{1}{N}\sum_{j=1}^N \chi(\lambda_j)\,\|\alpha_{\gamma_j} K_B^{\gamma_j} f_j\|^2 \\
&\quad = \frac{-1}{2i\pi N}\int_{z\in\partial\mathscr{R}} \sum_{(x_0,x_1)\in B}\ \sum_{(x_2;x_k),(y_2;y_k)} \chi(z)\alpha_{\eta_0}^z(x_0,x_1) \\
&\qquad K_{\eta_0}^z(x_0;x_k,y_k)\cdot\Big[\frac{g^z(x_k,y_k)}{\zeta_{x_{k-1}}^{z+i\eta_0}(x_k)\zeta_{y_{k-1}}^{z-i\eta_0}(y_k)} - \frac{g^z(x_k,y_{k-1})}{\zeta_{x_{k-1}}^{z+i\eta_0}(x_k)} \\
&\qquad\qquad - \frac{g^z(x_{k-1},y_k)}{\zeta_{y_{k-1}}^{z-i\eta_0}(y_k)} + g^z(x_{k-1},y_{k-1})\Big]\,\mathrm{d}z.
\end{aligned} \tag{4.6}$$

We now observe that the integral over the vertical segments of the contour do not contribute as $\eta,\eta_0 \downarrow 0$. More precisely,

LEMMA 4.2. *The integral* $\frac{-1}{2i\pi N}\int_{z\in\partial\mathscr{R}} F(z)\,\mathrm{d}z$ *in* (4.6) *may be replaced by* $\frac{1}{2i\pi N}(\int_{a-2\eta}^{b+2\eta} F(\lambda+i\eta^4)\,\mathrm{d}\lambda - \int_{a-2\eta}^{b+2\eta} F(\lambda-i\eta^4)\,\mathrm{d}\lambda)$, *up to an error term at most* $C_{k,D,A}\eta_0^{-3}\eta^{-4}|||K|||_{\eta_0}^2 e^{-1/\eta}$.

Proof. The error is the integral of $F(z)$ on the two vertical paths $\{\operatorname{Re} z = -A-D-\eta, \operatorname{Im} z \in [-\eta^4,\eta^4]\}$, $\{\operatorname{Re} z = A+D+\eta, \operatorname{Im} z\in[-\eta^4,\eta^4]\}$, and the four connected components of the set

$$\{\operatorname{Im} z = \pm\eta^4, \operatorname{Re} z \in [-A-D-\eta, A+D+\eta]\setminus(a-2\eta,b+2\eta)\}.$$

On these pieces, we know that $|\chi(z)| \le e^{-1/\eta}$. Moreover, $|K_{\eta_0}^z(x_0;x_k,y_k)| \le |||K|||_{\eta_0}^2$. Next, by (2.11),

$$|\alpha_{\eta_0}^z| = \frac{1}{2}\left|\frac{1}{\zeta_{x_1}^{z+i\eta_0}(x_0)} - \frac{1}{\zeta_{x_1}^{z-i\eta_0}(x_0)}\right| \le c_{D,A}\left(\frac{1}{\eta_0+\eta^4} + \frac{1}{\eta_0-\eta^4}\right).$$

Since $\eta \le \eta_0/2$ by assumption, this yields $|\alpha^z_{\eta_0}| \le C_{D,A}\eta_0^{-1}$. The Green functions and ζ terms may be bounded similarly by $4c_{D,A}\eta_0^{-2}\eta^{-4}$. A factor $C_{k,D}$ comes from the number of paths, divided by N. □

Our next aim is to lift this expression to the universal cover $\widetilde{G}$. In other words, we wish to replace g^z by $\tilde{g}^z$ everywhere, to be able to use the identities of Section 2.2.

LEMMA 4.3. *Denote* $z = \lambda + i\eta^4$. *Given* $R \in \mathbb{N}^*$, *there is* $d_{R,k,\eta} > 0$ *such that the integral* $\frac{1}{2i\pi N}\int_{a-2\eta}^{b+2\eta} F(z)\,\mathrm{d}\lambda$ *in Lemma 4.2 may be replaced by*

$$\frac{1}{2i\pi N}\int_{a-2\eta}^{b+2\eta} \sum_{\rho_G(x_0)\ge d_{R,k,\eta}} \sum_{x_1\sim x_0} \sum_{(x_2;x_k),(y_2;y_k)} \chi(z)\alpha^z_{\eta_0}(x_0,x_1)$$
$$K^z_{\eta_0}(x_0;x_k,y_k)\cdot\Big[\frac{\tilde g^z(\tilde x_k,\tilde y_k)}{\zeta^{z+i\eta_0}_{e_k}\zeta^{z-i\eta_0}_{e'_k}} - \frac{\tilde g^z(\tilde x_k,\tilde y_{k-1})}{\zeta^{z+i\eta_0}_{e_k}}$$
$$- \frac{\tilde g^z(\tilde x_{k-1},\tilde y_k)}{\zeta^{z-i\eta_0}_{e'_k}} + \tilde g^z(\tilde x_{k-1},\tilde y_{k-1})\Big]\,\mathrm{d}\lambda,$$

where $\zeta^\gamma_{e_k} = \zeta^\gamma_{x_{k-1}}(x_k)$ *and* $\zeta^\gamma_{e'_k} = \zeta^\gamma_{y_{k-1}}(y_k)$, *up to an error term*

$$\left(\frac{\#\{\rho_G(x_0) < d_{R,k,\eta}\}}{N}\eta^{-4} + \frac{1}{R}\right) C_{k,D,A}\eta_0^{-3}|||K|||^2_{\eta_0}e^{\eta^5}.$$

Similarly, $\frac{1}{2i\pi N}\int_{a-2\eta}^{b+2\eta} F(\bar z)\,\mathrm{d}\lambda$ *in Lemma 4.2 may be replaced by*

$$\frac{1}{2i\pi N}\int_{a-2\eta}^{b+2\eta} \sum_{\rho_G(x_0)\ge d_{R,k,\eta}} \sum_{x_1\sim x_0} \sum_{(x_2;x_k),(y_2;y_k)} \chi(\bar z)\alpha^{\bar z}_{\eta_0}(x_0,x_1)K^{\bar z}_{\eta_0}(x_0;x_k,y_k)$$
$$\cdot\Big[\frac{\tilde g^{\bar z}(\tilde x_k,\tilde y_k)}{\zeta^{\bar z+i\eta_0}_{e_k}\zeta^{\bar z-i\eta_0}_{e'_k}} - \frac{\tilde g^{\bar z}(\tilde x_k,\tilde y_{k-1})}{\zeta^{\bar z+i\eta_0}_{e_k}}$$
$$- \frac{\tilde g^{\bar z}(\tilde x_{k-1},\tilde y_k)}{\zeta^{\bar z-i\eta_0}_{e'_k}} + \tilde g^{\bar z}(\tilde x_{k-1},\tilde y_{k-1})\Big]\,\mathrm{d}\lambda$$

up to an error term $(\frac{\#\{\rho_G(x_0)<d_{R,k,\eta}\}}{N}\eta^{-4} + \frac{1}{R})C_{k,D,A}\eta_0^{-3}|||K|||^2_{\eta_0}e^{\eta^5}$.

Proof. We first approximate $\lambda \mapsto g^{\lambda+i\eta^4}(x,y)$ by a polynomial on the compact interval I_0. Let $h_\eta(t) = -(t-i\eta^4)^{-1}$, and choose a polynomial q_η with $\|h_\eta - q_\eta\|_\infty < \frac{1}{R}$. Then $\|h_\eta(H-\lambda) - q_\eta(H-\lambda)\| < \frac{1}{R}$, so

$$|g^{\lambda+i\eta}(x,y) - q_\eta(H-\lambda)(x,y)| < \frac{1}{R}$$

for any x, y and λ. So replacing each $g^{\lambda+i\eta^4}(x,y)$ by $q_\eta(H-\lambda)(x,y)$ in the sums gives an error term $\frac{C_{k,D,A}\eta_0^{-3}|||K|||^2_{\eta_0}e^{\eta^5}}{R}$ as in Lemma 4.2.

Denote $C_{k,D,A,\eta_0} = C_{k,D,A}\eta_0^{-3}\|K\|_{\eta_0}^2$. Let $d_{R,\eta}$ be the degree of q_η. Suppose $\rho_G(x_0) \geq d_{R,\eta} + k =: d_{R,k,\eta}$. Then it is easy to see that

$$q_\eta(H-\lambda)(x_k, y_k) = q_\eta(\widetilde{H}-\lambda)(\tilde{x}_k, \tilde{y}_k);$$

cf. Lemma A.1. The same holds for the other edges (x_k, y_{k-1}) and so on. The terms with $\rho_G(x_0) < d_{R,k,\eta}$ bring an error term $\frac{\#\{\rho_G(x_0)<d_{R,k,\eta}\}}{N}\eta^{-4}C_{k,D,A,\eta_0}$. Finally, we replace the $q_\eta(\widetilde{H}-\lambda)(\tilde{x},\tilde{y})$ by $\tilde{g}^{\lambda+i\eta^4}(\tilde{x},\tilde{y})$, which yields again an error of the form $\frac{C_{k,D,A,\eta_0}}{R}$. This proves the first statement, and the second one is proven similarly. □

We continue to simplify the expression and record the following.

LEMMA 4.4. *If we replace*

$$\alpha_{\eta_0}^z(x_0,x_1)K_{\eta_0}^z(x_0;x_k,y_k) \text{ and } \alpha_{\eta_0}^{\bar{z}}(x_0,x_1)K_{\eta_0}^{\bar{z}}(x_0;x_k,y_k)$$

in Lemma 4.3 by

$$|\alpha_{z+i\eta_0}(x_0,x_1)|^2 K^{z+i\eta_0}(x_0;x_k)\overline{K^{z+i\eta_0}(x_0;y_k)},$$

then as $N\to\infty$, the error we get is at most $C_{k,D,A}\eta_0^{-6}|\!|\!|K|\!|\!|_{\eta_0}^2 e^{\eta^5}\eta^4$.

We may also replace $\chi(\lambda \pm i\eta^4)$ by $\chi(\lambda)$, modulo the asymptotic error $C_{k,D,A}\eta_0^{-3}|\!|\!|K|\!|\!|_{\eta_0}^2 e^{\eta^5}\eta$. Finally, we may replace each $\zeta_{e_k}^{\bar{z}+i\eta_0}$ by $\zeta_{e_k}^{z+i\eta_0}$ and $\zeta_{e_k'}^{z-i\eta_0}$ by $\zeta_{e_k'}^{\bar{z}-i\eta_0}$, modulo an asymptotic error $C_{k,D,A}\eta_0^{-6}|\!|\!|K|\!|\!|_{\eta_0}^2 e^{\eta^5}\eta^4$.

Proof. We start with $\alpha_{\eta_0}^z(x_0,x_1)K_{\eta_0}^z(x_0;x_k,y_k)$. Denote $e=(x_0,x_1)$ and $\zeta_e^\gamma = \zeta_{x_1}^\gamma(x_0)$. We note that

$$\begin{aligned}\left|\alpha_{\eta_0}^z(x_0,x_1) - |\alpha^{z+i\eta_0}(x_0,x_1)|^2\right| &= \left|\frac{\zeta_e^{z-i\eta_0}-\zeta_e^{z+i\eta_0}}{2i\zeta_e^{z+i\eta_0}\zeta_e^{z-i\eta_0}} - \frac{\zeta_e^{\bar{z}-i\eta_0}-\zeta_e^{z+i\eta_0}}{2i\zeta_e^{z+i\eta_0}\zeta_e^{\bar{z}-i\eta_0}}\right| \\ &= \frac{1}{2}\left|\frac{1}{\zeta_e^{\bar{z}-i\eta_0}} - \frac{1}{\zeta_e^{z-i\eta_0}}\right| \leq C_{D,A}\eta_0^{-2}\left|\zeta_e^{z-i\eta_0}-\zeta_e^{\bar{z}-i\eta_0}\right| \\ &\leq C_{D,A}\eta_0^{-4}|z-\bar{z}| = 2C_{D,A}\eta_0^{-4}\eta^4,\end{aligned}$$

where we used (2.11) in the first inequality and the resolvent identity in the second one. Similarly, $K^{z+i\eta_0}(x_0;x_k)\overline{K^{z+i\eta_0}(x_0;y_k)}$ is the same as $K_{\eta_0}^z(x_0;x_k,y_k)$, but with each $z-i\eta_0$ replaced by $\bar{z}-i\eta_0$. It follows that

$$\begin{aligned}&|K_{\eta_0}^z(x_0;x_k,y_k) - K^{z+i\eta_0}(x_0;x_k)\overline{K^{z+i\eta_0}(x_0;y_k)}| \\ &\qquad \leq 2\sup|\partial_z K(v_0;v_k)|\sup|K(v_0;v_k)|\cdot|z-\bar{z}| \\ &\qquad \leq 4|\!|\!|K|\!|\!|_{\eta_0}^2\eta^4.\end{aligned}$$

Hence, $\alpha_{\eta_0}^z(x_0,x_1)K_{\eta_0}^z(x_0;x_k,y_k)$ is the same as

$$|\alpha_{z+i\eta_0}(x_0,x_1)|^2 K^{z+i\eta_0}(x_0;x_k)\overline{K^{z+i\eta_0}(x_0;y_k)},$$

modulo $C_{D,A}\eta_0^{-4}\|K\|_{\eta_0}^2\eta^4$. This error is further multiplied by the function χ. Bounding the ζ terms by some $c_{D,A}\eta_0^{-2}$ and $|\chi(z)|$ by e^{η^5}, we end up with an error term at most

$$\int_{a-2\eta}^{b+2\eta} \frac{C_{D,A}\eta_0^{-6}\|K\|_{\eta_0}^2 e^{\eta^5}\eta^4}{N} \sum_{(x_0,x_1)} \sum_{(x_2;x_k),(y_2;y_k)} \left|\tilde{g}^{\lambda\pm i\eta^4}(\tilde{x}_k,\tilde{y}_k)\right| \, \mathrm{d}\lambda$$

and a similar upper bound for each term involving $\tilde{g}^{\lambda\pm i\eta^4}$. Since $I_\eta = (a-2\eta, b+2\eta) \subset I_1$, we may use Remark A.5 to deduce that the integrand is uniformly bounded over $\lambda \in I_\eta$ by $C_{k,D,A}\eta_0^{-6}\|K\|_{\eta_0}^2 e^{\eta^5}\eta^4$ as $N\to\infty$. Note that $|I_\eta| \le |I_0| = 2(D+A)$.

This proves the first claim. The second claim is similar; for example,

$$|\alpha_{\eta_0}^{\bar{z}}(x_0,x_1) - |\alpha^{z+i\eta_0}(x_0,x_1)|^2| \le C_{D,A}\eta_0^{-2}|\zeta_e^{z+i\eta_0} - \zeta_e^{\bar{z}+i\eta_0}| \le 2C_{D,A}\eta_0^{-4}\eta^4.$$

Moreover, $K_{\eta_0}^{\bar{z}}(x_0;x_k,y_k)$ is the same as $K^{z+i\eta_0}(x_0;x_k)\overline{K^{z+i\eta_0}(x_0;y_k)}$ with each $z+i\eta_0$ replaced by $\bar{z}+i\eta_0$, so the proof carries on. For the third claim, note that

$$|\chi(\lambda\pm i\eta^4) - \chi(\lambda)| \le \sup_{z\in\mathscr{R}} \left|\frac{\partial\chi}{\partial x_2}(z)\right| \cdot \eta^4 \le Ce^{\eta^5}\eta.$$

For the last claim, $|(\zeta_e^{z\pm i\eta_0})^{-1} - (\zeta_e^{\bar{z}\pm i\eta_0})^{-1}| \le 2C_{D,A}\eta_0^{-4}\eta^4$ as we previously saw when analyzing $\alpha_{\eta_0}^z$, so we get a similar error. □

By virtue of Lemmas 4.3 and 4.4, denoting $z=\lambda+i\eta^4$, we know at this stage that modulo some error terms, the expression (4.6) may be replaced by

$$\begin{aligned}\frac{1}{\pi N}\int_{a-2\eta}^{b+2\eta} &\sum_{\rho_G(x_0)\ge d_{R,k,\eta}} \sum_{x_1\sim x_0} \sum_{(x_2;x_k),(y_2;y_k)} \chi(\lambda)|\alpha_{z+i\eta_0}(x_0,x_1)|^2 \\ &\cdot K^{z+i\eta_0}(x_0;x_k)\overline{K^{z+i\eta_0}(x_0;y_k)} \\ &\cdot\left(\frac{\operatorname{Im}\tilde{g}^z(\tilde{x}_k,\tilde{y}_k)}{\zeta_{e_k}^{z+i\eta_0}\zeta_{e_k'}^{\bar{z}-i\eta_0}} - \frac{\operatorname{Im}\tilde{g}^z(\tilde{x}_k,\tilde{y}_{k-1})}{\zeta_{e_k}^{z+i\eta_0}}\right. \\ &\left.-\frac{\operatorname{Im}\tilde{g}^z(\tilde{x}_{k-1},\tilde{y}_k)}{\zeta_{e_k'}^{\bar{z}-i\eta_0}} + \operatorname{Im}\tilde{g}^z(\tilde{x}_{k-1},\tilde{y}_{k-1})\right)\mathrm{d}\lambda.\end{aligned} \tag{4.7}$$

We now make the expression more homogeneous as follows:

Lemma 4.5. *Assume we have made all the replacements in Lemma 4.4. If we finally replace each of the four $\operatorname{Im}\tilde{g}^z(\tilde{x},\tilde{y})$ by $\operatorname{Im}\tilde{g}^{z+i\eta_0}(\tilde{x},\tilde{y})$ in (4.7), then the error term vanishes as $N\to\infty$, followed by $\eta\downarrow 0$, followed by $\eta_0\downarrow 0$.*

Proof. We only analyze the first error term; the other three are similar. Choose p, q, r such that $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$, and use the Hölder's inequality

$$\begin{aligned}
&\Bigg|\frac{1}{\pi N}\int_{a-2\eta}^{b+2\eta}\sum_{\rho_G(x_0)\geq d_{R,k,\eta}}\sum_{x_1\sim x_0}\sum_{(x_2;x_k),(y_2;y_k)}\chi(\lambda)K^{z+i\eta_0}(x_0;x_k)\overline{K^{z+i\eta_0}(x_0;y_k)}\\
&\quad\cdot\frac{|\alpha_{z+i\eta_0}(x_0,x_1)|^2}{\zeta_{e_k}^{z+i\eta_0}\zeta_{e'_k}^{\bar z-i\eta_0}}\Big(\operatorname{Im}\tilde g^{z}(\tilde x_k,\tilde y_k)-\operatorname{Im}\tilde g^{z+i\eta_0}(\tilde x_k,\tilde y_k)\Big)\,\mathrm{d}\lambda\Bigg|\\
&\leq\frac{e^{\eta^5}}{\pi N}\left(\int\sum_{(x_0,x_1)\in B}\sum_{(x_2;x_k),(y_2;y_k)}\Big|K^{z+i\eta_0}(x_0;x_k)K^{z+i\eta_0}(x_0;y_k)\Big|^p\,\mathrm{d}\lambda\right)^{1/p}\\
&\quad\times\left(\int\sum_{(x_0,x_1)\in B}\sum_{(x_2;x_k),(y_2;y_k)}\left|\frac{|\alpha_{z+i\eta_0}(x_0,x_1)|^2}{\zeta_{e_k}^{z+i\eta_0}\zeta_{e'_k}^{\bar z-i\eta_0}}\right|^q\,\mathrm{d}\lambda\right)^{1/q}\\
&\quad\times\left(\int\sum_{(x_0,x_1)\in B}\sum_{(x_2;x_k),(y_2;y_k)}\Big|\operatorname{Im}\tilde g^{z}(\tilde x_k,\tilde y_k)-\operatorname{Im}\tilde g^{z+i\eta_0}(\tilde x_k,\tilde y_k)\Big|^r\,\mathrm{d}\lambda\right)^{1/r}.
\end{aligned}$$

Here $\int=\int_{a-2\eta}^{b+2\eta}$. The first sum is bounded by

$$D^{k-1}\sum_{(x_0;x_k)\in B_k}|K^{z+i\eta_0}(x_0;x_k)|^{2p}.$$

Assumption (Hol) on K implies that

$$\sup_{\eta_0,\eta}\limsup_{N\to\infty}\frac{1}{N}\int\sum_{(x_0;x_k)\in B_k}|K^{\lambda+i\eta^4+i\eta_0}(x_0;x_k))|^{2p}\,\mathrm{d}\lambda<+\infty.$$

Next, by Remark A.3,

$$\begin{aligned}
\lim_{N\to\infty}\frac{1}{N}\int\sum_{(x_0,x_1)\in B}\sum_{(x_2;x_k),(y_2;y_k)}&\left|\frac{|\alpha_{z+i\eta_0}(x_0,x_1)|^2}{\zeta_{e_k}^{z+i\eta_0}\zeta_{e'_k}^{\bar z-i\eta_0}}\right|^q\,\mathrm{d}\lambda\\
&=\int\mathbb{E}\left(\sum_{(x_0;x_k),(y_0;y_k),x_0=y_0=o}\left|\frac{|\hat\alpha_{z+i\eta_0}(x_0,x_1)|^2}{\hat\zeta_{e_k}^{z+i\eta_0}\hat\zeta_{e'_k}^{\bar z-i\eta_0}}\right|^q\right)\mathrm{d}\lambda,
\end{aligned}$$

and the right-hand side is uniformly bounded in $\eta,\eta_0\in(0,1)$ by Remark A.4. Remember the convention that objects wearing a hat $\hat{}$ are defined on the limit $(\mathcal{T},\mathcal{W})$, by similar formulas to those on G_N. We also refer to Section 2.2 for notation related to Green functions.

Finally, again by Remark A.3 we have

$$\lim_{N\to\infty}\frac{1}{N}\int\sum_{(x_0,x_1)\in B}\sum_{(x_2;x_k),(y_2;y_k)}\Big|\operatorname{Im}\tilde g^{z}(\tilde x_k,\tilde y_k)-\operatorname{Im}\tilde g^{z+i\eta_0}(\tilde x_k,\tilde y_k)\Big|^r\,\mathrm{d}\lambda$$

$$=\int\mathbb{E}\left(\sum_{(v_0;v_k),(w_0;w_k),v_0=w_0=o}\Big|\operatorname{Im}\mathcal{G}^{z}(v_k,w_k)-\operatorname{Im}\mathcal{G}^{z+i\eta_0}(v_k,w_k)\Big|^r\right)\mathrm{d}\lambda.$$

We check that the right-hand side vanishes as $\eta,\eta_0\downarrow 0$. Let

$$X_\eta^{\eta_0}=\operatorname{Im}\mathcal{G}^{\lambda+i(\eta^4+\eta_0)}(v_k,w_k)-\operatorname{Im}\mathcal{G}^{\lambda+i\eta^4}(v_k,w_k),$$
$$X^{\eta_0}=\operatorname{Im}\mathcal{G}^{\lambda+i\eta_0}(v_k,w_k)-\operatorname{Im}\mathcal{G}^{\lambda+i0}(v_k,w_k),\ \text{ and } Y_\eta^{\eta_0}=X_\eta^{\eta_0}-X^{\eta_0}.$$

Denote $\sum_{v_k,w_k}=\sum_{(v_0;v_k),(w_0;w_k),v_0=w_0=o}$. For any $M>0$, we have

$$\int\mathbb{E}\sum_{v_k,w_k}|Y_\eta^{\eta_0}|^r=\int\mathbb{E}\sum_{v_k,w_k}|Y_\eta^{\eta_0}|^r1_{|Y_\eta^{\eta_0}|\le M}+\int\mathbb{E}\sum_{v_k,w_k}|Y_\eta^{\eta_0}|^r1_{|Y_\eta^{\eta_0}|>M}.$$

By Proposition 2.4, $\sum_{v_k,w_k}|Y_\eta^{\eta_0}|^r\to 0$ for Lebesgue-a.e. $\lambda\in\mathbb{R}$ and $\mathbb{P}$-a.e. $[\mathcal{T},o,\mathcal{W}]\in\mathscr{T}_*^{D,A}$ as $\eta\downarrow 0$. So the first term tends to 0 by dominated convergence. For the second, for any $s>r$,

$$\int\mathbb{E}\sum_{v_k,w_k}|Y_\eta^{\eta_0}|^r1_{|Y_\eta^{\eta_0}|>M}\le\frac{1}{M^{s-r}}\int\mathbb{E}\sum_{v_k,w_k}|Y_\eta^{\eta_0}|^s\le\frac{C_s}{M^{s-r}}$$

by Remark A.4 and Fatou's lemma. This vanishes as $M\to\infty$. Thus, $\int\mathbb{E}\sum_{v_k,w_k}|Y_\eta^{\eta_0}|^r\to 0$ as $\eta\downarrow 0$. Similarly, $\int\mathbb{E}\sum_{v_k,w_k}|X^{\eta_0}|^r\to 0$ as $\eta_0\downarrow 0$. Since $|X_\eta^{\eta_0}|^r\le 2^{r-1}(|Y_\eta^{\eta_0}|^r+|X^{\eta_0}|^r)$, it follows that $\int\mathbb{E}\sum_{v_k,w_k}|X_\eta^{\eta_0}|^r\to 0$ as $\eta\downarrow 0$ followed by $\eta_0\downarrow 0$. □

By virtue of Lemma 4.5, denoting $\Psi_{\gamma,v}(w)=\operatorname{Im}\tilde g^{\gamma}(v,w)$, the term in parentheses in (4.7) may be replaced by

$$\left(\frac{\Psi_{z+i\eta_0,\tilde x_k}(\tilde y_k)}{\zeta_{e_k}^{z+i\eta_0}\zeta_{e'_k}^{\bar z-i\eta_0}}-\frac{\Psi_{z+i\eta_0,\tilde x_k}(\tilde y_{k-1})}{\zeta_{e_k}^{z+i\eta_0}}-\frac{\Psi_{z+i\eta_0,\tilde x_{k-1}}(\tilde y_k)}{\zeta_{e'_k}^{\bar z-i\eta_0}}+\Psi_{z+i\eta_0,\tilde x_{k-1}}(\tilde y_{k-1})\right).\tag{4.8}$$

Recall that $e_k=(x_{k-1},x_k)$, $e'_k=(y_{k-1},y_k)$ and that there are non-backtracking paths $(x_0,x_1,\ldots,x_{k-1},x_k)$ and $(x_0,x_1,\ldots,y_{k-1},y_k)$. Moreover, $\rho_G(x_0)\ge d_{R,\eta,k}\ge k$.

Suppose $e'_k\ne e_k$. Then there is a path $(v_0,\ldots,v_s)$ with $v_0=\tilde x_k$, $v_1=\tilde x_{k-1}$, $v_{s-1}=\tilde y_{k-1}$ and $v_s=\tilde y_k$. Taking the complex conjugate in identity (2.13), noting that $\Psi_{z+i\eta_0,v}(w)$ is real, we see that (4.8) is zero. If $e_k=e'_k$, then (2.12) tells us that (4.8) equals $\frac{|\operatorname{Im}\zeta_{x_{k-1}}^{z+i\eta_0}(x_k)|}{|\zeta_{x_{k-1}}^{z+i\eta_0}(x_k)|^2}$.

Since $\rho_G(x_0) \geq k$ in Lemma 4.3, the paths $(x_0, x_1, x_2, \ldots, x_k)$ and $(x_0, x_1, y_2, \ldots, y_k)$ are determined by e_k and e'_k, respectively. So the terms in the sum are only nonzero if $(x_0, x_1, x_2, \ldots, x_k) = (x_0, x_1, y_2, \ldots, y_k)$. Hence, if we make all replacements in Lemmas 4.4 and 4.5, modulo the errors appearing in these lemmas, the expression (4.6) finally takes the form

$$\frac{1}{\pi N} \int_{a-2\eta}^{b+2\eta} \sum_{\rho_G(x_0) \geq d_{R,k,\eta}} \sum_{x_1 \sim x_0} \sum_{(x_2;x_k)} \chi(\lambda) |\alpha_{z+i\eta_0}(x_0, x_1)|^2 |K^{z+i\eta_0}(x_0; x_k)|^2 \cdot \frac{|\operatorname{Im} \zeta^{z+i\eta_0}_{x_{k-1}}(x_k)|}{|\zeta^{z+i\eta_0}_{x_{k-1}}(x_k)|^2} \, \mathrm{d}\lambda \leq \frac{1}{\pi} \int_{a-2\eta}^{b+2\eta} \|K^{z+i\eta_0}\|^2_{z+i\eta_0} \, \mathrm{d}\lambda,$$

where we used that $\chi(\lambda) \leq 1$ on $\mathbb{R}$. Collecting all estimates on the error terms, taking $N \to \infty$, then $\eta \downarrow 0$, then $\eta_0 \downarrow 0$, then $R \to \infty$, we finally get

$$\frac{1}{N} \sum_{j=1}^{N} \chi(\lambda_j) \|\alpha_{\gamma_j} K_B^{\gamma_j} f_j\|^2 \lesssim \frac{1}{\pi} \int_{a-2\eta}^{b+\eta} \|K^{z+i\eta_0}\|^2_{z+i\eta_0} \, \mathrm{d}\lambda.$$

Recalling (4.5), if we prove (4.3), this will complete the proof of Theorem 4.1.

We have

$$\|\overline{\alpha_{\gamma_j}}^{-1} f_j^*\|^2 = \sum_{(x_0,x_1) \in B} \frac{1}{|\operatorname{Im} \zeta^{\gamma_j}_{x_1}(x_0)|} |\psi_j(x_0) - \zeta^{\gamma_j}_{x_1}(x_0) \psi_j(x_1)|^2.$$

Repeating the same arguments, we see that modulo asymptotically vanishing error terms, we have

$$\frac{1}{N} \sum_{\lambda_j \in I} \|\overline{\alpha_{\gamma_j}}^{-1} f_j^*\|^2 \lesssim \frac{3}{\pi N} \int_{a-2\eta}^{b+2\eta} \sum_{\rho_G(x_0) \geq d_{R,\eta}} \sum_{x_1 \sim x_0} \frac{\chi(\lambda)}{|\operatorname{Im} \zeta^{z+i\eta_0}_{x_1}(x_0)|} \cdot \Big[\Psi_{z+i\eta_0, \tilde{x}_0}(\tilde{x}_0) - \zeta^{z+i\eta_0}_{x_1}(x_0) \Psi_{z+i\eta_0, \tilde{x}_1}(\tilde{x}_0) - \overline{\zeta^{z+i\eta_0}_{x_1}(x_0)} \Psi_{z+i\eta_0, \tilde{x}_0}(\tilde{x}_1) + |\zeta^{z+i\eta_0}_{x_1}(x_0)|^2 \Psi_{z+i\eta_0, \tilde{x}_1}(\tilde{x}_1) \Big] \, \mathrm{d}\lambda.$$

The term in square brackets is just $|\operatorname{Im} \zeta^{z+i\eta_0}_{x_1}(x_0)|$ by (2.12). Hence, using $\chi(\lambda) \leq 1$ we get $\frac{1}{N} \sum_{\lambda_j \in I} \|\overline{\alpha_{\gamma_j}}^{-1} f_j^*\|^2 \lesssim \frac{3(|I|+4\eta)D}{\pi}$ for any small $\eta > 0$, and (4.3) follows.

5. Step 2: Invariance property of the quantum variance

In the scheme of Section 1.6, we are now in Step 2. Using the functional equations (3.1) and (3.2) satisfied by f_j, f_j^*, we show that there are certain transformations $\mathcal{R}^{\gamma}_{n,r} : \mathscr{H}_k = \mathbb{C}^{B_k} \to \mathscr{H}_{n+k} = \mathbb{C}^{B_{n+k}}$ that leave the quantum variance (3.3) unchanged.

Recall from Section 3 that $\mathcal{B}(\zeta^{\gamma_j} f_j) = f_j - i\eta_0 \, \tau_+ \psi_j$ and $\mathcal{B}^*(\iota \zeta^{\gamma_j} f_j^*) = f_j^* - i\eta_0 \, \tau_- \psi_j$ if $\gamma_j = \lambda_j + i\eta_0$. So

$$(\mathcal{B}\zeta^{\gamma_j})^2 f_j = \mathcal{B}\zeta^{\gamma_j} f_j - i\eta_0 \mathcal{B}\zeta^{\gamma_j} \tau_+ \psi_j = f_j - i\eta_0 (I + \mathcal{B}\zeta^{\gamma_j}) \tau_+ \psi_j.$$

Iterating r times,

$$(\mathcal{B}\zeta^{\gamma_j})^r f_j = f_j - i\eta_0 \sum_{t=0}^{r-1} (\mathcal{B}\zeta^{\gamma_j})^t \tau_+ \psi_j.$$

Similarly,

$$(\mathcal{B}^*\iota\zeta^{\gamma_j})^{n-r} f_j^* = f_j^* - i\eta_0 \sum_{t'=0}^{n-r-1} (\mathcal{B}^*\iota\zeta^{\gamma_j})^{t'} \tau_- \psi_j.$$

If we define for $r \le n$ and $\gamma \in \mathbb{C} \setminus \mathbb{R}$ the operator $\mathcal{R}_{n,r}^\gamma : \mathscr{H}_k \to \mathscr{H}_{n+k}$ by

$$\begin{aligned}(\mathcal{R}_{n,r}^\gamma K)(x_0; x_{n+k}) &= \overline{\zeta_{x_1}^\gamma(x_0)\zeta_{x_2}^\gamma(x_1)\cdots\zeta_{x_{n-r}}^\gamma(x_{n-r-1})} K(x_{n-r}; x_{n-r+k}) \\ &\quad \cdot \zeta_{x_{n-r+k}}^\gamma(x_{n-r+k+1})\zeta_{x_{n-r+k+1}}^\gamma(x_{n-r+k+2})\cdots\zeta_{x_{n+k-1}}^\gamma(x_{n+k}),\end{aligned}$$

we thus get

$$\begin{aligned}&\langle f_j^*, (\mathcal{R}_{n,r}^{\gamma_j} K)_B f_j \rangle \\ &\quad = \sum_{(x_{n-r};x_{n-r+k})} \overline{\left[(\mathcal{B}^*\iota\zeta^{\gamma_j})^{n-r} f_j^*\right](x_{n-r}, x_{n-r+1})} K(x_{n-r}; x_{n-r+k}) \\ &\quad\quad \cdot \left[(\mathcal{B}\zeta^{\gamma_j})^r f_j\right](x_{n-r+k-1}, x_{n-r+k}) \\ &\quad = \left\langle (\mathcal{B}^*\iota\zeta^{\gamma_j})^{n-r} f_j^*, K_B(\mathcal{B}\zeta^{\gamma_j})^r f_j \right\rangle = \langle f_j^*, K_B f_j\rangle - \mathcal{E}_{n,r,j}(\eta_0, K),\end{aligned}$$

where the $\mathcal{E}$ stands for an "error term" that should vanish as $\eta_0 \downarrow 0$:

$$\begin{aligned}\mathcal{E}_{n,r,j}(\eta_0, K) &= i\eta_0 \sum_{t=0}^{r-1} \langle f_j^*, K_B(\mathcal{B}\zeta^{\gamma_j})^t \tau_+\psi_j\rangle + i\eta_0 \sum_{t'=0}^{n-r-1} \langle (\mathcal{B}^*\iota\zeta^{\gamma_j})^{t'}\tau_-\psi_j, K_B f_j\rangle \\ &\quad + \eta_0^2 \sum_{t=0}^{r-1}\sum_{t'=0}^{n-r-1} \langle (\mathcal{B}^*\iota\zeta^{\gamma_j})^{t'}\tau_-\psi_j, K_B(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j\rangle.\end{aligned}$$

Since this holds for each $1 \le r \le n$ and $K = K^\gamma$, by the triangle inequality we get

$$\operatorname{Var}_{\mathrm{nb},\eta_0}^{\mathrm{I}}(K^\gamma) \le \operatorname{Var}_{\mathrm{nb},\eta_0}^{\mathrm{I}}\Big(\frac{1}{n}\sum_{r=1}^n \mathcal{R}_{n,r}^\gamma K^\gamma\Big) + \frac{1}{N}\sum_{\lambda_j\in I}\Big|\frac{1}{n}\sum_{r=1}^n \mathcal{E}_{n,r,j}(\eta_0, K^\gamma)\Big|. \tag{5.1}$$

We first show that the latter term may be neglected.

Lemma 5.1. *Suppose $K^\gamma \in \mathscr{H}_k$ satisfies assumptions* (Hol), *and let $\bar{I} \subseteq I_1$. Then for all $n \in \mathbb{N}$,*

$$\lim_{\eta_0\downarrow 0}\limsup_{N\to\infty}\left(\frac{1}{N}\sum_{\lambda_j\in I}\Big|\frac{1}{n}\sum_{r=1}^n \mathcal{E}_{n,r,j}(\eta_0, K^\gamma)\Big|\right)^2 = 0.$$

Proof. We have

$$\left(\frac{1}{N}\sum_{\lambda_j\in I}\Big|\frac{1}{n}\sum_{r=1}^n \mathcal{E}_{n,r,j}\Big|\right)^2 \le \frac{1}{n}\sum_{r=1}^n\left(\frac{1}{N}\sum_{\lambda_j\in I}|\mathcal{E}_{n,r,j}|\right)^2.$$

Now, letting $\gamma_j = \lambda_j + i\eta_0$ as above,

$$\Big(\sum_{\lambda_j\in I}|\mathcal{E}_{n,r,j}|\Big)^2 \le \eta_0^2 c_{n,r}\Big\{\sum_{t=0}^{r-1}\Big(\sum_{\lambda_j\in I}\Big|\Big\langle f_j^*, K_B^{\gamma_j}(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j\Big\rangle\Big|\Big)^2 + \sum_{t'=0}^{n-r-1}\Big(\sum_{\lambda_j\in I}\Big|\Big\langle(\mathcal{B}^*\iota\zeta^{\gamma_j})^{t'}\tau_-\psi_j, K_B^{\gamma_j}f_j\Big\rangle\Big|\Big)^2 + \eta_0^2\sum_{t,t'}\Big(\sum_{\lambda_j\in I}\Big|\Big\langle(\mathcal{B}^*\iota\zeta^{\gamma_j})^{t'}\tau_-\psi_j, K_B^{\gamma_j}(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j\Big\rangle\Big|\Big)^2\Big\},$$

where $c_{n,r} = n+r(n-r)$. So it suffices to show that $\limsup_N \big(\frac{1}{N}\sum_{\lambda_j\in I}|\langle\cdot,\cdot\rangle|\big)^2$ is uniformly bounded in η_0 for each t, t'. For the first term, we have

$$\Big(\frac{1}{N}\sum_{\lambda_j\in I}|\langle f_j^*, K_B^{\gamma_j}(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j\rangle|\Big)^2 \le \frac{1}{N}\sum_{\lambda_j\in I}\|\overline{\alpha_{\gamma_j}}^{-1}f_j^*\|^2 \cdot \frac{1}{N}\sum_{\lambda_j\in I}\|\alpha_{\gamma_j}K_B^{\gamma_j}(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j\|^2.$$

The first sum is uniformly bounded as $\eta_0 \downarrow 0$, by (4.3). Next, by (2.3), we have

$$\|\alpha_{\gamma_j}K_B^{\gamma_j}(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j\|^2 = \sum_{(x_0,x_1)\in B}\sum_{(x_2;x_k),(y_2;y_k)}|\alpha_{\gamma_j}(x_0,x_1)|^2K^{\gamma_j}(x_0;x_k) \cdot\overline{K^{\gamma_j}(x_0;y_k)}\cdot[(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j](x_{k-1},x_k)\overline{[(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j](y_{k-1},y_k)}.$$

Arguing as in Section 4, applying Lemma 4.2 to Lemma 4.4, for $z = \lambda + i\eta^4$ we get

$$\frac{1}{N}\sum_{\lambda_j\in I}\|\alpha_{\gamma_j}K_B^{\gamma_j}(\mathcal{B}\zeta^{\gamma_j})^t\tau_+\psi_j\|^2 \lesssim \frac{3}{\pi N}\int_{a-2\eta}^{b+2\eta}\sum_{\rho_G(x_0)\ge d_{R,k,t,\eta}}\sum_{x_1\sim x_0}\sum_{(x_2;x_{k+t}),(y_2;y_{k+t})} \cdot\chi(\lambda)|\alpha_{z+i\eta_0}(x_0,x_1)|^2K^{z+i\eta_0}(x_0;x_k)\overline{K^{z+i\eta_0}(x_0;y_k)}\zeta_{x_k}^{z+i\eta_0}(x_{k+1}) \cdots\zeta_{x_{k+t-1}}^{z+i\eta_0}(x_{k+t})\overline{\zeta_{y_k}^{z+i\eta_0}(y_{k+1})\cdots\zeta_{y_{k+t-1}}^{z+i\eta_0}(y_{k+t})}\Psi_{z,\tilde{x}_{k+t}}(\tilde{y}_{k+t})\,\mathrm{d}\lambda.$$

Using Hölder's inequality as in Lemma 4.5, we see that as $N \to \infty$, this quantity is uniformly bounded in η, η_0 by (Hol) and (Green). One bounds

$\frac{1}{N}\sum_{\lambda_j}\|K_B^{\gamma_j}f_j\|^2$ similarly. Finally,

$$\begin{aligned}\frac{1}{N}\sum_{\lambda_j\in I}&\|(\mathcal{B}^*\iota\zeta^{\gamma_j})^{t'}\tau_-\psi_j\|^2\\ &\le\frac{D^{t'}}{N}\sum_{\lambda_j\in I}\sum_{(x_0;x_{t'+1})}|\psi_j(x_0)|^2|\zeta_{x_1}^{\gamma_j}(x_0)\cdots\zeta_{x_{t'}}^{\gamma_j}(x_{t'-1})|^2\\ &\lesssim\frac{3D^n}{\pi N}\int_{a-2\eta}^{b+2\eta}\sum_{(x_0;x_{t'+1}),\rho_G(x_0)\ge d_{R,\eta,t'}}\chi(\lambda)\Psi_{z,\tilde{x}_0}(\tilde{x}_0)|\zeta_{x_1}^{z+i\eta_0}(x_0)\\ &\quad\cdots\zeta_{x_{t'}}^{z+i\eta_0}(x_{t'-1})|^2\,\mathrm{d}\lambda,\end{aligned}$$

which is asymptotically bounded using Hölder's inequality as before. □

Using the invariance law (5.1), Theorem 4.1 with $\tilde{K}^\gamma=\frac{1}{n}\sum_{r=1}^n\mathcal{R}_{n,r}^\gamma K^\gamma$, and Lemma 5.1, we deduce the following statement:

PROPOSITION 5.2. *Under the assumptions of Theorem 4.1,*

$$\begin{aligned}&\lim_{\eta_0\downarrow 0}\limsup_{N\to+\infty}\mathrm{Var}_{\mathrm{nb},\eta_0}^{\mathrm{I}}(K^\gamma)^2\\ &\le D\,|I|\lim_{\eta_0\downarrow 0}\lim_{\eta\downarrow 0}\limsup_{N\to\infty}\int_{a-2\eta}^{b+2\eta}\left\|\frac{1}{n}\sum_{r=1}^n\mathcal{R}_{n,r}^{\lambda+i(\eta^4+\eta_0)}K^{\lambda+i(\eta^4+\eta_0)}\right\|_{\lambda+i(\eta^4+\eta_0)}^2\mathrm{d}\lambda.\end{aligned}$$

6. Step 3: A stationary Markov chain appears

Denoting $\gamma=\lambda+i(\eta^4+\eta_0)$ in Proposition 5.2, we are now concerned with estimating

$$\left\|\frac{1}{n}\sum_{r=1}^n\mathcal{R}_{n,r}^\gamma K^\gamma\right\|_\gamma^2=\frac{1}{n^2}\sum_{r,r'=1}^n\left\langle\mathcal{R}_{n,r}^\gamma K^\gamma,\mathcal{R}_{n,r'}^\gamma K^\gamma\right\rangle_\gamma. \tag{6.1}$$

Suppose $r\ge r'$, so that $n-r\le n-r'$. Then

$$\begin{aligned}&\langle\mathcal{R}_{n,r}^\gamma K,\mathcal{R}_{n,r'}^\gamma K\rangle_\gamma\\ &\quad=\frac{1}{N}\sum_{(x_0;x_{n+k})\in B_{n+k}}\frac{|\operatorname{Im}\zeta_{x_1}^\gamma(x_0)|}{|\zeta_{x_1}^\gamma(x_0)|^2}\cdot|\zeta_{x_1}^\gamma(x_0)\cdots\zeta_{x_{n-r}}^\gamma(x_{n-r-1})|^2\\ &\qquad\cdot|\zeta_{x_{n-r'+k}}^\gamma(x_{n-r'+k+1})\cdots\zeta_{x_{n+k-1}}^\gamma(x_{n+k})|^2\\ &\qquad\cdot\overline{K(x_{n-r};x_{n-r+k})\,\zeta_{x_{n-r+k}}^\gamma(x_{n-r+k+1})\cdots\zeta_{x_{n-r'+k-1}}^\gamma(x_{n-r'+k})}\\ &\qquad\cdot\overline{\zeta_{x_{n-r+1}}^\gamma(x_{n-r})\cdots\zeta_{x_{n-r'}}^\gamma(x_{n-r'-1})}K(x_{n-r'};x_{n-r'+k})\\ &\qquad\cdot\frac{|\operatorname{Im}\zeta_{x_{n+k-1}}^\gamma(x_{n+k})|}{|\zeta_{x_{n+k-1}}^\gamma(x_{n+k})|^2}.\end{aligned}$$

Letting $\eta_1 = \operatorname{Im}\gamma$, (2.9) tells us that

$$\sum_{x_0 \in \mathcal{N}_{x_1} \setminus \{x_2\}} |\operatorname{Im} \zeta^{\gamma}_{x_1}(x_0)| = \frac{|\operatorname{Im} \zeta^{\gamma}_{x_2}(x_1)|}{|\zeta^{\gamma}_{x_2}(x_1)|^2} - \eta_1.$$

Similarly, we have

$$\sum_{x_{n+k} \in \mathcal{N}_{x_{n+k-1}} \setminus \{x_{n+k-2}\}} |\operatorname{Im} \zeta^{\gamma}_{x_{n+k-1}}(x_{n+k})| = \frac{|\operatorname{Im} \zeta^{\gamma}_{x_{n+k-2}}(x_{n+k-1})|}{|\zeta^{\gamma}_{x_{n+k-2}}(x_{n+k-1})|^2} - \eta_1.$$

By iteration, this induces some simplifications:

$$\begin{aligned}
&\langle \mathcal{R}^{\gamma}_{n,r} K, \mathcal{R}^{\gamma}_{n,r'} K\rangle_{\gamma} \\
&\quad = \frac{1}{N} \sum_{(x_{n-r}; x_{n-r'+k}) \in B_{k+r-r'}} \frac{|\operatorname{Im} \zeta^{\gamma}_{x_{n-r+1}}(x_{n-r})|}{|\zeta^{\gamma}_{x_{n-r+1}}(x_{n-r})|^2} \overline{K(x_{n-r}; x_{n-r+k})} \\
&\qquad \cdot K(x_{n-r'}; x_{n-r'+k}) \cdot \overline{\zeta^{\gamma}_{x_{n-r+k}}(x_{n-r+k+1}) \cdots \zeta^{\gamma}_{x_{n-r'+k-1}}(x_{n-r'+k})} \\
&\qquad \cdot \overline{\zeta^{\gamma}_{x_{n-r+1}}(x_{n-r}) \cdots \zeta^{\gamma}_{x_{n-r'}}(x_{n-r'-1})} \\
&\qquad \cdot \frac{|\operatorname{Im} \zeta^{\gamma}_{x_{n-r'+k-1}}(x_{n-r'+k})|}{|\zeta^{\gamma}_{x_{n-r'+k-1}}(x_{n-r'+k})|^2} - \mathbf{E}_{n,r,r'}(\eta_1, K),
\end{aligned} \tag{6.2}$$

with the error term

$$\begin{aligned}
\mathbf{E}_{n,r,r'}(\eta_1, K) &= \frac{\eta_1}{N} \sum_{s=1}^{n-r} \sum_{(x_s; x_{n+k})} |\zeta^{\gamma}_{x_{s+1}}(x_s) \cdots \zeta^{\gamma}_{x_{n-r}}(x_{n-r-1})|^2 \\
&\quad \cdot |\zeta^{\gamma}_{x_{n-r'+k}}(x_{n-r'+k+1}) \cdots \zeta^{\gamma}_{x_{n+k-2}}(x_{n+k-1})|^2 \cdot |\operatorname{Im} \zeta^{\gamma}_{x_{n+k-1}}(x_{n+k})| \\
&\quad \cdot \overline{K(x_{n-r}; x_{n-r+k})\, \zeta^{\gamma}_{x_{n-r+k}}(x_{n-r+k+1}) \cdots \zeta^{\gamma}_{x_{n-r'+k-1}}(x_{n-r'+k})} \\
&\quad \cdot \overline{\zeta^{\gamma}_{x_{n-r+1}}(x_{n-r}) \cdots \zeta^{\gamma}_{x_{n-r'}}(x_{n-r'-1})} K(x_{n-r'}; x_{n-r'+k}) \\
&\quad + \frac{\eta_1}{N} \sum_{s'=n-r'+k}^{n+k-1} \sum_{(x_{n-r'}; x_{s'})} \frac{|\operatorname{Im} \zeta^{\gamma}_{x_{n-r+1}}(x_{n-r})|}{|\zeta^{\gamma}_{x_{n-r+1}}(x_{n-r})|^2} \\
&\quad \cdot |\zeta^{\gamma}_{x_{n-r'+k}}(x_{n-r'+k+1}) \cdots \zeta^{\gamma}_{x_{s'-1}}(x_{s'})|^2 \\
&\quad \cdot \overline{K(x_{n-r}; x_{n-r+k})\, \zeta^{\gamma}_{x_{n-r+k}}(x_{n-r+k+1}) \cdots \zeta^{\gamma}_{x_{n-r'+k-1}}(x_{n-r'+k})} \\
&\quad \cdot \overline{\zeta^{\gamma}_{x_{n-r+1}}(x_{n-r}) \cdots \zeta^{\gamma}_{x_{n-r'}}(x_{n-r'-1})} K(x_{n-r'}; x_{n-r'+k}).
\end{aligned}$$

The expression is slightly nicer if we replace K by $Z_{\gamma} K$ defined by

$$(Z_{\gamma} K)(x_0; x_k) = \zeta^{\gamma}_{x_0}(x_1) \cdots \zeta^{\gamma}_{x_{k-1}}(x_k) K(x_0; x_k). \tag{6.3}$$

If $\gamma \mapsto K^\gamma$ satisfies (Hol), then so does $\gamma \mapsto Z_\gamma K^\gamma$. Using (2.7), in that case we get

$$\begin{aligned}
&\langle \mathcal{R}^\gamma_{n,r} Z_\gamma K^\gamma, \mathcal{R}^\gamma_{n,r'} Z_\gamma K^\gamma \rangle_\gamma \\
&\quad = \frac{1}{N} \sum_{(x_{n-r};x_{n-r'+k})\in B_{k+r-r'}} \frac{|\operatorname{Im} \zeta^\gamma_{x_{n-r+1}}(x_{n-r})|}{|m^\gamma_{x_{n-r+1}}|^2 |\zeta^\gamma_{x_{n-r}}(x_{n-r+1})|^2} \\
&\qquad \cdot |\zeta_{x_{n-r}}(x_{n-r+1}) \cdots \zeta_{x_{n-r'+k-1}}(x_{n-r'+k})|^2 \overline{m^\gamma_{x_{n-r}} K^\gamma(x_{n-r}; x_{n-r+k})} \\
&\qquad \cdot m^\gamma_{x_{n-r'}} K^\gamma(x_{n-r'}; x_{n-r'+k}) \cdot u^\gamma_{x_{n-r+1}}(x_{n-r}) \cdots u^\gamma_{x_{n-r'}}(x_{n-r'-1}) \\
&\qquad \cdot \frac{|\operatorname{Im} \zeta^\gamma_{x_{n-r'+k-1}}(x_{n-r'+k})|}{|\zeta^\gamma_{x_{n-r'+k-1}}(x_{n-r'+k})|^2} - \mathbf{E}_{n,r,r'}(\eta_1, Z_\gamma K^\gamma),
\end{aligned} \tag{6.4}$$

where $u^\gamma_x(y)$ is the complex number of modulus 1 given by

$$u^\gamma_x(y) = \overline{\zeta^\gamma_x(y)} \zeta^\gamma_x(y)^{-1}. \tag{6.5}$$

Let us define a positive measure μ^γ_k on the set B_k of non-backtracking paths of length k, by putting

$$\mu^\gamma_k[(x_0;x_k)] = \frac{|\operatorname{Im} \zeta^\gamma_{x_1}(x_0)|}{|m^\gamma_{x_1} \zeta^\gamma_{x_0}(x_1)|^2} \cdot |\zeta_{x_0}(x_1) \cdots \zeta_{x_{k-1}}(x_k)|^2 \cdot \frac{|\operatorname{Im} \zeta^\gamma_{x_{k-1}}(x_k)|}{|\zeta^\gamma_{x_{k-1}}(x_k)|^2}. \tag{6.6}$$

Let us also introduce the operator

$$\begin{aligned}
&(\mathcal{S}_{u^\gamma} K)(x_0; x_k) \\
&\quad = \frac{|\zeta^\gamma_{x_1}(x_0)|^2}{|\operatorname{Im} \zeta^\gamma_{x_1}(x_0)|} \sum_{x_{-1}\in \mathcal{N}_{x_0}\setminus\{x_1\}} |\operatorname{Im} \zeta^\gamma_{x_0}(x_{-1})| \, \overline{u^\gamma_{x_0}(x_{-1})} K(x_{-1}; x_{k-1}).
\end{aligned} \tag{6.7}$$

Then, using (2.7) again, we see that (6.4) takes the nicer form

$$\langle \mathcal{R}^\gamma_{n,r} Z_\gamma K^\gamma, \mathcal{R}^\gamma_{n,r'} Z_\gamma K^\gamma \rangle_\gamma = \frac{1}{N} \langle \mathcal{S}^{r-r'}_{u^\gamma} m^\gamma K^\gamma, m^\gamma K^\gamma \rangle_{\ell^2(\mu^\gamma_k)} - \mathbf{E}_{n,r,r'}(\eta_1, Z_\gamma K^\gamma), \tag{6.8}$$

where we let $(m^\gamma K)(x;y) = m^\gamma_x K(x;y)$. Let us also define

$$(\mathcal{S}_\gamma K)(x_0; x_k) = \frac{|\zeta^\gamma_{x_1}(x_0)|^2}{|\operatorname{Im} \zeta^\gamma_{x_1}(x_0)|} \sum_{x_{-1}\in \mathcal{N}_{x_0}\setminus\{x_1\}} |\operatorname{Im} \zeta^\gamma_{x_0}(x_{-1})| \, K(x_{-1}; x_{k-1}). \tag{6.9}$$

Such operators would be called "transfer operators" in ergodic theory, or "transition matrices" in the theory of Markov chains. Note that $\mathcal{S}_\gamma$ has non-negative coefficients and that $\mathcal{S}_{u^\gamma}$ just differs from $\mathcal{S}_\gamma$ by the "phases" $\overline{u^\gamma_{x_0}(x_{-1})}$. The effect of adding a phase to a stochastic operator is a much studied subject in the theory of Markov chains, or more generally in ergodic theory. (See Wielandt's theorem [35, Ch. 8], or in the context of hyperbolic dynamical systems [36, Ch. 4].)

The matrix elements of $\mathcal{S}_\gamma$ are given by

$$\mathcal{S}_\gamma(\omega,\omega') = \frac{|\zeta^\gamma_{x_1}(x_0)|^2}{|\operatorname{Im}\zeta^\gamma_{x_1}(x_0)|}|\operatorname{Im}\zeta^\gamma_{x_0}(x_{-1})| \tag{6.10}$$

if $\omega = (x_0;x_k)$, $\omega' = (x_{-1};x_{k-1})$ and $\omega' \rightsquigarrow \omega$, and $\mathcal{S}_\gamma(\omega,\omega') = 0$ otherwise. Recall from Section 2.1 that if $\omega = (x_0;x_k)$, we write $\omega' \rightsquigarrow \omega$ if $\omega' = (x_{-1},x_0,\dots,x_{k-1})$ for some $x_{-1} \in \mathcal{N}_{x_0} \setminus \{x_1\}$.

Note that $\mathcal{S}_\gamma$ is substochastic: $\sum_{\omega'\in B_k}\mathcal{S}_\gamma(\omega,\omega') \le 1$ for any $\omega \in B_k$, by (2.9). More precisely, if $\omega = (x_0;x_k)$ and $\eta_1 = \operatorname{Im}\gamma > 0$, then

$$\sum_{\omega'\in B_k}\mathcal{S}_\gamma(\omega,\omega') = 1 - \eta_1\frac{|\zeta^\gamma_{x_1}(x_0)|^2}{|\operatorname{Im}\zeta^\gamma_{x_1}(x_0)|}. \tag{6.11}$$

Taking the adjoint in $\ell^2(\mu^\gamma_k)$, a direct calculation gives

$$(\mathcal{S}^*_\gamma K)(x_0;x_k) = \frac{|\zeta^\gamma_{x_{k-1}}(x_k)|^2}{|\operatorname{Im}\zeta^\gamma_{x_{k-1}}(x_k)|}\sum_{x_{k+1}\in\mathcal{N}_{x_k}\setminus\{x_{k-1}\}}|\operatorname{Im}\zeta^\gamma_{x_k}(x_{k+1})|\,K(x_1;x_{k+1}).$$

The adjoint $\mathcal{S}^*_\gamma$ is also substochastic, with

$$\sum_{\omega'\in B_k}\mathcal{S}^*_\gamma(\omega,\omega') = 1 - \eta_1\frac{|\zeta^\gamma_{x_{k-1}}(x_k)|^2}{|\operatorname{Im}\zeta^\gamma_{x_{k-1}}(x_k)|}. \tag{6.12}$$

Remark 6.1. By (2.9), for any $(x_0;x_{k-1}) \in B_{k-1}$, we have

$$\sum_{x_k\in\mathcal{N}_{x_{k-1}}\setminus\{x_{k-2}\}}\mu^\gamma_k\left[(x_0;x_k)\right] \le \mu^\gamma_{k-1}\left[(x_0;x_{k-1})\right] \tag{6.13}$$

and for any $(x_1;x_k) \in B_{k-1}$,

$$\sum_{x_0\in\mathcal{N}_{x_1}\setminus\{x_2\}}\mu^\gamma_k\left[(x_0;x_k)\right] \le \mu^\gamma_{k-1}\left[(x_1;x_k)\right]. \tag{6.14}$$

In (6.1) we take $\gamma = \lambda + i(\eta^4+\eta_0)$ (cf. Proposition 5.2), and thus $\eta_1 = \operatorname{Im}\gamma = \eta^4 + \eta_0$. In the limiting case $\eta_1 = 0$, (6.13) and (6.14) turn into equalities. Equation (6.13) is then the Kolmogorov compatibility condition: it tells us that the family of measures (μ^γ_k) may be extended to a positive measure (actually, a Markov measure) on the set B_∞ of infinite non-backtracking paths. Equality in condition (6.14) means that this Markov chain is stationary. This stationarity is the property that makes the measures μ^γ_k nice, and this is the reason for introducing (somewhat artificially) the weight $\frac{\operatorname{Im}\zeta^\gamma_x(y)}{|\zeta^\gamma_x(y)|^2}$ in (4.1).

This family of stationary Markov chains (indexed by γ) is in some sense the "classical dynamical system" that we were seeking in Section 1.6.

Since $\eta_1 = \eta^4 + \eta_0$ is non-zero (but small), we do not actually have exact equality in (6.13) and (6.14). This causes some error terms that we need to control as $\eta, \eta_0 \longrightarrow 0$.

7. Spectral gap and mixing

In this section, we convert the expanding assumption (EXP) into an estimate on the rate of mixing of the "Markov chains" (μ_k^γ) defined in (6.6). Every transitive Markov chain is mixing, but here we need estimates that are uniform both as $N \longrightarrow +\infty$ *and* as γ approaches the real axis.

A technical difficulty is that the measures (μ_k^γ) are not a priori bounded from above, and the transition probabilities are not bounded from below as γ approaches the real axis. Peaks of (μ_k^γ), as well as small transition probabilities, tend to "disconnect" the graph and are bad for mixing. So we will need to show that there are few peaks and few small transitions (Proposition 7.6).

Let

$$\nu_k^\gamma = \frac{1}{\mu_k^\gamma(B_k)} \mu_k^\gamma \tag{7.1}$$

be the normalized measure. We denote by $\ell^2(\nu_k^\gamma)$ the set $\ell^2(B_k)$ endowed with the scalar product $\langle f, g\rangle_{\nu_k^\gamma} = \sum_{\omega \in B_k} \nu_k^\gamma(\omega)\overline{f(\omega)}g(\omega)$.

We anticipate the calculations of Section 10, where we will need to consider the non-backtracking quantum variance of operators K_γ of the form $K_\gamma = \mathcal{F}_\gamma K$ where K is independent of γ, and $\mathcal{F}_\gamma : \mathscr{H}_m \to \mathscr{H}_k$ is a γ-dependent operator for some $1 \le k \le m+1$, having the form $\mathcal{F}_\gamma = \mathcal{L}^\gamma d^{-1} \mathcal{S}_{T,\gamma}$, $\mathcal{T}^\gamma$, $\mathcal{O}_1^\gamma$, $\mathcal{U}_j^\gamma$, $\mathcal{O}_j^\gamma$, $\mathcal{P}_j^\gamma$, $j \ge 2$, or a polynomial combination thereof. See (10.3), (10.4), (10.5), (10.7), (10.8), and (10.9) for the definitions. In the case $\mathcal{F}_\gamma = \mathcal{L}^\gamma d^{-1}\mathcal{S}_{T,\gamma}$, the operator depends on an additional parameter $T \in \mathbb{N}^*$, which has to be taken arbitrarily large in Corollary 10.3.

Comparing with equation (6.8), this means that we will need to deal with $\langle \mathcal{S}_{u^\gamma}^{r-r'} K^\gamma, K^\gamma\rangle_{\mu_k^\gamma}$, where now $K^\gamma = B_\gamma K$, K is γ-independent, and $B_\gamma : \mathscr{H}_m \to \mathscr{H}_k$ is defined by

$$B_\gamma = m^\gamma Z_\gamma^{-1} \mathcal{F}_\gamma.$$

For simplicity, the calculations below are written for $k = 1$. This suffices for our purposes, as we shall see in Section 9. Like in the statement of Theorem 1.3, we will always assume that the γ-independent operator K satisfies $\|K\|_\infty := \sup_{x,y\in V} |K(x,y)| \le 1$.

The main results of this section are the two following propositions, which estimate the norm of the transfer operator $\mathcal{S}_\gamma$ (6.10) on proper subspaces. We call F the space of functions f on B such that $f(e)$ "depends only on the terminus," that is, $f(e) = f(e')$ if $t_e = t_{e'}$. The first proposition estimates the norm of $\mathcal{S}_\gamma$ on the orthogonal of F, and the second one estimates the norm of $\mathcal{S}_\gamma^2$ on the orthogonal of constant functions.

We denote by $\ell^2(B_1, U)$ the set $\ell^2(B_1)$ endowed with the scalar product $\langle f, g\rangle_U = \frac{1}{N}\sum_{e\in B_1} \overline{f(e)}g(e)$. Let $P_{F,U}$ be the orthogonal projector on F in $\ell^2(B_1, U)$:

$$P_{F,U}K(e) = \frac{1}{d(t_e)} \sum_{e' : t_{e'}=t_e} K(e'). \tag{7.2}$$

We use as a "reference operator" the transfer operator $\mathcal{S}$ defined by

$$\mathcal{S} : \ell^2(B, U) \longrightarrow \ell^2(B, U)$$
$$\mathcal{S}f(e) = \frac{1}{q(o_e)} \sum_{e' \rightsquigarrow e} f(e'),$$

where $q(x) = d(x) - 1$. Both $\mathcal{S}$ and $\mathcal{S}^*$ are stochastic if the adjoint of $\mathcal{S}$ is taken in $\ell^2(B_1, U)$. The influence of the spectral gap assumption (EXP) on the spectrum of $\mathcal{S}$ is studied in [6], and we will use these results below.

We denote $\mathcal{Q} = \mathcal{S}^*\mathcal{S}$ and $\mathcal{Q}_2 = \mathcal{S}^{2\,*}\mathcal{S}^2$. Note that $\mathcal{Q}(e, e') = 0$ unless there exists e'' such that $e \rightsquigarrow e''$ and $e' \rightsquigarrow e''$. In this case, we say that $[e, e']$ is a *pair*; $[e, e']$ form a pair if and only if they share the same terminus. The set of pairs is denoted by $P(B_1)$.

PROPOSITION 7.1. *Let $B_\gamma K \in \mathscr{H}_1$. Let $w = P_{F^\perp,\nu} B_\gamma K$ be the orthogonal projection of $B_\gamma K$ on $F^\perp$ in $\ell^2(\nu_1^\gamma)$. Then for any $M > 0$, we have*

$$\|\mathcal{S}_\gamma w\|^2_{\nu_1^\gamma} \le (1 - \tfrac{3}{4}M^{-2}) \cdot \|w\|^2_{\nu_1^\gamma} + C_{N,M}(B_\gamma) \cdot \|K\|^2_\infty,$$

where

$$\begin{aligned} C_{N,M}(B_\gamma) = \sup_{\|K\|_\infty=1} \frac{M^{-1}}{2N} &\sum_{[e,e']\in \mathrm{Badp}(M)} \mathcal{Q}(e, e')|B_\gamma K(e) - B_\gamma K(e')|^2 \\ &+ M^{-2} \sum_{e \in \mathrm{Bad}(M)} \nu_1^\gamma(e)|B_\gamma K(e) - P_{F,U} B_\gamma K(e)|^2. \end{aligned} \tag{7.3}$$

The sets $\mathrm{Bad}(M)$ of bad edges and $\mathrm{Badp}(M)$ of bad pairs of edges will be defined in the course of the proof. They correspond to the aforementioned peaks of μ_1^γ and problems of small transition probabilities. If there were no bad edges and bad pairs, Proposition 7.1 would be a genuine spectral gap estimate.

PROPOSITION 7.2. *Let $B_\gamma K \in \mathscr{H}_1$. Let $f = P_{\mathbf{1}^\perp,\nu} B_\gamma K$ be the orthogonal projection of $B_\gamma K$ on $\mathbf{1}^\perp$ in $\ell^2(\nu_1^\gamma)$. Then for any $M > 0$ we have*

$$\|\mathcal{S}_\gamma^2 f\|^2_{\nu_1^\gamma} \le (1 - M^{-2}c(D, \beta)) \cdot \|f\|^2_{\nu_1^\gamma} + C_{N,M,2}(B_\gamma) \cdot \|K\|^2_\infty,$$

where $c(D,\beta) > 0$ *is explicit and depends only on* D (*upper bound on the degree*) *and the spectral gap* β *of* (EXP), *and*

$$C_{N,M,2}(B_\gamma) = \sup_{\|K\|_\infty=1} \frac{M^{-1}}{2N} \sum_{[e,e']\in \mathrm{Badp}(2,M)} \mathcal{Q}_2(e,e')|B_\gamma K(e) - B_\gamma K(e')|^2 + M^{-2} \sum_{e\,\in \mathrm{Bad}(M)} \nu_1^\gamma(e)|B_\gamma K(e) - P_{\mathbf{1},U} B_\gamma K(e)|^2,$$

where $P_{\mathbf{1},U}$ *is the orthogonal projector on* $\mathbf{1}$ *in* $\ell^2(B_1, U)$.

Here, $\mathrm{Badp}(2,M)$ is another set of bad edge-couples defined in the proof. The quantities $C_{N,M}(B_\gamma), C_{N,M,2}(B_\gamma)$ are estimated in Proposition 7.7.

Proof of Proposition 7.1. Let $\mathcal{Q}^\gamma = \mathcal{S}_\gamma^* \mathcal{S}_\gamma$ (where now the adjoint is considered in $\ell^2(\nu_1^\gamma)$). The operator $\mathcal{Q}^\gamma$ being self-adjoint on $\ell^2(\nu_1^\gamma)$ is equivalent to the relation

$$\nu_1^\gamma(e)\mathcal{Q}^\gamma(e,e') = \nu_1^\gamma(e')\mathcal{Q}^\gamma(e',e) \tag{7.4}$$

for all $e, e' \in B_1$. Note that $\mathcal{Q}^\gamma(e,e') = 0$ unless $[e,e']$ is a pair.

Define $D^\gamma(e) = \sum_{e'} \mathcal{Q}^\gamma(e,e') \le 1$ and $\mathcal{M}^\gamma(e,e') = D^\gamma(e)\delta_{e=e'} - \mathcal{Q}^\gamma(e,e')$. Then using (7.4), we have the *Dirichlet identity*

$$\frac{1}{2}\sum_{e,e'} \nu_1^\gamma(e)\mathcal{Q}^\gamma(e,e')|K(e) - K(e')|^2 = \langle K, \mathcal{M}^\gamma K\rangle_{\nu_1^\gamma}. \tag{7.5}$$

We observe that for any $K \in \ell^2(\nu_1^\gamma)$,

$$\|\mathcal{S}_\gamma K\|_{\nu_1^\gamma} \le \|K\|_{\nu_1^\gamma}. \tag{7.6}$$

Indeed, denoting $\langle\cdot,\cdot\rangle_\nu := \langle\cdot,\cdot\rangle_{\nu_1^\gamma}$, by Dirichlet we have $\|\mathcal{S}_\gamma K\|_\nu^2 = \langle K, \mathcal{Q}^\gamma K\rangle_\nu$ and $\langle K, \mathcal{M}^\gamma K\rangle_\nu \ge 0$, so $\|K\|_\nu^2 \ge \langle K, D^\gamma K\rangle_\nu \ge \langle K, \mathcal{Q}^\gamma K\rangle_\nu$ as claimed.

Remark 7.3. The Dirichlet identity shows that $F = \{K \in \mathbb{C}^B : \mathcal{M}^\gamma K = 0\} = \{K \in \mathbb{C}^B : (I - \mathcal{Q})K = 0\}$.

Remark 7.4. If $J \perp F$ in $\ell^2(B_1, U)$, then $\langle J, (I-\mathcal{Q})J\rangle_U \ge \frac{3}{4}\|J\|_U^2$. Indeed, $\langle \tau_+\delta_y, J\rangle_U = 0$ for all $y \in V$, so $\sum_{x\sim y} J(x,y) = 0$ for all $y \in V$ and thus $(\mathcal{Q}J)(x_0,x_1) = (\mathcal{S}^*\mathcal{S}J)(x_0,x_1) = \frac{J(x_0,x_1)}{q(x_1)^2}$. (Recall that $q(x) = d(x) - 1$ where $d(x)$ is the degree of x.) As $\min q(x) \ge 2$, we get $\|\mathcal{Q}J\|_U \le \frac{1}{4}\|J\|_U$ and the claim follows.

Fix a large $M > 0$. We call $e \in B_1$ *bad* if $\nu_1^\gamma(e) > \frac{M}{N}$. We call a pair $[e,e'] \in P(B_1)$ *bad* if $\nu_1^\gamma(e)\mathcal{Q}^\gamma(e,e') < \frac{M^{-1}}{N}$. We call $\mathrm{Bad}(M)$ and $\mathrm{Badp}(M)$ the sets of bad e and $[e,e']$, respectively.

To prove Proposition 7.1, we first note that by (7.5), and letting $K_\gamma = B_\gamma K$, we have

$$\begin{aligned}
\|w\|_\nu^2 - \|\mathcal{S}_\gamma w\|_\nu^2 &\geq \langle w, \mathcal{M}^\gamma w\rangle_\nu = \langle K_\gamma, \mathcal{M}^\gamma K_\gamma\rangle_\nu \\
&= \frac{1}{2}\sum_{[e,e']\in P(B_1)} \nu_1^\gamma(e)\mathcal{Q}^\gamma(e,e')|K_\gamma(e)-K_\gamma(e')|^2 \\
&\geq \frac{M^{-1}}{2N}\sum_{[e,e']\notin \mathrm{Badp}(M)} \mathcal{Q}(e,e')|K_\gamma(e)-K_\gamma(e')|^2 \qquad (7.7)\\
&= M^{-1}\langle K_\gamma, (I-\mathcal{Q})K_\gamma\rangle_U \\
&\quad - \frac{M^{-1}}{2N}\sum_{[e,e']\in \mathrm{Badp}(M)} \mathcal{Q}(e,e')|K_\gamma(e)-K_\gamma(e')|^2,
\end{aligned}$$

where we used $\mathcal{Q}(e,e') \leq 1$. By Remark 7.4,

$$\begin{aligned}
\langle K_\gamma, (I-\mathcal{Q})K_\gamma\rangle_U &= \langle K_\gamma - P_{F,U}K_\gamma, (I-\mathcal{Q})(K_\gamma - P_{F,U}K_\gamma)\rangle_U \\
&\geq \frac{3}{4}\cdot \|K_\gamma - P_{F,U}K_\gamma\|_U^2.
\end{aligned}$$

Now

$$\begin{aligned}
&\|K_\gamma - P_{F,U}K_\gamma\|_U^2 \geq M^{-1}\sum_{e\notin \mathrm{Bad}(M)} \nu_1^\gamma(e)|K_\gamma(e) - P_{F,U}K_\gamma(e)|^2 \qquad (7.8)\\
&\quad = M^{-1}\|K_\gamma - P_{F,U}K_\gamma\|_\nu^2 - M^{-1}\sum_{e\in \mathrm{Bad}(M)} \nu_1^\gamma(e)|K_\gamma(e) - P_{F,U}K_\gamma(e)|^2 \\
&\quad \geq M^{-1}\|w\|_\nu^2 - M^{-1}\sum_{e\in \mathrm{Bad}(M)} \nu_1^\gamma(e)|K_\gamma(e) - P_{F,U}K_\gamma(e)|^2.
\end{aligned}$$

We used that $\|K_\gamma - P_{F,U}K_\gamma\|_\nu^2 \geq \|w\|_\nu^2$ since $w = P_{F^\perp,\nu}(K_\gamma - P_{F,U}K_\gamma)$. The result is obtained by putting together (7.7) and (7.8). □

Proof of Proposition 7.2. We now let $\mathcal{Q}_2^\gamma = \mathcal{S}_\gamma^{2*}\mathcal{S}_\gamma^2$ (where the adjoint is taken in $\ell^2(\nu_1^\gamma)$). Then $\mathcal{Q}_2^\gamma(e,e') \neq 0$ if and only if there exists e'', e_1, e_1' such that $e \rightsquigarrow e_1 \rightsquigarrow e''$ and $e' \rightsquigarrow e_1' \rightsquigarrow e''$. We denote the set of such couples $[e,e']$ by $P_2(B_1)$, and let $\mathcal{M}_2^\gamma(e,e') = D_2\delta_{e=e'} - \mathcal{Q}_2(e,e')$, where

$$D_2(e) = \sum_{e'} \mathcal{Q}_2^\gamma(e,e') \leq 1.$$

Fix $M > 0$. We say that $[e,e'] \in P_2(B_1)$ is *bad* if $\nu_1^\gamma(e)\mathcal{Q}_2(e,e') < \frac{M^{-1}}{N}$. We call $\mathrm{Badp}(2,M)$ the set of bad couples in $P_2(B_1)$.

The proof is then similar to Proposition 7.1, replacing the space F by the space of constant functions and using [6, Th. 1.1] instead of Remark 7.4. In particular, the quantity $c(\beta, D)$ is the one appearing in [6, Th. 1.1]. □

Later on, we will need to iterate the result of Proposition 7.2, considering $\mathcal{S}_\gamma^{2r}$ instead of $\mathcal{S}_\gamma^2$. Since $\mathcal{S}_\gamma^*$ is not exactly stochastic, $\mathcal{S}_\gamma$ does not preserve the orthogonal of constants. Still, we can iterate (6.12) to get $\mathcal{S}_\gamma^{*\,l}\mathbf{1} = 1 - \eta_1 \sum_{s=0}^{l-1} \mathcal{S}_\gamma^{*\,s}\xi^\gamma$, where $\xi^\gamma(x_0,x_1) = \frac{|\zeta_{x_0}^\gamma(x_1)|^2}{|\operatorname{Im}\zeta_{x_0}^\gamma(x_1)|}$. Hence, for any K, we have $\langle \mathbf{1}, \mathcal{S}_\gamma^l K\rangle_\nu = \langle \mathbf{1}, K\rangle_\nu - \eta_1 \langle \sum_{s=0}^{l-1}\mathcal{S}_\gamma^{*\,s}\xi^\gamma, K\rangle_\nu$. Denoting

$$\mathcal{Z}_l K := \xi^\gamma \sum_{s=0}^{2l-1} \mathcal{S}_\gamma^s K, \qquad \mathcal{Z}_0 K := 0,$$

we see that if $K \perp \mathbf{1}$, then $(\mathcal{S}_\gamma^{2l} K + \eta_1 \mathcal{Z}_l K) \perp \mathbf{1}$.

Proposition 7.5. *Let $K \in \mathscr{H}_m$. Let $f = P_{\mathbf{1}^\perp,\nu} B_\gamma K$ be the orthogonal projection of $B_\gamma K$ on $\mathbf{1}^\perp$ in $\ell^2(\nu_1^\gamma)$. Then for any $M > 0$, we have*

$$\|\mathcal{S}_\gamma^{2r} f\|_\nu \le \left(1 - M^{-2}c(D,\beta)\right)^{r/2} \|f\|_\nu + \sum_{l=0}^{r-1} C_{N,M,l,2}(B_\gamma)^{1/2}\|K\|_\infty + 2\eta_1 \sum_{l=1}^{r-1} \|\mathcal{Z}_l f\|_\nu,$$

where $C_{N,M,l,2}(B_\gamma) = C_{N,M,2}((\mathcal{S}_\gamma^{2l} + \eta_1 \mathcal{Z}_l) P_{\mathbf{1}^\perp,\nu} B_\gamma)$.

Proof. The proof is by induction on r. This holds for $r = 1$ by Proposition 7.2. Assume the result holds for r. If $f \perp \mathbf{1}$, we have just seen that $(\mathcal{S}_\gamma^{2r} + \eta_1 \mathcal{Z}_r) f \perp \mathbf{1}$ in $\ell^2(\nu_1^\gamma)$. So using Proposition 7.2 and (7.6),

$$\begin{aligned}\|\mathcal{S}_\gamma^{2(r+1)} f\|_\nu &\le \|\mathcal{S}_\gamma^2(\mathcal{S}_\gamma^{2r} + \eta_1\mathcal{Z}_r) f\|_\nu + \eta_1\|\mathcal{Z}_r f\|_\nu \\ &\le \left(1 - M^{-2}c(D,\beta)\right)^{1/2} \|(\mathcal{S}_\gamma^{2r} + \eta_1 \mathcal{Z}_r) f\|_\nu \\ &\quad + C_{N,M,r,2}(B_\gamma)^{1/2}\|K\|_\infty + \eta_1 \|\mathcal{Z}_r f\|_\nu.\end{aligned}$$

Since $\|(\mathcal{S}_\gamma^{2r} + \eta_1\mathcal{Z}_r) f\| \le \|\mathcal{S}_\gamma^{2r} f\| + \eta_1 \|\mathcal{Z}_r f\|$, the claim follows. ◻

The rest of this section is devoted to estimating the "bad" quantities.

Proposition 7.6. *Under assumptions* (BSCT) *and* (Green), *for any $s \ge 1$, there exists C_s such that for all $M > 1$, we have*

$$\sup_{\eta_1\in(0,1)} \limsup_{N\to\infty} \sup_{\operatorname{Re}\gamma\in I_1, \operatorname{Im}\gamma=\eta_1} \nu_1^\gamma(\mathrm{Bad}(M)) \le C_s M^{-s}$$

and

$$\limsup_{N\to\infty} \frac{\#\mathrm{Badp}(M)}{N} \le C_s M^{-s}.$$

Proof. We have $\nu_1^\gamma(\mathrm{Bad}) = \nu_1^\gamma\{e : \nu_1^\gamma(e) > \frac{M}{N}\}$, so

$$\nu_1^\gamma(\mathrm{Bad}) \le M^{-s}N^s \sum_{e\in B_1} \nu_1^\gamma(e)\nu_1^\gamma(e)^s = M^{-s}\Big(\frac{N}{\mu_1^\gamma(B_1)}\Big)^{s+1}\frac{1}{N}\sum_{e\in B_1}\mu_1^\gamma(e)^{s+1}.$$

Recalling the definition of μ_1^γ (6.6), and using Remark A.3, we get

$$\Big(\frac{N}{\mu_1^\gamma(B_k)}\Big)^{s+1}\frac{1}{N}\sum_{e\in B_1}\mu_1^\gamma(e)^{s+1}$$
$$\underset{N\longrightarrow+\infty}{\longrightarrow}\frac{1}{\mathbb{E}[\sum_{o'\sim o}\hat\mu_1^\gamma(o,o')]^{s+1}}\,\mathbb{E}\left[\sum_{o'\sim o}\hat\mu_1^\gamma(o,o')^{s+1}\right]$$

uniformly in $\operatorname{Re}\gamma\in I_1$, for any fixed $\operatorname{Im}\gamma=\eta_1$. By Remark A.4, this is bounded by some constant C_s. The second assertion is proved similarly. □

PROPOSITION 7.7. *For all $t\in\mathbb{N}$,*

$$\begin{aligned}C_{N,M}(\mathcal{S}^t_{u^\gamma}B_\gamma)\le{}&\frac{2M^{-1}}{N}\#\mathrm{Badp}(M)^{1/3}\left(\sum_e\frac{1}{\nu_1^\gamma(e)}\right)^{1/3}\\&\cdot\left(\sum_e\nu_1^\gamma(e)\Big(\sum_w|B_\gamma(e,w)|\Big)^6\right)^{1/3}\\&+2M^{-2}\nu_1^\gamma(\mathrm{Bad}(M))^{1/2}\left(\sum_e\nu_1^\gamma(e)\Big(\sum_w|B_\gamma(e,w)|\Big)^4\right)^{1/2}\\&+2M^{-2}\nu_1^\gamma(\mathrm{Bad}(M))^{1/2}\left(\sum_e\frac{[(P_{F,U}\nu_1^\gamma)(e)]^2}{\nu_1^\gamma(e)}\right)^{1/4}\\&\cdot\left(\sum_e\nu_1^\gamma(e)\Big(\sum_w|B_\gamma(e,w)|\Big)^8\right)^{1/4},\end{aligned}$$

where $(P_{F,U}\nu_1^\gamma)(e)=\frac{1}{d(t_e)}\sum_{t_{e'}=t_e}\nu_1^\gamma(e')$, and

(7.9)
$$\begin{aligned}C_{N,M,2}(\mathcal{S}^t_{u^\gamma}B_\gamma)\le{}&\frac{2M^{-1}}{N}\#\mathrm{Badp}(2,M)^{1/3}\\&\cdot\left(\sum_e\frac{1}{\nu_1^\gamma(e)}\right)^{1/3}\left(\sum_e\nu_1^\gamma(e)\Big(\sum_w|B_\gamma(e,w)|\Big)^6\right)^{1/3}\\&+2M^{-2}\nu_1^\gamma(\mathrm{Bad}(M))^{1/2}\left(\sum_e\nu_1^\gamma(e)\Big(\sum_w|B_\gamma(e,w)|\Big)^4\right)^{1/2}\\&+2M^{-2}\nu_1^\gamma(\mathrm{Bad}(M))^{1/2}\left(\frac{1}{N^2}\sum_e\frac{1}{\nu_1^\gamma(e)}\right)^{1/4}\\&\cdot\left(\sum_e\nu_1^\gamma(e)\Big(\sum_w|B_\gamma(e,w)|\Big)^8\right)^{1/4}.\end{aligned}$$

Similar estimates hold if B_γ is replaced by $P_{\mathbf{1}^\perp,\nu}B_\gamma$, where $P_{\mathbf{1}^\perp,\nu}$ is the projection on the orthogonal of constants in $\ell^2(\nu_1^\gamma)$.

We first deduce the following corollary. Recall that the operators $\mathcal{F}_\gamma$ from Corollary 10.3 depend on a parameter $T \in \mathbb{N}^*$, and $B_\gamma = m^\gamma Z_\gamma^{-1}\mathcal{F}_\gamma$. In this section, T is fixed, but it will be taken to $+\infty$ in Section 10.

Corollary 7.8. *For any $s > 0$, there exists $C_{s,T}$ such that, for all M,*

$$\sup_{\eta_1\in(0,1)} \limsup_{N\to\infty} \sup_{\mathrm{Re}\,\gamma\in I_1, \mathrm{Im}\,\gamma=\eta_1} \sup_{t\in\mathbb{N}} C_{N,M}(\mathcal{S}^t_{u^\gamma} B_\gamma) \le C_{s,T} M^{-s}$$

and

$$\sup_{\eta_1\in(0,1)} \limsup_{N\to\infty} \sup_{\mathrm{Re}\,\gamma\in I_1, \mathrm{Im}\,\gamma=\eta_1} \sup_{t\in\mathbb{N}} C_{N,M,2}(\mathcal{S}^t_{u^\gamma} B_\gamma) \le C_{s,T} M^{-s}.$$

Similar estimates hold if B_γ is replaced by $P_{\mathbf{1}^\perp,\nu} B_\gamma$.

Proof of Corollary 7.8. This will follow from Propositions 7.6 and 7.7 if we show that

$$\sup_{\eta_1\in(0,1)} \limsup_{N\to\infty} \sup_{\mathrm{Re}\,\gamma\in I_1, \mathrm{Im}\,\gamma=\eta_1} N^{-2} \sum_e \frac{1}{\nu_1^\gamma(e)} < +\infty, \tag{7.10}$$

$$\sup_{\eta_1\in(0,1)} \limsup_{N\to\infty} \sup_{\mathrm{Re}\,\gamma\in I_1, \mathrm{Im}\,\gamma=\eta_1} \sum_e \nu_1^\gamma(e) \Big(\sum_w |B_\gamma(e,w)| \Big)^\alpha < +\infty \tag{7.11}$$

($\alpha = 4, 6, 8$), and

$$\sup_{\eta_1\in(0,1)} \limsup_{N\to\infty} \sup_{\mathrm{Re}\,\gamma\in I_1, \mathrm{Im}\,\gamma=\eta_1} \sum_e \frac{1}{\nu_1^\gamma(e)} \frac{1}{d(t_e)^2} \Big(\sum_{t_{e'}=t_e} \nu_1^\gamma(e') \Big)^2 < +\infty. \tag{7.12}$$

For (7.10), we have by Remark A.3 that

$$N^{-2} \sum_e \frac{1}{\nu_1^\gamma(e)} = \frac{\sum_e \mu_1^\gamma(e)}{N} \cdot \frac{1}{N} \sum_e \frac{1}{\mu_1^\gamma(e)}$$
$$\xrightarrow[N\to\infty]{} \mathbb{E}\left(\sum_{o'\sim o} \hat{\mu}_1^\gamma(o,o') \right) \cdot \mathbb{E}\left(\sum_{o'\sim o} \frac{1}{\hat{\mu}_1^\gamma(o,o')} \right)$$

uniformly in $\mathrm{Re}\,\gamma \in I_1$, for any fixed $\mathrm{Im}\,\gamma = \eta_1$. So the claim follows Remark A.4.

For (7.12), we have

$$\frac{1}{d(t_e)^2} \Big(\sum_{t_{e'}=t_e} \nu_1^\gamma(e') \Big)^2 \le \sum_{t_{e'}=t_e} \nu_1^\gamma(e')^2,$$

so we deduce the upper bound $D(\sum_e \frac{1}{\nu_1^\gamma(e)^2})^{1/2} (\sum_e \nu_1^\gamma(e)^4)^{1/2}$, which is uniformly bounded by

$$\frac{D}{\mathbb{E}\left(\sum_{o'\sim o} \hat{\mu}_1^\gamma(o,o')\right)} \mathbb{E}\left(\sum_{o'\sim o} \frac{1}{\hat{\mu}_1^\gamma(o,o')^2} \right)^{1/2} \mathbb{E}\left(\sum_{o'\sim o} \hat{\mu}_1^\gamma(o,o')^4 \right)^{1/2}.$$

Finally, for (7.11), we write

$$\sum_e \nu_1^\gamma(e)\Big(\sum_w |B_\gamma(e,w)|\Big)^\alpha \leq \frac{N}{\sum_e \mu_1^\gamma(e)}\left(\frac{1}{N}\sum_e \frac{\mu_1^\gamma(e)^2|m_{o_e}^\gamma|^{2\alpha}}{|\zeta_{o_e}^\gamma(t_e)|^{2\alpha}}\right)^{1/2}\left(\frac{1}{N}\sum_e\Big(\sum_w|\mathcal{F}_\gamma(e,w)|\Big)^{2\alpha}\right)^{1/2}.$$

The first two terms are bounded by $\frac{1}{\mathbb{E}(\sum_{o'\sim o}\hat{\mu}_1^\gamma(o,o'))}(\mathbb{E}\sum_{o'\sim o}\frac{\hat{\mu}_1^\gamma(o,o')^2|\hat{m}_o^\gamma|^{2\alpha}}{|\hat{\zeta}_o^\gamma(o')|^{2\alpha}})^{1/2}$ and the last term is shown to be uniformly bounded in Remark 10.4. This completes the proof. □

Proof of Proposition 7.7. An important point here is to obtain a bound that does not depend on t. Recalling (7.3), we first estimate

$$\begin{aligned}&\sum_{[e,e']\in\mathrm{Badp}(M)} \mathcal{Q}(e,e')|\mathcal{S}_{u^\gamma}^t B_\gamma K(e)-\mathcal{S}_{u^\gamma}^t B_\gamma K(e')|^2\\ &\leq 4\sum_{[e,e']\in\mathrm{Badp}(M)}\mathcal{Q}(e,e')|\mathcal{S}_{u^\gamma}^t B_\gamma K(e)|^2 = 4\sum_e n(e)|\mathcal{S}_{u^\gamma}^t B_\gamma K(e)|^2,\end{aligned} \tag{7.13}$$

where $n(e)=\sum_{e':[e,e']\in\mathrm{Badp}(M)}\mathcal{Q}(e,e')$. Using Hölder, this is less than

$$4\left(\sum_e n^3(e)\right)^{1/3}\left(\sum_e\frac{1}{\nu_1^\gamma(e)}\right)^{1/3}\left(\sum_e \nu_1^\gamma(e)|\mathcal{S}_{u^\gamma}^t B_\gamma K(e)|^6\right)^{1/3}.$$

But again by Hölder and the fact that $\mathcal{Q}$ is stochastic, we have

$$\sum_e n^3(e)\leq\sum_e\Big(\sum_{e'}\mathbb{1}_{[e,e']\in\mathrm{Badp}(M)}\Big)\Big(\sum_{e'}\mathcal{Q}(e,e')^{3/2}\Big)^2\leq \#\mathrm{Badp}(M).$$

Next, recalling (6.7) and (6.9), we have $|\mathcal{S}_{u^\gamma}^t B_\gamma K(e)|\leq(\mathcal{S}_\gamma^t|B_\gamma K|)(e)$. As $\mathcal{S}_\gamma^t$ and $\mathcal{S}_\gamma^{*\,t}$ are substochastic, and $\nu_1^\gamma(e)\mathcal{S}_\gamma^t(e,e')=\nu_1^\gamma(e')\mathcal{S}_\gamma^{*\,t}(e',e)$, we have

$$\begin{aligned}\sum_e \nu_1^\gamma(e)[\mathcal{S}_\gamma^t|B_\gamma K|(e)]^6 &\leq \sum_e\nu_1^\gamma(e)\Big(\sum_{e'}\mathcal{S}_\gamma^t(e,e')\Big)^5\Big(\sum_{e'}\mathcal{S}_\gamma^t(e,e')[|B_\gamma K|(e')]^6\Big)\\ &\leq \sum_{e,e'}\nu_1^\gamma(e')\mathcal{S}_\gamma^{*\,t}(e',e)[|B_\gamma K|(e')]^6\\ &\leq \sum_{e'}\nu_1^\gamma(e')[|B_\gamma K|(e')]^6.\end{aligned}$$

Collecting the estimates, we showed that (7.13) is bounded by

$$4\,(\#\mathrm{Badp}(M))^{1/3}\left(\sum_e\frac{1}{\nu_1^\gamma(e)}\right)^{1/3}\left(\sum_e\nu_1^\gamma(e)\,[|B_\gamma K|(e)]^6\right)^{1/3}.$$

For the second term in (7.3), we have

$$\begin{aligned}(7.14)\qquad &\sum_{e\in \mathrm{Bad}(M)} \nu_1^\gamma(e)|\mathcal{S}^t_{u^\gamma}B_\gamma K(e)-P_{F,U}\mathcal{S}^t_{u^\gamma}B_\gamma K(e)|^2\\ &\le 2\sum_{e\in \mathrm{Bad}(M)} \nu_1^\gamma(e)\Big(\big[\mathcal{S}^t_\gamma|B_\gamma K|(e)\big]^2+\big[P_{F,U}\mathcal{S}^t_\gamma|B_\gamma K|(e)\big]^2\Big)\end{aligned}$$

and again, as $\mathcal{S}^t_\gamma$ and $\mathcal{S}^{*\,t}_\gamma$ are substochastic,

$$\sum_{e\in \mathrm{Bad}(M)} \nu_1^\gamma(e)\left[\mathcal{S}^t_\gamma\,|B_\gamma K|(e)\right]^2 \le \nu_1^\gamma(\mathrm{Bad}(M))^{1/2}\left(\sum_e \nu_1^\gamma(e)\,[|B_\gamma K|(e)]^4\right)^{1/2}.$$

Also,

$$\begin{aligned}&\sum_{e\in \mathrm{Bad}(M)} \nu_1^\gamma(e)\left[P_{F,U}\mathcal{S}^t_\gamma\,|B_\gamma K|(e)\right]^2\\ &\qquad\le \nu_1^\gamma(\mathrm{Bad}(M))^{1/2}\left(\sum_e \nu_1^\gamma(e)\left[P_{F,U}\mathcal{S}^t_\gamma\,|B_\gamma K|(e)\right]^4\right)^{1/2}.\end{aligned}$$

Using that $P_{F,U}$ is stochastic and $\mathcal{S}^t_\gamma$ is substochastic, we have

$$\begin{aligned}\sum_e \nu_1^\gamma(e)\left[P_{F,U}\mathcal{S}^t_\gamma|B_\gamma K|(e)\right]^4 &\le \sum_{e,e'}\nu_1^\gamma(e)P_{F,U}(e,e')\left[\mathcal{S}^t_\gamma\,|B_\gamma K|(e')\right]^4\\ &\le \left(\sum_{e'}\frac{[(P_{F,U}\nu_1^\gamma)(e')]^2}{\nu_1^\gamma(e')}\right)^{1/2}\left(\sum_{e'}\nu_1^\gamma(e')\left[\mathcal{S}^t_\gamma\,|B_\gamma K|(e')\right]^8\right)^{1/2}\\ &\le \left(\sum_{e}\frac{[(P_{F,U}\nu_1^\gamma)(e)]^2}{\nu_1^\gamma(e)}\right)^{1/2}\left(\sum_{e}\nu_1^\gamma(e)\,[|B_\gamma K|(e)]^8\right)^{1/2}.\end{aligned}$$

This yields the first inequality. The second one is proven similarly. □

Remark 7.9. Note that if $\|K\|_\infty\le 1$, then

$$(7.15)\qquad \|B_\gamma K\|^2_{\nu_1^\gamma}=\sum_{e\in B}\nu_1^\gamma(e)|B_\gamma K(e)|^2\le\sum_e\nu_1^\gamma(e)\Big(\sum_w|B_\gamma(e,w)|\Big)^2,$$

so $\sup_{\eta_1>0}\limsup_{N\to\infty}\sup_{\mathrm{Re}\,\gamma\in I_1,\mathrm{Im}\,\gamma=\eta_1}\|B_\gamma K\|^2_{\nu_1^\gamma}\le C_T$ by the proof in Corollary 7.8; see also Remark 10.4.

For a quantity $A(N,\gamma,\kappa)$ depending on N,γ (and possibly on an additional parameter κ), we will write $A(N,\gamma,\kappa)=O_\kappa(1)_{N\longrightarrow+\infty,\gamma}$ to mean that, for any given κ,

$$\sup_{\eta_1\in(0,1)}\limsup_{N\to\infty}\sup_{\mathrm{Re}\,\gamma\in I_1,\mathrm{Im}\,\gamma=\eta_1}A(N,\gamma,\kappa)<+\infty.$$

For instance, if $\|K\|_\infty\le 1$, then $\|B_\gamma K\|^2_{\nu_1^\gamma}=O_T(1)_{N\longrightarrow+\infty,\gamma}$. This is true more generally for $\|B_\gamma K\|^2_{\nu_k^\gamma}$, with $B_\gamma=\frac{m^\gamma}{Z_\gamma}\mathcal{F}_\gamma:\mathscr{H}_m\to\mathscr{H}_k$, and $\mathcal{F}_\gamma$ as in Corollary 10.3.

Similarly, for the operator $\mathcal{Z}_l$ appearing in Proposition 7.5, the arguments in Corollary 7.8 and Remark 10.4 show that $\|\mathcal{Z}_l f\|_{\nu_1^\gamma} = O_{l,T}(1)_{N\longrightarrow+\infty,\gamma}$.

Finally, by Corollary 7.8, $\sup_t C_{N,M,2}(\mathcal{S}^t_{u^\gamma} B_\gamma)$ is uniformly bounded by $C_{s,T}M^{-s}$ for any M and s, as $N \to +\infty$. To express this, we use the notation $O_T(M^{-\infty})_{N\longrightarrow+\infty,\gamma}$.

8. Transition matrices with phases

We now consider the operator $\mathcal{S}_{u^\gamma}$ given in (6.7). If we denote by M_{u^γ} the multiplication operator $(M_{u^\gamma}K)(x_0; x_k) = \overline{u^\gamma_{x_1}(x_0)}K(x_0; x_k)$, where $u^\gamma_{x_1}(x_0)$ is the function of modulus 1 defined in (6.5), then $\mathcal{S}_{u^\gamma} = \mathcal{S}_\gamma M_{u^\gamma}$.

It is well known that putting phases into a matrix with positive entries will strictly diminish its spectral radius, unless the phases satisfy very special relations: this is the contents of Wielandt's theorem [35, Ch. 8]. This is reflected in Proposition 8.1 below. Without the error term, part (i) says that the norm of $\mathcal{S}^4_{u^\gamma}$ is strictly smaller than one, in contrast to $\mathcal{S}^4_\gamma$. (The latter only contracts the norm on proper subspaces; see Section 7.) The contraction property of $\mathcal{S}^4_{u^\gamma}$ holds true except in special cases, described in part (ii) of Proposition 8.1.

Note that we are not using Wielandt's theorem directly, as we want some information on the norm of the operator $\mathcal{S}^4_{u^\gamma}$ instead of its spectral radius. In addition, as in Section 7, we need estimates that are uniform both as $N \to \infty$ and as γ approaches the real axis.

Recall from Section 7 that B_γ is an operator $\mathscr{H}_m \to \mathscr{H}_k$ with $1 \le k \le m$. As in Section 7, the case $k = 1$ suffices for our purposes, but we need more general operators $A_\gamma : \mathscr{H}_m \to \mathscr{H}_1$ defined in terms of B_γ. The quantities $C_{N,M}(A_\gamma), C_{N,M,2}(A_\gamma)$ were introduced in Propositions 7.1 and 7.2. In particular, $C_{N,M,2}(I)$ corresponds to the case where A_γ is the identity operator. The measure ν_1^γ is defined in (6.6) and (7.1).

PROPOSITION 8.1. *Fix $\gamma \in \mathbb{C}^+$, $A_\gamma K \in \mathscr{H}_1$, $\varepsilon \in (0,1)$, $M > 0$ and a graph $G = G_N$. Denote $\eta_1 = \operatorname{Im}\gamma$. Then*

(i) *Either we have*

$$\|\mathcal{S}^4_{u^\gamma} A_\gamma K\|^2_{\nu_1^\gamma} \le (1-\varepsilon)^2 \|A_\gamma K\|^2_{\nu_1^\gamma} + \tilde{C}_{N,M,2}(A_\gamma)\cdot \|K\|^2_\infty \tag{8.1}$$

with

$$\tilde{C}_{N,M,2}(A_\gamma) = \max\{C_{N,M}(A_\gamma), C_{N,M,2}(A_\gamma), C_{N,M}(\mathcal{S}_{u^\gamma}A_\gamma), C_{N,M,2}(\mathcal{S}^2_{u^\gamma}A_\gamma)\},$$

(ii) *or there exist $\theta : V \to \mathbb{R}$ and constants s_j with $|s_j| \le 1$, $j = 1, 2$, such that*

$$\left\| u^\gamma_{x_1}(x_0) - s_2 \frac{e^{-i[\theta(x_0)+\theta(x_1)]}}{n^\gamma_{x_0}} \right\|^2_{\nu_1^\gamma} \le c_{M,\beta}\left[\varepsilon^{1/2} + \eta_1 \|\xi^\gamma\|_{\nu_1^\gamma} + \eta_1^2 \|\xi^\gamma\|^2_{\nu_1^\gamma}\right] + C'_{N,M},$$

and

$$\|u^\gamma_{x_1}(x_0) - s_1 n^\gamma_{x_1} e^{i[\theta(x_0)+\theta(x_1)]}\|^2_{\nu_1^\gamma} \leq c_{M,\beta}\left[\varepsilon^{1/2} + \eta_1 \|\xi^\gamma\|_{\nu_1^\gamma} + \eta_1^2 \|\xi^\gamma\|^2_{\nu_1^\gamma}\right] + C'_{N,M},$$

where

$$\xi^\gamma(x_0,x_1) = \frac{|\zeta^\gamma_{x_0}(x_1)|^2}{|\operatorname{Im}\zeta^\gamma_{x_0}(x_1)|}, \quad n^\gamma_x = (\overline{m^\gamma_x})(m^\gamma_x)^{-1}, \quad C'_{N,M} = \frac{8M^2 C_{N,M,2}(I)}{c(D,\beta)}.$$

Moreover, there is an explicit $f(\beta,D)$, depending only on the spectral gap β and on the degree, such that $c_{M,\beta} \leq f(\beta,D)M^3$ as $M \to +\infty$.

In particular, in case (ii),

(8.2)

$$\|u^\gamma_{x_0}(x_1)u^\gamma_{x_1}(x_0) - s_1 s_2\|^2_{\nu_1^\gamma} \leq 4c_{M,\beta}\left[\varepsilon^{1/2} + \eta_1 \|\xi^\gamma\|_{\nu_1^\gamma} + \eta_1^2 \|\xi^\gamma\|^2_{\nu_1^\gamma}\right] + 4C'_{N,M}.$$

Proof. (a) We start with some preliminary inequalities. Denote $\langle\cdot,\cdot\rangle_\nu = \langle\cdot,\cdot\rangle_{\nu_1^\gamma}$. Recall that we denote by F the space of functions on B that depend only on the terminus.

Let $\delta_1 = \frac{3}{4}M^{-2}$, $K_\gamma = A_\gamma K$, and let $w = P_{F^\perp}K_\gamma$ be the orthogonal projection of K_γ on $F^\perp$ in $\ell^2(\nu_1^\gamma)$. By the proof of Proposition 7.1,

$$\langle w, \mathcal{M}^\gamma w\rangle_\nu \geq \delta_1 \|w\|^2_\nu - C_{N,M}(A_\gamma)\|K\|^2_\infty.$$

By Remark 7.3 and the fact that $\mathcal{M}^{\gamma *} = \mathcal{M}^\gamma$, we have

$$\langle w, \mathcal{M}^\gamma w\rangle_\nu = \langle K_\gamma, \mathcal{M}^\gamma K_\gamma\rangle_\nu \leq \|K_\gamma\|^2_\nu - \|\mathcal{S}_\gamma K_\gamma\|^2_\nu.$$

So if $f = P_F K_\gamma = K_\gamma - w \in F$ is the projection of K_γ on F, we have

$$\|K_\gamma - f\|^2_\nu \leq \delta_1^{-1}\left(\|K_\gamma\|^2_\nu - \|\mathcal{S}_\gamma K_\gamma\|^2_\nu + C_{N,M}(A_\gamma)\|K\|^2_\infty\right). \tag{8.3}$$

Similarly, if $\delta_2 = M^{-2}c(D,\beta)$ and $C\,\mathbf{1} = P_{\mathbf{1}}|K_\gamma|$ is the projection of $|K_\gamma|$ on $\mathbf{1}$, then using Proposition 7.2, we get

$$\|\,|K_\gamma| - C\,\mathbf{1}\|^2_\nu \leq \delta_2^{-1}\left(\|K_\gamma\|^2_\nu - \|\mathcal{S}^2_\gamma\,|K_\gamma|\,\|^2_\nu + C_{N,M,2}(A_\gamma)\|K\|^2_\infty\right). \tag{8.4}$$

Now

$$\left\|K_\gamma - \|K_\gamma\|_\nu \frac{f}{|f|}\right\|_\nu \leq \|K_\gamma - f\|_\nu + \left\|f - \|K_\gamma\|_\nu \frac{f}{|f|}\right\|_\nu$$

and

$$\left\|f - \|K_\gamma\|_\nu \frac{f}{|f|}\right\|_\nu = \|\,|f| - \|K_\gamma\|_\nu\,\mathbf{1}\|_\nu\,.$$

(This is true even if f vanishes, if we give an arbitrary value of modulus 1 to $\frac{f}{|f|}$ in this case.) Also,

$$\|\,|f| - \|K_\gamma\|_\nu\,\mathbf{1}\|_\nu \leq \|\,|K_\gamma| - |f|\|_\nu + \|\,|K_\gamma| - \|K_\gamma\|_\nu\,\mathbf{1}\|_\nu$$

and

$$\|\,|K_\gamma| - |f|\|_\nu \leq \|K_\gamma - f\|_\nu.$$

Finally,

$$\| \, |K_\gamma| - \|K_\gamma\|_\nu \, \mathbf{1} \|_\nu \leq \| \, |K_\gamma| - C \, \mathbf{1} \|_\nu + | \, \|K_\gamma\|_\nu - C \, | \leq 2 \, \| \, |K_\gamma| - C \, \mathbf{1} \|_\nu \, .$$

Putting all these inequalities together, we obtain

$$\left\| K_\gamma - \|K_\gamma\|_\nu \frac{f}{|f|} \right\|_\nu \leq 2 \, \|K_\gamma - f\|_\nu + 2 \, \| \, |K_\gamma| - C \, \mathbf{1} \|_\nu \, . \tag{8.5}$$

Comparing with (8.3) and (8.4), this says the following: If $\|\mathcal{S}_\gamma^2 \, |K_\gamma| \, \|_\nu$ is close to $\|K_\gamma\|_\nu$ and if $\|\mathcal{S}_\gamma K_\gamma\|_\nu$ is close to $\|K_\gamma\|_\nu$, then K_γ must be close to $\|K_\gamma\|_\nu \frac{f}{|f|}$, where f is a function that depends only on the terminus.

Repeating the arguments of (8.3) with $M_{u^\gamma} \mathcal{S}_{u^\gamma} K_\gamma$ instead of K_γ, then taking $\tilde{f} = P_F M_{u^\gamma} \mathcal{S}_{u^\gamma} K_\gamma \in F$, we get

$$\|M_{u^\gamma} \mathcal{S}_{u^\gamma} K_\gamma - \tilde{f}\|_\nu^2 \leq \delta_1^{-1} \left(\|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu^2 - \|\mathcal{S}_{u^\gamma}^2 K_\gamma\|_\nu^2 + C_{N,M}(\mathcal{S}_{u^\gamma} A_\gamma) \|K\|_\infty^2 \right) . \tag{8.6}$$

Similarly to (8.4), if $\tilde{C} \, \mathbf{1} = P_{\mathbf{1}} |\mathcal{S}_{u^\gamma} K_\gamma|$, we get

$$\begin{aligned} &\left\| \, |\mathcal{S}_{u^\gamma} K_\gamma| - \tilde{C} \, \mathbf{1} \right\|_\nu^2 \\ &\leq \delta_2^{-1} \left(\|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu^2 - \left\| \mathcal{S}_\gamma^2 \, |\mathcal{S}_{u^\gamma} K_\gamma| \, \right\|_\nu^2 + C_{N,M,2}(\mathcal{S}_{u^\gamma} A_\gamma) \|K\|_\infty^2 \right) . \end{aligned} \tag{8.7}$$

Finally, arguing as in (8.5), we have

$$\begin{aligned} &\left\| M_{u^\gamma} \mathcal{S}_{u^\gamma} K_\gamma - \|K_\gamma\|_\nu \frac{\tilde{f}}{|\tilde{f}|} \right\|_\nu \\ &\leq 2 \left\| M_{u^\gamma} \mathcal{S}_{u^\gamma} K_\gamma - \tilde{f} \right\|_\nu + 2 \left\| \, |\mathcal{S}_{u^\gamma} K_\gamma| - \tilde{C} \, \mathbf{1} \right\|_\nu + \|K_\gamma\|_\nu - \|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu . \end{aligned} \tag{8.8}$$

(b) We can now start the proof itself. Suppose (i) is not true:

$$\|\mathcal{S}_{u^\gamma}^4 K_\gamma\|_\nu^2 > (1-\varepsilon)^2 \|K_\gamma\|_\nu^2 + \tilde{C}_{N,M,2}(A_\gamma) \cdot \|K\|_\infty^2 .$$

Using

$$\begin{aligned} &\|\mathcal{S}_{u^\gamma}^4 K_\gamma\|_\nu \leq \|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu = \|\mathcal{S}_\gamma M_{u^\gamma} K_\gamma\|_\nu , \\ &\|\mathcal{S}_{u^\gamma}^4 K_\gamma\|_\nu \leq \|\mathcal{S}_\gamma^2 \, |K_\gamma| \|_\nu = \|\mathcal{S}_\gamma^2 \, |M_{u^\gamma} K_\gamma| \|_\nu , \\ &\|\mathcal{S}_{u^\gamma}^4 K_\gamma\|_\nu \leq \|\mathcal{S}_\gamma^2 \, |\mathcal{S}_{u^\gamma} K_\gamma| \|_\nu \text{ and } \|K_\gamma\|_\nu \geq \|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu , \end{aligned}$$

we see that we must also have

$$\begin{aligned} \|K_\gamma\|_\nu^2 - \|\mathcal{S}_\gamma M_{u^\gamma} K_\gamma\|_\nu^2 &< 2\varepsilon \, \|K_\gamma\|_\nu^2 - \tilde{C}_{N,M,2}(A_\gamma) \cdot \|K\|_\infty^2 , \\ \|K_\gamma\|_\nu^2 - \|\mathcal{S}_\gamma^2 \, |M_{u^\gamma} K_\gamma| \|_\nu^2 &< 2\varepsilon \, \|K_\gamma\|_\nu^2 - \tilde{C}_{N,M,2}(A_\gamma) \cdot \|K\|_\infty^2 , \\ \|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu^2 - \|\mathcal{S}_\gamma^2 \, |\mathcal{S}_{u^\gamma} K_\gamma| \|_\nu^2 &< 2\varepsilon \, \|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu^2 - \tilde{C}_{N,M,2}(A_\gamma) \cdot \|K\|_\infty^2 \end{aligned}$$

as well as

$$\|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu^2 - \|\mathcal{S}_{u^\gamma}^2 K_\gamma\|_\nu^2 < 2\varepsilon \, \|\mathcal{S}_{u^\gamma} K_\gamma\|_\nu^2 - \tilde{C}_{N,M,2}(A_\gamma) \cdot \|K\|_\infty^2 .$$

Applying (8.3), (8.4) and (8.5) to $M_{u^\gamma}K_\gamma$ instead of K_γ, and $f = P_F M_{u^\gamma} K_\gamma$, it follows that

$$\left\| M_{u^\gamma}K_\gamma - \|K_\gamma\|_\nu \frac{f}{|f|}\right\|_\nu^2 \leq 16(\delta_1^{-1}+\delta_2^{-1})\,\varepsilon\cdot\|K_\gamma\|_\nu^2. \tag{8.9}$$

Applying (8.6), (8.7) and (8.8) yields

$$\left\| M_{u^\gamma}\mathcal{S}_{u^\gamma}K_\gamma - \|K_\gamma\|_\nu \frac{\tilde f}{|\tilde f|}\right\|_\nu^2 \leq 24(\delta_1^{-1}+\delta_2^{-1})\,\varepsilon\cdot\|K_\gamma\|_\nu^2 + 6\varepsilon\cdot\|K_\gamma\|_\nu^2. \tag{8.10}$$

As $f, \tilde f \in F$, we have $\frac{f}{|f|}(x_0,x_1) = e^{i\theta(x_1)}$ and $\frac{\tilde f}{|\tilde f|}(x_0,x_1) = e^{i\theta'(x_1)}$ for some $\theta,\theta' : V\to\mathbb{R}$. Note that in this case,

$$\left(\mathcal{S}_\gamma \frac{f}{|f|}\right)(x_0,x_1) = e^{i\theta(x_0)} - \eta_1\xi^\gamma(x_1,x_0)e^{i\theta(x_0)},$$

where $\xi^\gamma(x_0,x_1) = \frac{|\zeta^\gamma_{x_0}(x_1)|^2}{|\operatorname{Im}\zeta^\gamma_{x_0}(x_1)|}$, using (6.11). Applying $\mathcal{S}_\gamma$ to (8.9), we thus get

$$\begin{aligned}\left\|\mathcal{S}_{u^\gamma}K_\gamma - \|K_\gamma\|_\nu e^{i\theta(x_0)}\right\|_\nu^2 &\leq 2\left\|\mathcal{S}_\gamma M_{u^\gamma}K_\gamma - \|K_\gamma\|_\nu \mathcal{S}_\gamma\frac{f}{|f|}\right\|_\nu^2 \\ &\quad + 2\eta_1^2\,\|\xi^\gamma\|_\nu^2\cdot\|K_\gamma\|_\nu^2 \\ &\leq 32(\delta_1^{-1}+\delta_2^{-1})\,\varepsilon\cdot\|K_\gamma\|_\nu^2 + 2\eta_1^2\,\|\xi^\gamma\|_\nu^2\cdot\|K_\gamma\|_\nu^2.\end{aligned}$$

Applying M_{u^γ} and comparing with (8.10), it follows that

$$\begin{aligned}\left\|\overline{u^\gamma_{x_1}(x_0)}e^{i\theta(x_0)} - e^{i\theta'(x_1)}\right\|_\nu^2 &\leq (2\times 32 + 2\times 24)(\delta_1^{-1}+\delta_2^{-1})\cdot\varepsilon \\ &\quad + 4\eta_1^2\,\|\xi^\gamma\|_\nu^2 + 12\varepsilon.\end{aligned} \tag{8.11}$$

Repeating the procedure with K_γ replaced by $\mathcal{S}_{u^\gamma}K_\gamma$, and f replaced by $\tilde f$, the same arguments show that there exists $\theta'' : V\to\mathbb{R}$ such that

$$\left\|\overline{u^\gamma_{x_1}(x_0)}e^{i\theta'(x_0)} - e^{i\theta''(x_1)}\right\|_\nu^2 \leq (112\delta_1^{-1} + 112\delta_2^{-1} + 12)\cdot\varepsilon + 4\eta_1^2\|\xi^\gamma\|_\nu^2. \tag{8.12}$$

Thus we have proved $u^\gamma_{x_1}(x_0)$ is close to both $e^{i(\theta(x_0)-\theta'(x_1))}$ and $e^{i(\theta'(x_0)-\theta''(x_1))}$.

(c) Because of relation (2.7), the function u satisfies $u^\gamma_{x_1}(x_0) = u^\gamma_{x_0}(x_1)\frac{n^\gamma_{x_1}}{n^\gamma_{x_0}}$, where $n^\gamma_x = (\overline{m^\gamma_x})(m^\gamma_x)^{-1}$.

To conclude the proof, we show that if $e^{i(\theta(x_0)-\theta'(x_1))}$ and $e^{i(\theta'(x_0)-\theta''(x_1))}$ are close to u^γ, and if the function $u^\gamma_{x_1}(x_0)$ satisfies the relation above, then this gives constraints on θ,θ',θ'' that imply part (ii) of the proposition.

Let $g(x_0,x_1) = e^{i(\theta(x_0)-\theta'(x_1))}$ and $\mathbf{c} = (112\delta_1^{-1} + 112\delta_2^{-1} + 12)$. We have shown in (b) that $\|u^\gamma_{x_1}(x_0) - g\|_\nu^2 \leq \mathbf{c}\varepsilon + 4\eta_1^2\|\xi^\gamma\|_\nu^2$. Recall that we denote by

ι the involution of edge reversal. Hence, if we define $\tilde{g}(x_0, x_1) = g(x_1, x_0)\frac{n^\gamma_{x_1}}{n^\gamma_{x_0}}$, we get

$$\|\tilde{g} - u^\gamma_{x_1}(x_0)\|^2_\nu = \|\iota g - u^\gamma_{x_0}(x_1)\|^2_\nu \leq \mathbf{c}\varepsilon + 4\eta_1^2\,\|\xi^\gamma\|^2_\nu. \tag{8.13}$$

Thus, $\|\tilde{g} - g\|^2_\nu \leq 4\mathbf{c}\varepsilon + 16\eta_1^2\,\|\xi^\gamma\|^2_\nu$. Hence, defining

$$h_1(x_0, x_1) = n^\gamma_{x_1} e^{i[\theta(x_1)+\theta'(x_1)]} \quad \text{and} \quad h_2(x_0, x_1) = n^\gamma_{x_0} e^{i[\theta(x_0)+\theta'(x_0)]},$$

we get

$$\|h_1 - h_2\|^2_\nu = \|\tilde{g} - g\|^2_\nu \leq 4\mathbf{c}\varepsilon + 16\eta_1^2\,\|\xi^\gamma\|^2_\nu.$$

Note that the functions h_1, h_2 have modulus 1, and $\mathcal{S}_\gamma h_1 = h_2 - \eta_1\iota\xi^\gamma h_2$, so

$$\begin{aligned}\|\mathcal{S}^2_\gamma h_1 - h_1\|_\nu \leq 2\,\|\mathcal{S}_\gamma h_1 - h_1\|_\nu &\leq 2\left(\|h_2 - h_1\|_\nu + \eta_1\|\xi^\gamma\|_\nu\right)\\ &\leq 4\mathbf{c}^{1/2}\varepsilon^{1/2} + 8\eta_1\,\|\xi^\gamma\|_\nu.\end{aligned}$$

Consider $P_{\mathbf{1},\nu}h_1 = s\,\mathbf{1}$, the projection of h_1 to the space of constant functions. Arguing as in (8.4), we can write

$$\|h_1 - s\,\mathbf{1}\|^2_\nu \leq \delta_2^{-1}(\|h_1\|^2_\nu - \|\mathcal{S}^2_\gamma h_1\|^2_\nu + 4C_{N,M,2}(I)).$$

But $\|h_1\|^2 - \|\mathcal{S}^2_\gamma h_1\|^2 = (\|h_1\| + \|\mathcal{S}^2_\gamma h_1\|)(\|h_1\| - \|\mathcal{S}^2_\gamma h_1\|) \leq 2\,\|\mathcal{S}^2_\gamma h_1 - h_1\|$. Hence,

$$\|h_1 - s\,\mathbf{1}\|^2_\nu \leq 8\delta_2^{-1}\mathbf{c}^{1/2}\varepsilon^{1/2} + 16\eta_1\delta_2^{-1}\|\xi^\gamma\|_\nu + 4\delta_2^{-1}C_{N,M,2}(I).$$

We observe that

$$\begin{aligned}\|h_1 - s\,\mathbf{1}\| &= \|n^\gamma_{x_1} e^{i(\theta(x_1)+\theta'(x_1))} - s\,\mathbf{1}\|\\ &= \|\tilde{g} n^\gamma_{x_0} e^{i(\theta'(x_0)+\theta'(x_1))} - s\,\mathbf{1}\| = \left\|\tilde{g} - \frac{e^{-i(\theta'(x_0)+\theta'(x_1))}}{n^\gamma_{x_0}} s\right\|.\end{aligned}$$

Thus, comparing with (8.13),

$$\begin{aligned}\left\|u^\gamma_{x_1}(x_0) - s\,\frac{e^{-i(\theta'(x_0)+\theta'(x_1))}}{n^\gamma_{x_0}}\right\|^2_\nu &\leq 16\delta_2^{-1}\mathbf{c}^{1/2}\varepsilon^{1/2} + 32\eta_1\delta_2^{-1}\|\xi^\gamma\|_\nu\\ &\quad + 8\delta_2^{-1}C_{N,M,2}(I) + 2\mathbf{c}\varepsilon + 8\eta_1^2\|\xi^\gamma\|^2_\nu.\end{aligned}$$

This is the first half of (ii) with

$$c_{M,\beta} = \max\{16\delta_2^{-1}\mathbf{c}^{1/2}, 2\mathbf{c}, 32\delta_2^{-1}, 8\}. \tag{8.14}$$

Remembering that $\delta_1 = \frac{3}{4}M^{-2}$, $\delta_2 = M^{-2}c(D, \beta)$ and $\mathbf{c} = (112\delta_1^{-1} + 112\delta_2^{-1} + 12)$, we see that there is an explicit $f(\beta, D)$ such that $c_{M,\beta} \leq f(\beta, D)M^3$ as $M \to +\infty$. Note that $|s| \leq 1$ since $\|h_1\|_\nu = 1$.

The second half of (ii) is proven similarly, using (8.12) instead of (8.11). Here we take $g'(x_0, x_1) = e^{i(\theta'(x_0)-\theta''(x_1))}$, $h'_1(x_0, x_1) = \frac{1}{n^\gamma_{x_1}} e^{-i[\theta'(x_1)+\theta''(x_1)]}$, $s'\mathbf{1} = P_{\mathbf{1}}h'_1$ and $h'_2(x_0, x_1) = \frac{1}{n^\gamma_{x_0}} e^{-i[\theta'(x_0)+\theta''(x_0)]}$.

To prove (8.2), we write

$$\Big\|u^\gamma_{x_1}(x_0)^2 - ss'\frac{n^\gamma_{x_1}}{n^\gamma_{x_0}}\Big\|^2 \le 2\,\Big\|u^\gamma_{x_1}(x_0)[u^\gamma_{x_1}(x_0) - s\frac{e^{-i\widetilde{\theta}(x_0,x_1)}}{n^\gamma_{x_0}}]\Big\|^2 + 2\,\Big\|s\frac{e^{-i\widetilde{\theta}(x_0,x_1)}}{n^\gamma_{x_0}}[u^\gamma_{x_1}(x_0) - s'e^{i\widetilde{\theta}(x_0,x_1)}n^\gamma_{x_1}]\Big\|^2,$$

where we put $\widetilde{\theta}(x_0,x_1) = \theta'(x_0)+\theta'(x_1)$. Since $u^\gamma_{x_1}(x_0)^2\frac{n^\gamma_{x_0}}{n^\gamma_{x_1}} = u^\gamma_{x_1}(x_0)u^\gamma_{x_0}(x_1)$, the proof is complete. □

9. Step 4: End of the proof of Theorem 3.3

Our aim is to show that $\lim_{\eta_0\downarrow 0}\lim_{N\to+\infty}\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K) = 0$ for the operators $\mathcal{F}_\gamma$ that appear in Corollary 10.3. A main step was carried out in Proposition 5.2, and the upper bound was put in a convenient form in (6.8). We now use the estimates of Sections 7 and 8 to complete the proof. We denote $B_\gamma = \frac{m^\gamma}{Z_\gamma}\mathcal{F}_\gamma : \mathscr{H}_m \to \mathscr{H}_k$ as in Section 7, where Z_γ is defined in (6.3). It should be kept in mind that $\mathcal{F}_\gamma$ may depend on a parameter T that is fixed in this section, but will be taken arbitrarily large in the next one.

Recall that we take $\gamma = \lambda + i(\eta^4+\eta_0)$, where λ,η,η_0 come from Proposition 5.2. In other words, $\gamma = \lambda + i\eta_1 \in \mathbb{C}^+$ with $\lambda \in I_1$ and $\eta_1 = \eta^4+\eta_0$. Let $K \in \mathscr{H}_m$ so that $B_\gamma K \in \mathscr{H}_k$. Applying (6.8), recalling that $\nu^\gamma_k = \frac{1}{\mu^\gamma_k(B_k)}\mu^\gamma_k$, we obtain

$$\begin{aligned}&\frac{1}{n^2}\sum_{r,r'=1}^{n}\langle \mathcal{R}^\gamma_{n,r}\mathcal{F}_\gamma K, \mathcal{R}^\gamma_{n,r'}\mathcal{F}_\gamma K\rangle_\gamma\\ &= \frac{\mu^\gamma_k(B_k)}{Nn^2}\sum_{r'\le r\le n}\langle \mathcal{S}^{r-r'}_{u^\gamma}B_\gamma K, B_\gamma K\rangle_{\nu^\gamma_k}\\ &\quad+\frac{\mu^\gamma_k(B_k)}{Nn^2}\sum_{r<r'\le n}\langle B_\gamma K, \mathcal{S}^{r'-r}_{u^\gamma}B_\gamma K\rangle_{\nu^\gamma_k} + \frac{1}{n^2}\sum_{r,r'=1}^{n}\mathbf{E}_{n,r,r'}(\eta_1,\mathcal{F}_\gamma K).\end{aligned}\tag{9.1}$$

Fix M very large and take $n = M^9$. We apply Proposition 8.1 with $\varepsilon = M^{-8}$ to the family of operators $\{\mathcal{S}^{4j}_{u^\gamma}B_\gamma K\}_{j=1}^{M^9}$. Let

$$\tilde{\tilde{C}}_{N,M}(B_\gamma) = \max_{1\le j\le M^9}\tilde{C}_{N,M,2}(\mathcal{S}^{4j+k-1}_{u^\gamma}B_\gamma)^{1/2}\cdot\sqrt{\frac{\mu^\gamma_1(B)}{\mu^\gamma_k(B_k)}}.$$

We use the notation in Remark 7.9 throughout the section. In particular, $\tilde{\tilde{C}}_{N,M}(B_\gamma) = O_T(M^{-\infty})_{N\longrightarrow+\infty,\gamma}$ thanks to Corollary 7.8.

Remark 9.1. It is useful to note that the norm $\|\mathcal{S}^j_{u^\gamma}\|_{\nu^\gamma_k\to\nu^\gamma_k}$ for $k>1$ is controlled by the same norm for $k=1$. To see this, note that for $K \in \ell^2(\nu^\gamma_k)$, we have

$$(\mathcal{S}_{u^\gamma}^{k-1} K)(x_0; x_k) = \sum_{(x_{-k+1};x_{-1})_{x_{0,1}}} \Lambda(x_{-k+1}; x_1) K(x_{-k+1}; x_1)$$

for some function $\Lambda(x_{-k+1}; x_1)$. Here the sum is over those $(x_{-k+1}; x_{-1})$ for which the path $(x_{-k+1}, x_{-k+2}, \ldots, x_{-1}, x_0, x_1)$ does not backtrack; cf. (2.3). So $(\mathcal{S}_{u^\gamma}^{k-1} K)(x_0; x_k)$ only depends on (x_0, x_1). We may define $\phi_K \in \ell^2(\nu_1^\gamma)$ by $\phi_K(x_0, x_1) = (\mathcal{S}_{u^\gamma}^{k-1} K)(x_0; x_k)$. If $\mathscr{I} : \ell^2(\nu_1^\gamma) \to \ell^2(\nu_k^\gamma)$ is the map $(\mathscr{I}\phi)(x_0; x_k) = \phi(x_0, x_1)$, we have for any $j \geq k$, $[\mathcal{S}_{u^\gamma}^{j-k+1} \mathscr{I}\phi_K](x_0; x_k) = (\mathcal{S}_{u^\gamma}^{j} K)(x_0; x_k)$. Moreover, $[\mathcal{S}_{u^\gamma}\mathscr{I}\phi](x_0; x_k) = [\mathscr{I}(\mathcal{S}_{u^\gamma}\phi)](x_0; x_k)$. Thus,

$$\|\mathcal{S}_{u^\gamma}^{j} K\|_{\nu_k}^2 = \|\mathcal{S}_{u^\gamma}^{j-k+1} \mathscr{I}\phi_K\|_{\nu_k}^2 = \|\mathscr{I}(\mathcal{S}_{u^\gamma}^{j-k+1}\phi_K)\|_{\nu_k}^2 \leq \frac{\mu_1^\gamma(B)}{\mu_k^\gamma(B_k)} \|\mathcal{S}_{u^\gamma}^{j-k+1}\phi_K\|_{\nu_1}^2,$$

where we used that $\sum_{x_{0,1}(x_2;x_k)} \mu_k(x_0; x_k) \leq \mu_1(x_0, x_1)$ by (6.13). Hence,

$$\|\mathcal{S}_{u^\gamma}^{j} K\|_{\nu_k}^2 \leq \frac{\mu_1^\gamma(B)}{\mu_k^\gamma(B_k)} \|\mathcal{S}_{u^\gamma}^{j-k+1}\|_{\nu_1 \to \nu_1}^2 \cdot \|\phi_K\|_{\nu_1}^2.$$

But using (2.9) repeatedly, we have

$$\begin{aligned} &\sum_{(x_{-k+1};x_{-1})_{x_{0,1}}} |\Lambda(x_{-k+1}; x_1)| \\ &\quad = \sum_{(x_{-k+1},x_{-1})_{x_{0,1}}} \frac{|\zeta_{x_1}^\gamma(x_0)\zeta_{x_0}^\gamma(x_{-1})\cdots\zeta_{x_{-k+3}}^\gamma(x_{-k+2})|^2 |\operatorname{Im}\zeta_{x_{-k+2}}^\gamma(x_{-k+1})|}{|\operatorname{Im}\zeta_{x_1}^\gamma(x_0)|} \\ &\quad \leq 1, \end{aligned}$$

and $\mu_1^\gamma(x_0, x_1)|\Lambda(x_{-k+1}; x_1)| = \mu_k^\gamma(x_{-k+1}; x_1)$ by (6.6) and (2.7). Hence,

$$\begin{aligned} \|\phi_K\|_{\mu_1}^2 &= \sum_{(x_0,x_1)} \mu_1^\gamma(x_0, x_1) \Bigg| \sum_{(x_{-k+1};x_{-1})_{x_{0,1}}} \Lambda(x_{-k+1}; x_1) K(x_{-k+1}; x_1) \Bigg|^2 \\ &\leq \sum_{(x_0,x_1)} \mu_1^\gamma(x_0, x_1) \sum_{(x_{-k+1};x_{-1})_{x_{0,1}}} |\Lambda(x_{-k+1}; x_1)| \cdot |K(x_{-k+1}; x_1)|^2 \\ &= \sum_{(x_{-k+1};x_1)} \mu_k^\gamma(x_{-k+1}; x_1) \cdot |K(x_{-k+1}; x_1)|^2 = \|K\|_{\mu_k}^2. \end{aligned}$$

So $\|\phi_K\|_{\nu_1}^2 \leq \frac{\mu_k^\gamma(B_k)}{\mu_1^\gamma(B)} \|K\|_{\nu_k^\gamma}^2$. Summarizing, we have shown that for any $j \geq k$, we have

$$\|\mathcal{S}_{u^\gamma}^{j}\|_{\nu_k \to \nu_k} \leq \|\mathcal{S}_{u^\gamma}^{j-k+1}\|_{\nu_1 \to \nu_1}.$$

First alternative. For γ, ε as above, assume that case (i) of Proposition 8.1 is satisfied for all the operators $\{\mathcal{S}_{u^\gamma}^{4j} B_\gamma K\}_{j=1}^{M^9}$. Applying (8.1) for $\mathcal{S}_{u^\gamma}^{4t} B_\gamma K$, $t \le j$, if $k = 1$, we obtain

$$\|\mathcal{S}_{u^\gamma}^{4j} B_\gamma K\|_{\nu_1^\gamma} \le (1-\varepsilon)^j \|B_\gamma K\|_{\nu_1^\gamma} + j \max_{1\le t\le j}\{\tilde{C}_{N,M,2}(\mathcal{S}_{u^\gamma}^{4t} B_\gamma)^{1/2}\} \cdot \|K\|_\infty. \tag{9.2}$$

For higher k, we apply (9.2) to

$$\phi_{B_\gamma K}(x_0,x_1) = (\mathcal{S}_{u^\gamma}^{k-1} B_\gamma K)(x_0;x_k) = (A_\gamma K)(x_0,x_1),$$

where $A_\gamma = \mathcal{S}_{u^\gamma}^{k-1} B_\gamma$, instead of $B_\gamma K$. By Remark 9.1 we get

$$\|\mathcal{S}_{u^\gamma}^{4j+k-1} B_\gamma K\|_{\nu_k^\gamma} \le (1-\varepsilon)^j \|B_\gamma K\|_{\nu_k^\gamma} + j\tilde{\tilde{C}}_{N,M}(B_\gamma) \cdot \|K\|_\infty.$$

Using the euclidean division $r' - r - k + 1 = 4m_{r,r'} + n_{r,r'}$ with $n_{r,r'} < 4$, we see that for $r' - r \ge 4 + k - 1$,

$$\begin{aligned}|\langle B_\gamma K, \mathcal{S}_{u^\gamma}^{r'-r} B_\gamma K\rangle_{\nu_k^\gamma}| \le c_k (1-\varepsilon)^{(r'-r)/4} \|B_\gamma K\|_{\nu_k^\gamma}^2 \\ + n\tilde{\tilde{C}}_{N,M}(B_\gamma) \cdot \|K\|_\infty \|B_\gamma K\|_{\nu_k^\gamma},\end{aligned}$$

where $c_k = \frac{1}{(1-\varepsilon)^{(k-1+n_{r,r'})/4}} \le 2^{\frac{k+2}{4}}$ if $\varepsilon \le \frac12$. Note that $(1-\varepsilon)^{1/4} \le (1-\frac{\varepsilon}{5})$. Hence, since $4 + k - 1 \le 4k$, we have

$$\begin{aligned}&\Big|\sum_{r'\le n}\sum_{r<r'} \langle B_\gamma K, \mathcal{S}_{u^\gamma}^{r'-r} B_\gamma K\rangle_{\nu_k^\gamma}\Big| \\ &\quad\le \Big(\sum_{r'\le n}\sum_{r\le r'-4k} |\langle B_\gamma K, \mathcal{S}_{u^\gamma}^{r'-r} B_\gamma K\rangle_{\nu_k^\gamma}| + 4nk\|B_\gamma K\|_{\nu_k}^2\Big) \\ &\quad\le \Big[4nk + nc_k \sum_{m=1}^n (1-\varepsilon)^{m/4}\Big] \|B_\gamma K\|_{\nu_k^\gamma}^2 + n^3 \tilde{\tilde{C}}_{N,M}(B_\gamma) \cdot \|K\|_\infty \|B_\gamma K\|_{\nu_k^\gamma} \\ &\quad\le \frac{n(5c_k + 4k)}{\varepsilon} \|B_\gamma K\|_{\nu_k^\gamma}^2 + n^3 \tilde{\tilde{C}}_{N,M}(B_\gamma) \cdot \|K\|_\infty \|B_\gamma K\|_{\nu_k^\gamma}.\end{aligned}$$

Recall that $\varepsilon = M^{-8}$ and $n = M^9$. Comparing with (9.1), we get

$$\begin{aligned}&\Big\|\frac1n \sum_{r=1}^n \mathcal{R}_{n,r}^\gamma \mathcal{F}_\gamma K\Big\|_\gamma^2 \\ &\quad\le \frac{\mu_k^\gamma(B_k)}{N}\Big(\frac{c_k'}{M}\|B_\gamma K\|_{\nu_k^\gamma}^2 + M^9 \tilde{\tilde{C}}_{N,M}(B_\gamma) \cdot \|K\|_\infty \|B_\gamma K\|_{\nu_k^\gamma}\Big) \\ &\qquad + \frac{1}{n^2}\sum_{r,r'=1}^n \mathbf{E}_{n,r,r'}(\eta_1, \mathcal{F}_\gamma K).\end{aligned} \tag{9.3}$$

Second alternative. Now assume case (ii) of Proposition 8.1 is satisfied (with some complex numbers $s_j = s_j(N)$ and some function θ). We denote $\| \,\|_\nu = \| \,\|_{\ell^2(\nu_k^\gamma)}$, $\theta_0(x_0;x_k) = \theta(x_0)$, $\theta_1(x_0;x_k) = \theta(x_1)$, $n_0^\gamma(x_0;x_k) = n_{x_0}^\gamma$ and $n_1^\gamma(x_0;x_k) = n_{x_1}^\gamma$. Then we have

PROPOSITION 9.2. *Let $\|K\|_\infty \le 1$. For $A_\gamma K = \mathcal{S}_{u^\gamma}^\ell B_\gamma K$, we have for any $t \in \mathbb{N}^*$,*

$$\begin{aligned}\Big|\langle B_\gamma K, \mathcal{S}_{u^\gamma}^{2t} A_\gamma K\rangle_\nu - (\overline{s_1 s_2})^t \langle B_\gamma K, e^{i\theta_0}\mathcal{S}_\gamma^{2t} e^{-i\theta_0} A_\gamma K\rangle_\nu\Big| \\ \le t\left(c_{M,\beta}\left[\varepsilon^{1/2} + \eta_1 O(1)_{N\longrightarrow+\infty,\gamma}\right] + C'_{N,M}\right)^{1/4} O_T(1)_{N\longrightarrow+\infty,\gamma}.\end{aligned}$$

Proof. Recall that $\mathcal{S}_{u^\gamma} = \mathcal{S}_\gamma M_{u^\gamma}$ with M_{u^γ} the multiplication by $\overline{u_{x_1}^\gamma(x_0)}$. We have

$$\begin{aligned}&\Big\|\mathcal{S}_{u^\gamma}^2 A_\gamma K - \overline{s_1 s_2} e^{i\theta_0}\mathcal{S}_\gamma^2 e^{-i\theta_0} A_\gamma K\Big\|_\nu \\ &= \left\|\mathcal{S}_{u^\gamma}^2 A_\gamma K - \overline{s_1 s_2}\mathcal{S}_\gamma n_0^\gamma e^{i[\theta_0+\theta_1]}\mathcal{S}_\gamma \frac{e^{-i[\theta_0+\theta_1]}}{n_1^\gamma} A_\gamma K\right\|_\nu \\ &\le \Big\|\mathcal{S}_\gamma M_{u^\gamma}\mathcal{S}_\gamma M_{u^\gamma} A_\gamma K - \overline{s_2}\mathcal{S}_\gamma n_0^\gamma e^{i[\theta_0+\theta_1]}\mathcal{S}_\gamma M_{u^\gamma} A_\gamma K\Big\|_\nu \\ &\quad + \left\|\overline{s_2}\mathcal{S}_\gamma n_0^\gamma e^{i[\theta_0+\theta_1]}\mathcal{S}_\gamma M_{u^\gamma} A_\gamma K - \overline{s_1 s_2}\mathcal{S}_\gamma n_0^\gamma e^{i[\theta_0+\theta_1]}\mathcal{S}_\gamma \frac{e^{-i[\theta_0+\theta_1]}}{n_1^\gamma} A_\gamma K\right\|_\nu.\end{aligned}$$

Using (7.6) and Cauchy-Schwarz, the first term is bounded by

$$\Big\|\overline{u_{x_1}^\gamma(x_0)} - \overline{s_2} n_0^\gamma e^{i[\theta_0+\theta_1]}\Big\|_{\ell^4(\nu_k^\gamma)} \|\mathcal{S}_\gamma M_{u^\gamma} A_\gamma K\|_{\ell^4(\nu_k^\gamma)}.$$

But $u^\gamma, s_2, n_0^\gamma$ all have modulus bounded by 1, so

$$|\overline{u_{x_1}^\gamma(x_0)} - \overline{s_2} n_0^\gamma e^{i[\theta_0+\theta_1]}|^4 \le 4\,|\overline{u_{x_1}^\gamma(x_0)} - \overline{s_2} n_0^\gamma e^{i[\theta_0+\theta_1]}|^2.$$

Hence,

$$\|\overline{u_{x_1}^\gamma(x_0)} - \overline{s_2} n_0^\gamma e^{i[\theta_0+\theta_1]}\|_{\ell^4(\nu_1^\gamma)} \le (4c_{M,\beta}\left[\varepsilon^{1/2} + \eta_1 O(1)_{N\longrightarrow+\infty,\gamma}\right] + 4C'_{N,M})^{1/4}$$

by the first part of (ii). For higher k, using $\sum_{x_{0,1}(x_2;x_k)} \mu_k(x_0;x_k) \le \mu_1(x_0,x_1)$ by (6.13), we get

$$\Big\|\overline{u_{x_1}^\gamma(x_0)} - \overline{s_2} n_0^\gamma e^{i[\theta_0+\theta_1]}\Big\|_{\ell^4(\nu_k^\gamma)} \le \left(\frac{\mu_1^\gamma(B)}{\mu_k^\gamma(B_k)}\right)^{1/4} \Big\|\overline{u_{x_1}^\gamma(x_0)} - \overline{s_2} n_0^\gamma e^{i[\theta_0+\theta_1]}\Big\|_{\ell^4(\nu_1^\gamma)}.$$

Next, $\|\mathcal{S}_\gamma M_{u^\gamma} A_\gamma K\|_{\ell^4(\nu_k^\gamma)} = \|\mathcal{S}_{u^\gamma}^{\ell+1} B_\gamma K\|_{\ell^4(\nu_k^\gamma)}$. Arguing as in Proposition 7.7 and Corollary 7.8, we see this is $O_T(1)_{N\longrightarrow+\infty,\gamma}$. Bounding the second term similarly, we get

$$\left\|\mathcal{S}_{u^\gamma}^2 A_\gamma K - \overline{s_1 s_2} e^{i\theta_0} \mathcal{S}_\gamma^2 e^{-i\theta_0} A_\gamma K\right\|_\nu \le \left(c_{M,\beta}\left[\varepsilon^{1/2} + \eta_1 O(1)_{N\longrightarrow+\infty,\gamma}\right] + C'_{N,M}\right)^{1/4} O_T(1)_{N\longrightarrow+\infty,\gamma}.$$

Since $\|B_\gamma K\|_\nu = O_T(1)_{N\longrightarrow+\infty,\gamma}$ (see Remark 7.9), this proves the result for $t=1$.

For higher t, let $X = \overline{s_1 s_2} e^{i\theta_0} \mathcal{S}_\gamma^2 e^{-i\theta_0}$ and $Y = \mathcal{S}_{u^\gamma}^2$. Then

$$\|(X^t - Y^t) A_\gamma K\| = \left\|\sum_{i=1}^t X^{t-i}(X-Y)Y^{i-1} A_\gamma K\right\| \le \sum_{i=1}^t \|(X-Y)Y^{i-1}A_\gamma K\|.$$

Again, $\|Y^{i-1}A_\gamma K\|_{\ell^4(\nu_k^\gamma)} = O_T(1)_{N\longrightarrow+\infty,\gamma}$ for each i, and the claim follows. □

In sums like (9.1), we can make packets of size $2t$, and we have for all m and for any t,

$$\left|\sum_{r=0}^{t-1}\langle B_\gamma K, \mathcal{S}_{u^\gamma}^{2r+m} B_\gamma K\rangle_\nu - \sum_{r=0}^{t-1}(\overline{s_1 s_2})^r \langle B_\gamma K, e^{i\theta_0}\mathcal{S}_\gamma^{2r} e^{-i\theta_0}\mathcal{S}_{u^\gamma}^m B_\gamma K\rangle_\nu\right| \le t^2\left(c_{M,\beta}\left[\varepsilon^{1/2} + \eta_1 O(1)_{N\longrightarrow+\infty,\gamma}\right] + C'_{N,M}\right)^{1/4} O_T(1)_{N\longrightarrow+\infty,\gamma}. \tag{9.4}$$

As we will see below, the size $2t$ of packets should be chosen so that $t(c_{M,\beta}\varepsilon^{1/2})^{1/4}$ is small as M gets large. Remembering that $c_{M,\beta} \le f(D,\beta)M^3$ and $\varepsilon = M^{-8}$, we take $t = M^\alpha$ with $0 < \alpha < 1/4$. We then group the sum (9.1) into packets and write

$$\left|\sum_{r'\le r\le n}\langle \mathcal{S}_{u^\gamma}^{r-r'} B_\gamma K, B_\gamma K\rangle_\nu\right| = \left|\sum_{r'=1}^n \sum_{r=0}^{n-r'}\langle \mathcal{S}_{u^\gamma}^r B_\gamma K, B_\gamma K\rangle_\nu\right| \le \left|\sum_{r'=1}^n \sum_{a=0}^{\lfloor\frac{n-r'}{2t}\rfloor-2} \sum_{r=2ta}^{2t(a+1)-1}\langle \mathcal{S}_{u^\gamma}^r B_\gamma K, B_\gamma K\rangle_\nu\right| + 4nt\,\|B_\gamma K\|_\nu^2,$$

where we estimated

$$\left|\sum_{r'=1}^n \sum_{r=2t(\lfloor\frac{n-r'}{2t}\rfloor-1)}^{n-r'}\langle \mathcal{S}_{u^\gamma}^r B_\gamma K, B_\gamma K\rangle_\nu\right| \le 4nt\|B_\gamma K\|_\nu^2.$$

Note that

$$\sum_{r=2ta}^{2t(a+1)-1}\langle \mathcal{S}_{u^\gamma}^r \cdot, \cdot\rangle = \sum_{r=0}^{t-1}\langle \mathcal{S}_{u^\gamma}^{2r+2ta}\cdot,\cdot\rangle + \sum_{r=0}^{t-1}\langle \mathcal{S}_{u^\gamma}^{2r+1+2ta}\cdot,\cdot\rangle.$$

So using (9.4),

(9.5)

$$\begin{aligned}
&\Big|\sum_{r'=0}^{n}\sum_{a=0}^{\lfloor\frac{n-r'}{2t}\rfloor-2}\sum_{r=2ta}^{2t(a+1)-1}\langle\mathcal{S}_{u^\gamma}^r B_\gamma K, B_\gamma K\rangle_\nu\Big|\\
&\quad\le\Big|\sum_{r'=0}^{n}\sum_{a=0}^{\lfloor\frac{n-r'}{2t}\rfloor-2}\sum_{r=0}^{t-1}(\overline{s_1s_2})^r\left(\langle B_\gamma K, e^{i\theta_0}\mathcal{S}_\gamma^{2r}e^{-i\theta_0}(\mathcal{S}_{u^\gamma}^{2ta}+\mathcal{S}_{u^\gamma}^{2ta+1})B_\gamma K\rangle_\nu\right)\Big|\\
&\qquad+n\cdot\frac{n}{t}\cdot t^2\left(c_{M,\beta}\left[\varepsilon^{1/2}+\eta_1 O(1)_{N\longrightarrow+\infty,\gamma}\right]+C'_{N,M}\right)^{1/4}O_T(1)_{N\longrightarrow+\infty,\gamma}.
\end{aligned}$$

LEMMA 9.3. *Let* $\|K\|_\infty\le 1$. *For* $A_\gamma K=\mathcal{S}_{u^\gamma}^{2ta}B_\gamma K$ *or* $\mathcal{S}_{u^\gamma}^{2ta+1}B_\gamma K$ *we have for any* L,

$$\begin{aligned}
&\left|\sum_{r=0}^{t-1}(\overline{s_1s_2})^r\langle B_\gamma K, e^{i\theta_0}\mathcal{S}_\gamma^{2r}e^{-i\theta_0}A_\gamma K\rangle_\nu\right|\\
&\quad\le\frac{L^2c_k}{c(D,\beta)}O_T(1)_{N\longrightarrow\infty,\gamma}+t\,O_T(L^{-\infty})_{N\longrightarrow\infty,\gamma}\\
&\qquad+\eta_1 O_{M,T}(1)_{N\longrightarrow+\infty,\gamma}+\frac{1}{|s_1s_2-1|}O_T(1)_{N\longrightarrow\infty,\gamma}.
\end{aligned}$$

Proof. First assume $k=1$. We decompose $e^{-i\theta_0}A_\gamma K=C\mathbf{1}+f$, where $f\perp\mathbf{1}$ in $\ell^2(\nu_1^\gamma)$. So $\mathcal{S}_\gamma^{2r}e^{-i\theta_0}A_\gamma K=C\mathcal{S}_\gamma^{2r}\mathbf{1}+\mathcal{S}_\gamma^{2r}f$.

For the term $\mathcal{S}_\gamma^{2r}f$, we use Proposition 7.5, which yields, for any L,

$$\begin{aligned}
\|\mathcal{S}_\gamma^{2r}f\|_\nu\le&\left(1-L^{-2}c(D,\beta)\right)^{r/2}\|f\|_\nu\\
&+\sum_{l=0}^{r-1}C_{N,L,l,2}(e^{-i\theta_0}A_\gamma)^{1/2}+2\eta_1\sum_{l=1}^{r-1}\|\mathcal{Z}_l f\|_\nu.
\end{aligned}$$

By Corollary 7.8 (recalling that $r\le t\le M^\alpha$), we have

$$\sum_{l=0}^{r-1}C_{N,L,l,2}(e^{-i\theta_0}A_\gamma)^{1/2}=t\,O_T(L^{-\infty})_{N\longrightarrow+\infty,\gamma}.$$

Indeed, the term $e^{-i\theta_0}$ has no impact, as it can be bounded by 1 in the proof of Proposition 7.7. We also have $\|f\|_\nu\le\|A_\gamma K\|_\nu\le\|B_\gamma K\|_\nu=O_T(1)_{N\longrightarrow\infty,\gamma}$, and $\|\mathcal{Z}_l f\|_\nu=O_{l,T}(1)_{N\longrightarrow\infty,\gamma}$ by Remark 7.9. Thus,

$$\begin{aligned}
&\left|\sum_{r=0}^{t-1}(\overline{s_1s_2})^r\langle B_\gamma K, e^{i\theta_0}\mathcal{S}_\gamma^{2r}f\rangle_\nu\right|\\
&\quad\le\frac{2L^2}{c(D,\beta)}O_T(1)_{N\longrightarrow\infty,\gamma}+t\,O_T(L^{-\infty})_{N\longrightarrow\infty,\gamma}+\eta_1 O_{M,T}(1)_{N\longrightarrow\infty,\gamma}.
\end{aligned}$$

For the term $C\mathcal{S}_\gamma^{2r}\mathbf{1}$, we have $\mathcal{S}_\gamma^l\mathbf{1} = \mathbf{1} - \eta_1\sum_{s=0}^{l-1}\mathcal{S}_\gamma^s\iota\xi^\gamma = \mathbf{1} + \eta_1 O_l(1)_{N\longrightarrow\infty,\gamma}$ by (6.11). Thus,

$$\begin{aligned}\Big|\sum_{r=0}^{t-1}(\overline{s_1s_2})^r\langle B_\gamma K, e^{i\theta_0}\mathcal{S}_\gamma^{2r}\mathbf{1}\rangle_\nu\Big| &\le \Big|\sum_{r=0}^{t-1}(\overline{s_1s_2})^r\langle B_\gamma K, e^{i\theta_0}\mathbf{1}\rangle_\nu\Big| + \eta_1 O_M(1)_{N\longrightarrow\infty,\gamma}\|B_\gamma K\|_\nu \\ &= \Big|\frac{(\overline{s_1s_2})^t - 1}{\overline{s_1s_2}-1}\langle B_\gamma K, e^{i\theta_0}\mathbf{1}\rangle_\nu\Big| + \eta_1 O_M(1)_{N\longrightarrow\infty,\gamma}\|B_\gamma K\|_\nu \\ &\le \left(\frac{2}{|s_1s_2-1|} + \eta_1 O_M(1)_{N\longrightarrow\infty,\gamma}\right)\|B_\gamma K\|_\nu.\end{aligned}$$

Since $|C| \le \|A_\gamma K\|_\nu \le \|B_\gamma K\|_\nu$, this completes the proof for $k=1$.

For higher k, as in Remark 9.1, we have

$$\|\mathcal{S}_\gamma^{2r} f\|_{\nu_k} \le \sqrt{\frac{\mu_1^\gamma(B)}{\mu_k^\gamma(B_k)}}\|\mathcal{S}_\gamma^{2r-k+1}\phi_f\|_{\nu_1},$$

where now $\phi_f(x_0,x_1) = (\mathcal{S}_\gamma^{k-1}f)(x_0;x_k)$. We then note that $f\perp\mathbf{1}$ in $\ell^2(\nu_k^\gamma)$ if and only if $\phi_f\perp\mathbf{1}$ in $\ell^2(\nu_1^\gamma)$. Indeed, $\langle\mathbf{1},\phi_f\rangle_{\nu_1} = \frac{\mu_k^\gamma(B_k)}{\mu_1^\gamma(B)}\langle\mathbf{1},f\rangle_{\nu_k}$, since $\langle\mathbf{1},\phi_f\rangle_{\nu_1} = \sum_{(x_0,x_1)}\nu_1(x_0,x_1)(\mathcal{S}_\gamma^{k-1}f)(x_0;x_k)$, so applying (6.9), (6.6) and (2.7), the claim follows. Hence, $\|\mathcal{S}_\gamma^{2r-k+1}\phi_f\|_{\nu_1} \lesssim c(1-L^{-2}C)^{r/2}\|\phi_f\|_{\nu_1}$, where $c = \frac{1}{(1-L^{-2})^{(k+3)/4}} \le 2^{k+1}$ for large L. The error terms are the same, this time with $\|\mathcal{Z}_l\phi_f\|_{\nu_1} = O_{l,T}(1)_{N\longrightarrow\infty,\gamma}$. Finally, $\|\phi_f\|_{\nu_1} \le \sqrt{\frac{\mu_k^\gamma(B_k)}{\mu_1^\gamma(B)}}\|f\|_{\nu_k}$. □

Starting from (9.5) and applying the lemma, we obtain for $\|K\|_\infty \le 1$,

$$\begin{aligned}&\frac{1}{n^2}\Big|\sum_{r'\le n}\sum_{r\ge r'}\langle\mathcal{S}_{u^\gamma}^{r-r'}B_\gamma K, B_\gamma K\rangle_\nu\Big| \\ &\quad\le \frac{1}{t}\Big[\frac{2L^2}{c(D,\beta)}O_T(1)_{N\longrightarrow\infty,\gamma} + t\,O_T(L^{-\infty})_{N\longrightarrow\infty,\gamma} \\ &\qquad + \eta_1 O_{M,T}(1)_{N\longrightarrow+\infty,\gamma} + \frac{1}{|s_1s_2-1|}O_T(1)_{N\longrightarrow+\infty,\gamma}\Big] \\ &\qquad + t\left(c_{M,\beta}\left[\varepsilon^{1/2} + \eta_1 O(1)_{N\longrightarrow+\infty,\gamma}\right] + O_T(M^{-\infty})_{N\longrightarrow\infty,\gamma}\right)^{1/4} \\ &\qquad \cdot O_T(1)_{N\longrightarrow+\infty,\gamma} + 4n^{-1}t\,\|B_\gamma K\|_\nu^2.\end{aligned} \tag{9.6}$$

Remember that $n = M^9$ and $t = M^\alpha$ with $0<\alpha<1/4$. For the term $\frac{1}{t}\frac{2L^2}{c(D,\beta)}$ to be small, we choose $L = M^{\alpha'}$ with $0 < 2\alpha' < \alpha$. For instance, take $\alpha = 3/16$ and $\alpha' = 1/16$. For the other terms, we have $t(c_{M,\beta}\varepsilon^{1/2})^{1/4} = O(M^{\alpha-1/4})$ and

$n^{-1}t = M^{-9+\alpha}$. The terms $\eta_1 O_{M,T}(1)_{N\longrightarrow+\infty,\gamma}$ tend to 0 as $\eta_1 = \eta_0 + \eta \longrightarrow 0$, M and T being fixed. Finally, $\|B_\gamma K\|_\nu^2 = O_T(1)_{N\longrightarrow+\infty,\gamma}$ assuming $\|K\|_\infty \leq 1$.

We can gather the first and second alternative into one statement:

PROPOSITION 9.4. *Let $A > 0$. For all M, for all γ that fall either into the first alternative or into the second one with $|s_1^\gamma(N)s_2^\gamma(N) - 1| \geq A$, we have for $\|K\|_\infty \leq 1$ and for $n = M^9$,*

$$\left\|\frac{1}{n}\sum_{r=1}^{n}\mathcal{R}_{n,r}^\gamma \mathcal{F}_\gamma K\right\|_\gamma^2 \leq \frac{1}{M^{3/16}}\left[\frac{2M^{1/8}}{c(D,\beta)}O_T(1)_{N\longrightarrow\infty,\gamma} + O_T(M^{-\infty})_{N\longrightarrow\infty,\gamma} + \eta_1 O_{M,T}(1)_{N\longrightarrow+\infty,\gamma} + \frac{1}{A}O_T(1)_{N\longrightarrow+\infty,\gamma}\right] + O_T(M^{-1/16})_{N\longrightarrow+\infty,\gamma} + \eta_1^{1/4} O_{M,T}(1)_{N\longrightarrow+\infty,\gamma}.$$

Proof. The arguments in Remark 10.4 readily show that

$$\frac{1}{n^2}\sum_{r,r'=1}^{n}\mathbf{E}_{n,r,r'}(\eta_1, F_\gamma K) = \eta_1\, O_{n,T}(1)_{N\longrightarrow\infty,\gamma}.$$

The assertion follows from (9.1), (9.3) and (9.6). ◻

PROPOSITION 9.5. *Let $I \subset I_1$ with $\bar{I} \subset I_1$. There exists a_0 such that, if $a \leq a_0$, M is large enough, η_1 is small enough ($M \geq M(a)$, $\eta_1 \leq \eta(a)$), and N is large enough. For any γ falling into the second alternative on G_N, the sequence $s^\gamma(N) = s_1^\gamma(N)s_2^\gamma(N)$ must satisfy $|s^\gamma(N) - 1| > a^{13}$, if γ stays in a set of the form*

$$A_{a,\eta_1} = \{\gamma : \operatorname{Re}\gamma \in I, \operatorname{Im}\gamma = \eta_1, \mathbb{P}(|\mathcal{W}(o) - \gamma| < a) \leq 1 - a\}.$$

Before proving the proposition, let us finally give the

Proof of Theorem 3.3. We apply Proposition 5.2 and use Proposition 9.5 to show that we are in the framework of Proposition 9.4.

Two cases may happen.

Case 1: *$\mathcal{W}(o)$ is deterministic*. There exists E_0 such that $\mathbb{P}(\mathcal{W}(o) = E_0) = 1$. In this case, we fix a small $a > 0$, let $J_1 = I \setminus [E_0 - 2a, E_0 + 2a]$ and $J_2 = I \cap [E_0 - 2a, E_0 + 2a]$. We then write $\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K) = \mathrm{Var}^{\mathrm{J}_1}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K) + \mathrm{Var}^{\mathrm{J}_2}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K)$. For $\operatorname{Re}\gamma \in J_1$, we have $|\gamma - E_0| > 2a$, so $\mathbb{P}(|\mathcal{W}(o) - \gamma| < a) = 0$ and Proposition 9.5 applies with a arbitrarily small. Proposition 9.4, applied with $A = a^{13}$, thus allows to control $\mathrm{Var}^{\mathrm{J}_1}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K)$, while $\mathrm{Var}^{\mathrm{J}_2}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K) = O_T(a)$.

Case 2: If $\mathcal{W}(o)$ is not deterministic, there exists a such that for all $E \in \mathbb{R}$, $\mathbb{P}(|\mathcal{W}(o) - E| < a) \leq 1 - a$. Thus, for any complex γ, $\mathbb{P}(|\mathcal{W}(o) - \gamma| < a) \leq 1 - a$. In this case Proposition 9.5 may be applied with the fixed value $A = a^{13}$ and all γ.

Either way, we showed that there exists a_0 such that, for all $a \le a_0$, $M \ge M(a)$, we have for any s and T,

(9.7)
$$\lim_{\eta_0\downarrow 0}\limsup_{N\to\infty} \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K)^2 \le |I|^2 \frac{1}{M^{3/16}}\left[\frac{2M^{1/8}}{c(D,\beta)}C_T + C_{s,T}M^{-s} + \frac{C_T}{a^{13}}\right] + |I|^2 C_T M^{-1/16} + aC_T.$$

Taking $M \to \infty$ followed by $a \downarrow 0$, this completes the proof of Theorem 3.3. $\square$

We conclude the section with the following:

Proof of Proposition 9.5. We use the following consequences of (Green):

- There exists $0 < c_0 < \infty$ such that for all $\gamma \in \mathbb{C}^+$, $\mathrm{Re}\,\gamma \in I_1$,

(9.8)
$$\mathbb{E}\left(\sum_{y\sim o}\hat\mu_1^\gamma(o,y)\right) \le c_0, \quad \mathbb{E}\left(\sum_{y\sim o}(\hat\mu_1^\gamma(o,y))^{-1}\right) \le c_0, \quad \mathbb{E}\left(\sum_{y\sim o}|\hat\zeta_y^\gamma(o)|^{-2}\right) \le c_0.$$

  In fact, $\hat\mu_1^\gamma(o,y) = \frac{|\mathrm{Im}\,\hat\zeta_y^\gamma(o)\,\mathrm{Im}\,\hat\zeta_o^\gamma(y)|}{|\hat m_y^\gamma \hat\zeta_o^\gamma(y)|^2}$, so this follows from (Green) and its consequences (A.9) and (A.10).
- There exists $0 < c_1 < \infty$, such that for all $\gamma \in \mathbb{C}^+$, $\mathrm{Re}\,\gamma \in I_1$,

(9.9a)
$$\mathbb{P}\left(|2\,\mathrm{Im}\,\hat m_o^\gamma| \ge 2r \text{ and } |2\hat m_o^\gamma| \le \frac{1}{2}r^{-1}\right) \ge 1 - c_1 r,$$

(9.9b)
$$\mathbb{P}\left(\sum_{y\sim o}|\hat\zeta_o^\gamma(y)| \le \frac{1}{2}r^{-1}\right) \ge 1 - c_1 r.$$

  In fact, $\mathbb{E}(|2\,\mathrm{Im}\,\hat m_o|^{-1}) + \mathbb{E}(|2\hat m_o^\gamma|) \le c_1/2$ by (A.9), so the first claim follows by Markov's inequality. The second one follows similarly from (A.10).

We may now begin the proof. If γ falls into the second alternative, then

(9.10)
$$\|u_{x_0}^\gamma(x_1)u_{x_1}^\gamma(x_0) - s^\gamma(N)\|_\nu^2 \le 4f(\beta,D)M^3\left[M^{-4} + \eta_1 O(1)_{N\longrightarrow+\infty,\gamma}\right] + 4C'_{N,M}.$$

Let $a_0 = (2c_0)^{-2}(6+3c_1)^{-12}$; this choice will become clear later on. Take $a \le a_0$. There exist $M(a)$, $\eta(a)$ and $N(a)$ such that if $M \ge M(a)$, $\eta_1 \le \eta(a)$ and $N \ge N(a)$, then the right-hand side side in (9.10) is $\le a^{26}$. We fix $\rho \ge a^{26}$.

So take any $a \le a_0$, $M \ge M(a)$, $\eta_1 \le \eta(a)$, and assume towards a contradiction that we can find a subsequence $N_k = N_k(\eta_1) \longrightarrow +\infty$ and a sequence $\gamma_k \in A_{a,\eta_1}$, each falling into the second alternative on G_{N_k}, such that $|s^{\gamma_k}(N_k) - 1|^2 \le \rho$. After extracting further subsequences, let $\lim_{N_k\to+\infty} s^{\gamma_k}(N_k) = s$ and $\gamma_0 = \lim_{N_k\to+\infty}\gamma_k \in \mathbb{C}$. Then $|s-1|^2 \le \rho$, $\mathrm{Re}\,\gamma_0 \in I_1, \mathrm{Im}\,\gamma_0 = \eta_1$, and by

(9.10) and Remark A.3,

$$\mathbb{E}\left(\sum_{y\sim o}|\hat{u}_o^{\gamma_0}(y)\hat{u}_y^{\gamma_0}(o)-s|^2\hat{\mu}_1^{\gamma_0}(o,y)\right)\leq\rho\,\mathbb{E}\left(\sum_{y\sim o}\hat{\mu}_1^{\gamma_0}(o,y)\right),$$

which implies, using (9.8),

$$\mathbb{E}\left(\sum_{y\sim o}|\hat{u}_o^{\gamma_0}(y)\hat{u}_y^{\gamma_0}(o)-1|^2\hat{\mu}_1^{\gamma_0}(o,y)\right)\leq 4\rho\,\mathbb{E}\left(\sum_{y\sim o}\hat{\mu}_1^{\gamma_0}(o,y)\right)\leq 4c_0\rho.$$

By the Cauchy-Schwarz inequality,

$$\mathbb{E}\left(\sum_{y\sim o}|\hat{u}_o^{\gamma_0}(y)\hat{u}_y^{\gamma_0}(o)-1|^2\hat{\mu}_1^{\gamma_0}(o,y)\right)^{1/2}\geq\frac{\mathbb{E}\left(\sum_{y\sim o}|\hat{u}_o^{\gamma_0}(y)\hat{u}_y^{\gamma_0}(o)-1|\right)}{\mathbb{E}\left(\sum_{y\sim o}(\hat{\mu}_1^{\gamma_0}(o,y))^{-1}\right)^{1/2}}$$

and thus, by (9.8),

$$\mathbb{E}\left(\sum_{y\sim o}|\hat{u}_o^{\gamma_0}(y)\hat{u}_y^{\gamma_0}(o)-1|\right)\leq\left(4c_0\rho\,\mathbb{E}\left(\sum_{y\sim o}(\hat{\mu}_1^{\gamma_0}(o,y))^{-1}\right)\right)^{1/2}\leq 2c_0\rho^{1/2}. \tag{9.11}$$

Since the value of γ_0 is now fixed, let us omit it from the notation.

Let us write $\hat{\zeta}_o^{\gamma_0}(y)=\hat{\zeta}_o(y)=r(o,y)e^{-i\theta(o,y)}$ with $r\in\mathbb{R}_+$ and $\theta\in\mathbb{R}$. This implies $\hat{u}_o(y)=e^{2i\theta(o,y)}$ and

$$|\hat{u}_o(y)\hat{u}_y(o)-1|=|(e^{i\theta(y,o)}+e^{-i\theta(o,y)})(e^{i\theta(y,o)}-e^{-i\theta(o,y)})|.$$

Now (9.11) implies that

$$\mathbb{E}\left(\sum_{y\sim o}\min_{\epsilon=\pm1}|e^{i\theta(y,o)}-\epsilon e^{-i\theta(o,y)}|^2\right)\leq 2c_0\rho^{1/2}. \tag{9.12}$$

Let us call $\epsilon(o,y)$ the value of ϵ achieving the min. By (2.7) we have

$$2\hat{m}_o=\hat{\zeta}_y(o)^{-1}-\hat{\zeta}_o(y)=r(y,o)^{-1}e^{i\theta(y,o)}-r(o,y)e^{-i\theta(o,y)}$$

for all $y\sim o$. Thus, using (9.8),

$$\begin{aligned}&\mathbb{E}\left(\sum_{y\sim o}\left|e^{-i\theta(o,y)}\left(\epsilon(o,y)r(y,o)^{-1}-r(o,y)\right)-2\hat{m}_o\right|\right)\\&\quad=\mathbb{E}\left(\sum_{y\sim o}\left|\left(e^{i\theta(y,o)}-\epsilon(o,y)e^{-i\theta(o,y)}\right)r(y,o)^{-1}\right|\right)\\&\quad\leq\sqrt{2c_0}\rho^{1/4}\mathbb{E}\left(\sum_{y\sim o}r(y,o)^{-2}\right)^{1/2}\leq 2c_0\rho^{1/4}=:r^6.\end{aligned} \tag{9.13}$$

Denote $t_{o,y}=\epsilon(o,y)r(y,o)^{-1}-r(o,y)\in\mathbb{R}$. It follows by Markov's inequality that

$$\sum_{y\sim o}\left|t_{o,y}e^{-i\theta(o,y)}-2\hat{m}_o\right|\leq r^5 \tag{9.14}$$

with probability $\geq 1-r$.

The probability that $|2\operatorname{Im}\hat{m}_o| \geq 2r$ and $|2\hat{m}_o| \leq \frac{1}{2}r^{-1}$ is at least $1-c_1 r$ by (9.9a). Thus, (9.14) implies that with probability $\geq 1-r-c_1 r$, we have for any $y \sim o$,

$$r \leq |t_{o,y}| \leq r^{-1}. \tag{9.15}$$

Combining (9.14) and (9.15), we see that for any $y, y' \sim o$,

$$\left|e^{-i\theta(o,y)} - t_{o,y'}t_{o,y}^{-1}e^{-i\theta(o,y')}\right| \leq r^4. \tag{9.16}$$

Now (2.4a) says that

$$\gamma_0 = \mathcal{W}(o) + \sum_{y\sim o}\zeta_o(y) + 2\hat{m}_o = \mathcal{W}(o) + \sum_{y\sim o} r(o,y)e^{-i\theta(o,y)} + 2\hat{m}_o.$$

Using (9.14) and (9.16), we get for any fixed $y' \sim o$,

$$\begin{aligned}
&\left|\gamma_0 - \mathcal{W}(o) - \left(t_{o,y'} + \sum_{y\sim o} r(o,y)t_{o,y'}t_{o,y}^{-1}\right)e^{-i\theta(o,y')}\right| \\
&\leq \left|2\hat{m}_o - t_{o,y'}e^{-i\theta(o,y')}\right| + \left|\sum_{y\sim o} r(o,y)\left(e^{-i\theta(o,y)} - t_{o,y'}t_{o,y}^{-1}e^{-i\theta(o,y')}\right)\right| \\
&\leq r^5 + r^4\sum_{y\sim o} r(o,y) \leq 2r^3
\end{aligned} \tag{9.17}$$

with probability $\geq 1-r-2c_1 r$. Here we used that $\sum_{y\sim o} r(o,y) \leq \frac{1}{2}r^{-1}$ with probability $\geq 1-c_1 r$; see (9.9b). Since $|\gamma_0 - \mathcal{W}(o)| \geq a$ with probability $\geq a$, it follows that

$$\left|t_{o,y'} + \sum_{y\sim o} r(o,y)t_{o,y'}t_{o,y}^{-1}\right| \geq a - 2r^3$$

with probability $\geq 1-r-2c_1 r-(1-a)$. Recall that $r(o,y)$ and $t_{o,u}$ are real. Taking the imaginary part in (9.17), we thus get $|\operatorname{Im} e^{-i\theta(o,y')}| \leq \frac{2r^3+\eta_1}{a-2r^3}$. Assume $\eta_1 \leq r^3$. Then if $r < a/5$, we get $|\operatorname{Im} e^{-i\theta(o,y')}| < r^2$. Hence, $\mathbb{P}(|\operatorname{Im} e^{-i\theta(o,y')}| \geq r^2) \leq (2c_1+1)r + 1 - a$. But we know that $|2\operatorname{Im}\hat{m}_o| \geq 2r$, so taking the imaginary part in (9.14) and using (9.15), we also have that $|\operatorname{Im} e^{-i\theta(o,y')}| \geq r^2$ with probability $\geq 1-r-c_1 r$. If $(2+3c_1)r < a$, this will give a contradiction.

To prove the proposition, we take $r = \frac{a}{6+3c_1}$ and choose

$$a_0 \leq (2c_0)^{-2}(6+3c_1)^{-12}.$$

Recalling that $2c_0\rho^{1/4} = r^6$, we get $\rho^{1/2} = (2c_0)^{-2}(\frac{a}{6+3c_1})^{12} \geq a^{13}$ for $a \leq a_0$, as required. We also take $M > M(a)$ and $\eta_1 \leq \min(r^3, \eta(a))$. □

10. Step 5: Back to the original eigenfunctions

In this section, we show that it suffices to consider the non-backtracking quantum variance in order to prove quantum ergodicity; in other words, Theorem 1.3 can be retrieved from Theorem 3.3. This part may be read before or after the others.

Given $K \in \mathscr{H}_k$, we define the *quantum variance* by

$$\mathrm{Var}^{\mathrm{I}}(K) = \frac{1}{N} \sum_{\lambda_j \in I} |\langle \psi_j, K_G \psi_j \rangle| \,, \tag{10.1}$$

where K_G is as in Section 2.1.

More generally, fix $\eta_0 > 0$ and suppose $K^\gamma \in \mathscr{H}_k$ satisfies conditions (Hol). We denote

$$\mathrm{Var}^{\mathrm{I}}_{\eta_0}(K^\gamma) = \frac{1}{N} \sum_{\lambda_j \in I} \left| \left\langle \psi_j, K_G^{\lambda_j + i\eta_0} \psi_j \right\rangle \right| ,$$

where the subscript η_0 indicates that inside the variance, $\operatorname{Im} \gamma$ is fixed and equal to η_0. Denote $\gamma_j = \lambda_j + i\eta_0$, and define

$$g_j(x_0, x_1) = \overline{\zeta_{x_0}^{\gamma_j}(x_1)}^{-1} \psi_j(x_1) - \psi_j(x_0)$$

and

$$g_j^*(x_0, x_1) = \overline{\zeta_{x_1}^{\gamma_j}(x_0)}^{-1} \psi_j(x_0) - \psi_j(x_1),$$

so g_j^* and g_j are defined like f_j^* and f_j (Section 3), respectively, with ζ replaced by $\overline{\zeta}$. Put

$$\widetilde{\mathrm{Var}}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K^\gamma) = \frac{1}{N} \sum_{\lambda_j \in I} \left| \left\langle g_j^*, K_B^{\gamma_j} g_j \right\rangle \right| .$$

Next, given $\gamma \in \mathbb{C}^+$, define the function $N_\gamma : V \longrightarrow \mathbb{R}_+$ by

$$N_\gamma(x) = \operatorname{Im} \tilde{g}^\gamma(\tilde{x}, \tilde{x}), \tag{10.2}$$

where $\tilde{x}$ is a point in $\widetilde{G}$ projecting down to $G = \Gamma \backslash \widetilde{G}$. Recall the Laplacian P defined in (1.1). We next introduce the operators $P_\gamma, \mathcal{S}_{T,\gamma}, \widetilde{\mathcal{S}}_{T,\gamma} : \mathbb{C}^V \to \mathbb{C}^V$ defined by

(10.3)
$$P_\gamma = \frac{d}{N_\gamma} P \frac{N_\gamma}{d}, \qquad \mathcal{S}_{T,\gamma} = \frac{1}{T} \sum_{s=0}^{T-1} (T-s) P_\gamma^s \qquad \text{and} \qquad \widetilde{\mathcal{S}}_{T,\gamma} = \frac{1}{T} \sum_{s=1}^{T} P_\gamma^s$$

for $T \in \mathbb{N}^*$, and the operators $\mathcal{L}^\gamma, \widetilde{\mathcal{L}}^\gamma : \mathbb{C}^V \to \mathbb{C}^B$ defined by

$$(\mathcal{L}^\gamma J)(x_0, x_1) = \frac{|\zeta_{x_0}^\gamma(x_1)|^2}{|2m_{x_0}^\gamma|^2} \left(\frac{J(x_0)}{N_\gamma(x_1)} - \frac{J(x_1)}{\overline{\zeta_{x_0}^\gamma(x_1)} \zeta_{x_1}^\gamma(x_0) N_\gamma(x_0)} \right),$$

$$(\widetilde{\mathcal{L}}^\gamma J)(x_0, x_1) = \frac{|\zeta_{x_0}^\gamma(x_1)|^2}{|2m_{x_0}^\gamma|^2} \left(\frac{J(x_0)}{N_\gamma(x_1)} - \frac{J(x_1)}{\zeta_{x_0}^\gamma(x_1) \overline{\zeta_{x_1}^\gamma(x_0)} N_\gamma(x_0)} \right).$$

Finally, denote $\mathrm{Var}^{\mathrm{I}}_{\eta_0}(K-\langle K\rangle_\gamma) := \mathrm{Var}^{\mathrm{I}}_{\eta_0}(K-\langle K\rangle_\gamma \mathbf{1})$, where $\mathbf{1}\in\mathscr{H}_0$ is the constant function equal to 1 (so that, with the notation of Section 2.1, $\widehat{\mathbf{1}}$ is the identity operator).

Proposition 10.1. *Fix $\eta_0>0$ and $T\in\mathbb{N}^*$. For any $J\in\mathscr{H}_0$, we have*

$$\begin{aligned}\mathrm{Var}^{\mathrm{I}}_{\eta_0}\left(J-\langle J\rangle_\gamma\right) \le{}& \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}\left(\mathcal{L}^\gamma d^{-1}\mathcal{S}_{T,\gamma}\left(J-\langle J\rangle_\gamma\right)\right)\\ &+\widetilde{\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}}\left(\widetilde{\mathcal{L}}^\gamma d^{-1}\mathcal{S}_{T,\gamma}\left(J-\langle J\rangle_\gamma\right)\right)\\ &+\mathrm{Var}^{\mathrm{I}}_{\eta_0}\left(\widetilde{\mathcal{S}}_{T,\gamma}\left(J-\langle J\rangle_\gamma\right)\right).\end{aligned}$$

Proof. We have

(10.4)
$$\begin{aligned}\langle f_j^*,(\mathcal{L}^{\gamma_j}J)_B f_j\rangle ={}& \sum_{(x_0,x_1)\in B}\left(\frac{(\mathcal{L}^{\gamma_j}J)(x_0,x_1)}{\overline{\zeta^{\gamma_j}_{x_1}(x_0)}\zeta^{\gamma_j}_{x_0}(x_1)}+(\mathcal{L}^{\gamma_j}J)(x_1,x_0)\right)\overline{\psi_j(x_0)}\psi_j(x_1)\\ &-\sum_{(x_0,x_1)\in B}(\mathcal{L}^{\gamma_j}J)(x_0,x_1)\left(\frac{|\psi_j(x_0)|^2}{\overline{\zeta^{\gamma_j}_{x_1}(x_0)}}+\frac{|\psi_j(x_1)|^2}{\zeta^{\gamma_j}_{x_0}(x_1)}\right).\end{aligned}$$

We calculate $\langle g_j^*,(\widetilde{\mathcal{L}}^{\gamma_j}J)_B g_j\rangle$ similarly. We then note that

$$\frac{(\mathcal{L}^{\gamma_j}J)(x_0,x_1)}{\overline{\zeta^{\gamma_j}_{x_1}(x_0)}\zeta^{\gamma_j}_{x_0}(x_1)}+(\mathcal{L}^{\gamma_j}J)(x_1,x_0)-\frac{(\widetilde{\mathcal{L}}^{\gamma_j}J)(x_0,x_1)}{\zeta^{\gamma_j}_{x_1}(x_0)\overline{\zeta^{\gamma_j}_{x_0}(x_1)}}-(\widetilde{\mathcal{L}}^{\gamma_j}J)(x_1,x_0)=0,$$

using that $\frac{|\zeta^{\gamma}_{x_1}(x_0)|^2}{|m^{\gamma}_{x_1}|^2}=\frac{|\zeta^{\gamma}_{x_0}(x_1)|^2}{|m^{\gamma}_{x_0}|^2}$, by (2.7). Hence,

$$\begin{aligned}&\langle f_j^*,(\mathcal{L}^{\gamma_j}J)_B f_j\rangle-\langle g_j^*,(\widetilde{\mathcal{L}}^{\gamma_j}J)_B g_j\rangle\\ &\quad=\sum_{(x_0,x_1)\in B}(\widetilde{\mathcal{L}}^{\gamma_j}J)(x_0,x_1)\left(\frac{|\psi_j(x_0)|^2}{\zeta^{\gamma_j}_{x_1}(x_0)}+\frac{|\psi_j(x_1)|^2}{\overline{\zeta^{\gamma_j}_{x_0}(x_1)}}\right)\\ &\qquad-\sum_{(x_0,x_1)\in B}(\mathcal{L}^{\gamma_j}J)(x_0,x_1)\left(\frac{|\psi_j(x_0)|^2}{\overline{\zeta^{\gamma_j}_{x_1}(x_0)}}+\frac{|\psi_j(x_1)|^2}{\zeta^{\gamma_j}_{x_0}(x_1)}\right).\end{aligned}$$

Let $\alpha^{x_1}_{x_0}=\frac{|\zeta^{\gamma}_{x_0}(x_1)|^2}{|2m^{\gamma}_{x_0}|^2 N_\gamma(x_1)}$, and note that $\alpha^{x_0}_{x_1}=\frac{|\zeta^{\gamma}_{x_0}(x_1)|^2}{|2m^{\gamma}_{x_0}|^2 N_\gamma(x_0)}$ by (2.7). Then

$$\begin{aligned}&\frac{(\widetilde{\mathcal{L}}^{\gamma_j}J)(x_0,x_1)}{\zeta^{\gamma_j}_{x_1}(x_0)}-\frac{(\mathcal{L}^{\gamma_j}J)(x_0,x_1)}{\overline{\zeta^{\gamma_j}_{x_1}(x_0)}}\\ &\qquad=-2i\left[\frac{\operatorname{Im}\zeta^{\gamma_j}_{x_1}(x_0)}{|\zeta^{\gamma_j}_{x_1}(x_0)|^2}\alpha^{x_1}_{x_0}J(x_0)-\frac{\operatorname{Im}\zeta^{\gamma_j}_{x_0}(x_1)}{|\zeta^{\gamma_j}_{x_1}(x_0)\zeta^{\gamma_j}_{x_0}(x_1)|^2}\alpha^{x_0}_{x_1}J(x_1)\right]\end{aligned}$$

and

$$\frac{(\widetilde{\mathcal{L}}^{\gamma_j}J)(x_0,x_1)}{\overline{\zeta_{x_0}^{\gamma_j}(x_1)}} - \frac{(\mathcal{L}^{\gamma_j}J)(x_0,x_1)}{\zeta_{x_0}^{\gamma_j}(x_1)} = 2i\left[\frac{\operatorname{Im}\zeta_{x_0}^{\gamma_j}(x_1)}{|\zeta_{x_0}^{\gamma_j}(x_1)|^2}\alpha_{x_0}^{x_1}J(x_0) - \frac{\operatorname{Im}\zeta_{x_1}^{\gamma_j}(x_0)}{|\zeta_{x_1}^{\gamma_j}(x_0)\zeta_{x_0}^{\gamma_j}(x_1)|^2}\alpha_{x_1}^{x_0}J(x_1)\right].$$

Hence,

$$\begin{aligned}\langle f_j^*, (\mathcal{L}^{\gamma_j}J)_B f_j\rangle - \langle g_j^*, (\widetilde{\mathcal{L}}^{\gamma_j}J)_B g_j\rangle &= -2i\sum_{x_0\in V}|\psi_j(x_0)|^2 J(x_0)\sum_{x_1\sim x_0}\left(\frac{\operatorname{Im}\zeta_{x_1}^{\gamma_j}(x_0)}{|\zeta_{x_1}^{\gamma_j}(x_0)|^2}\alpha_{x_0}^{x_1} + \frac{\operatorname{Im}\zeta_{x_0}^{\gamma_j}(x_1)}{|\zeta_{x_0}^{\gamma_j}(x_1)\zeta_{x_1}^{\gamma_j}(x_0)|^2}\alpha_{x_0}^{x_1}\right)\\ &\quad + 2i\sum_{x_0\in V}|\psi_j(x_0)|^2\sum_{x_1\sim x_0}\left(\frac{\operatorname{Im}\zeta_{x_0}^{\gamma_j}(x_1)}{|\zeta_{x_1}^{\gamma_j}(x_0)\zeta_{x_0}^{\gamma_j}(x_1)|^2}\alpha_{x_1}^{x_0} + \frac{\operatorname{Im}\zeta_{x_1}^{\gamma_j}(x_0)}{|\zeta_{x_1}^{\gamma_j}(x_0)|^2}\alpha_{x_1}^{x_0}\right)J(x_1).\end{aligned}$$

Now, by (2.7),

$$\begin{aligned}\operatorname{Im}\zeta_{x_0}^{\gamma}(x_1) + \operatorname{Im}\zeta_{x_1}^{\gamma}(x_0)\cdot|\zeta_{x_0}^{\gamma}(x_1)|^2 &= |\zeta_{x_0}^{\gamma}(x_1)|^2\Big[\frac{\operatorname{Im}\zeta_{x_0}^{\gamma}(x_1)}{|\zeta_{x_0}^{\gamma}(x_1)|^2} + \operatorname{Im}\zeta_{x_1}^{\gamma}(x_0)\Big]\\ &= -2\operatorname{Im} m_{x_1}^{\gamma}\cdot|\zeta_{x_0}^{\gamma}(x_1)|^2.\end{aligned}$$

Since $2\operatorname{Im} m_{x_1}^{\gamma} = N_\gamma(x_1)|2m_{x_1}^{\gamma}|^2$, we get

$$\frac{\operatorname{Im}\zeta_{x_0}^{\gamma}(x_1) + \operatorname{Im}\zeta_{x_1}^{\gamma}(x_0)|\zeta_{x_0}^{\gamma}(x_1)|^2}{|\zeta_{x_0}^{\gamma}(x_1)\zeta_{x_1}^{\gamma}(x_0)|^2} = \frac{-N_\gamma(x_1)|2m_{x_1}^{\gamma}|^2}{|\zeta_{x_1}^{\gamma_j}(x_0)|^2}.$$

Since $\alpha_{x_0}^{x_1} = \frac{|\zeta_{x_1}^{\gamma}(x_0)|^2}{N_\gamma(x_1)|2m_{x_1}^{\gamma}|^2}$ and $\alpha_{x_1}^{x_0} = \frac{|\zeta_{x_1}^{\gamma}(x_0)|^2}{N_\gamma(x_0)|2m_{x_1}^{\gamma}|^2}$ by (2.7), we thus have

$$\begin{aligned}\langle f_j^*, (\mathcal{L}^{\gamma_j}J)_B f_j\rangle - \langle g_j^*, (\widetilde{\mathcal{L}}^{\gamma_j}J)_B g_j\rangle &= 2i\sum_{x_0\in V}|\psi_j(x_0)|^2 d(x_0)J(x_0)\\ &\quad - 2i\sum_{x_0\in V}|\psi_j(x_0)|^2\frac{1}{N_\gamma(x_0)}\sum_{x_1\sim x_0}N_\gamma(x_1)J(x_1) = 2i\,\langle\psi_j, [(I-P_{\gamma_j})dJ]_G\psi_j\rangle.\end{aligned}$$

Hence,

$$\operatorname{Var}_{\eta_0}^{\mathrm{I}}[(I-P_\gamma)J] \le \operatorname{Var}_{\mathrm{nb},\eta_0}^{\mathrm{I}}(\mathcal{L}^{\gamma}d^{-1}J) + \widetilde{\operatorname{Var}_{\mathrm{nb}}^{\mathrm{I}}}(\widetilde{\mathcal{L}}^{\gamma}d^{-1}J).$$

Now note that $P_\gamma(\mathcal{S}_{T,\gamma}K) = \frac{1}{T}\sum_{s=1}^{T}(T-s+1)P_\gamma^s K = \mathcal{S}_{T,\gamma}K - K + \widetilde{\mathcal{S}}_{T,\gamma}K$. Hence,

$$K = (I-P_\gamma)\mathcal{S}_{T,\gamma}K + \widetilde{\mathcal{S}}_{T,\gamma}K$$

for any $K \in \mathscr{H}_0$. Taking $K_\gamma = J - \langle J\rangle_\gamma$, we thus get

$$\begin{aligned}\mathrm{Var}^{\mathrm{I}}_{\eta_0}(K_\gamma) &\le \mathrm{Var}^{\mathrm{I}}_{\eta_0}[(I-P_\gamma)\mathcal{S}_{T,\gamma}K_\gamma] + \mathrm{Var}^{\mathrm{I}}_{\eta_0}(\widetilde{\mathcal{S}}_{T,\gamma}K_\gamma)\\ &\le \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{L}^\gamma d^{-1}\mathcal{S}_{T,\gamma}K_\gamma)\\ &\quad + \widetilde{\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb}}}(\widetilde{\mathcal{L}}^\gamma d^{-1}\mathcal{S}_{T,\gamma}K_\gamma) + \mathrm{Var}^{\mathrm{I}}_{\eta_0}(\widetilde{\mathcal{S}}_{T,\gamma}K_\gamma). \qquad \square\end{aligned}$$

We now consider $K\in\mathscr{H}_m$ for $m>0$. Define $\mathcal{T}^\gamma : \mathscr{H}_1 \to \mathscr{H}_1$ and $\mathcal{O}_1^\gamma : \mathscr{H}_1 \to \mathscr{H}_0$ by

$$(\mathcal{T}^\gamma K)(x_0,x_1) = \frac{\overline{\zeta^\gamma_{x_1}(x_0)}\zeta^\gamma_{x_0}(x_1)}{\overline{\zeta^\gamma_{x_1}(x_0)}\zeta^\gamma_{x_0}(x_1)+1}K(x_0,x_1), \tag{10.5}$$

$$(\mathcal{O}_1^\gamma K)(x_0) = \sum_{x_{-1}\sim x_0}\frac{(\mathcal{T}^\gamma K)(x_{-1},x_0)}{\zeta^\gamma_{x_{-1}}(x_0)} + \sum_{x_1\sim x_0}\frac{(\mathcal{T}^\gamma K)(x_0,x_1)}{\overline{\zeta^\gamma_{x_1}(x_0)}}. \tag{10.6}$$

For $m\ge 2$, define $\mathcal{U}^\gamma_m : \mathscr{H}_m \to \mathscr{H}_m$, $\mathcal{O}^\gamma_m : \mathscr{H}_m \to \mathscr{H}_{m-1}$ and $\mathcal{P}^\gamma_m : \mathscr{H}_m \to \mathscr{H}_{m-2}$ by

$$(\mathcal{U}^\gamma_m K)(x_0;x_m) = \overline{\zeta^\gamma_{x_1}(x_0)}\zeta^\gamma_{x_{m-1}}(x_m)K(x_0;x_m), \tag{10.7}$$

$$\begin{aligned}(\mathcal{O}^\gamma_m K)(x_0;x_{m-1}) = &\sum_{x_{-1}\in\mathcal{N}_{x_0}\setminus\{x_1\}}\overline{\zeta^\gamma_{x_0}(x_{-1})}K(x_{-1};x_{m-1})\\ &+\sum_{x_m\in\mathcal{N}_{x_{m-1}}\setminus\{x_{m-2}\}}K(x_0;x_m)\zeta^\gamma_{x_{m-1}}(x_m).\end{aligned} \tag{10.8}$$

(10.9)
$$(\mathcal{P}^\gamma_m K)(x_1;x_{m-1}) = \sum_{x_0\in\mathcal{N}_{x_1}\setminus\{x_2\},x_m\in\mathcal{N}_{x_{m-1}}\setminus\{x_{m-2}\}}\overline{\zeta^\gamma_{x_1}(x_0)}K(x_0;x_m)\zeta^\gamma_{x_{m-1}}(x_m).$$

PROPOSITION 10.2. *Fix $\eta_0>0$. Suppose $\overline{\psi_j(x_0)}\psi_j(x_1)\in\mathbb{R}$ for any $j=1,\dots,N$ and $(x_0,x_1)\in B$. Then for any $K\in\mathscr{H}_1$, we have*

$$\mathrm{Var}^{\mathrm{I}}_{\eta_0}(K-\langle K\rangle_\gamma) \le \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{T}^\gamma K) + \mathrm{Var}^{\mathrm{I}}_{\eta_0}(\mathcal{O}^\gamma_1 K - \langle\mathcal{O}^\gamma_1 K\rangle_\gamma),$$

and for any $K\in\mathscr{H}_m$, $m\ge 2$, we have

$$\begin{aligned}\mathrm{Var}^{\mathrm{I}}_{\eta_0}(K-\langle K\rangle_\gamma) \le{}& \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{U}^\gamma_m K)\\ &+ \mathrm{Var}^{\mathrm{I}}_{\eta_0}(\mathcal{O}^\gamma_m K - \langle\mathcal{O}^\gamma_m K\rangle_\gamma) + \mathrm{Var}^{\mathrm{I}}_{\eta_0}(\mathcal{P}^\gamma_m K - \langle\mathcal{P}^\gamma_m K\rangle_\gamma).\end{aligned}$$

Proof. Let $K\in\mathscr{H}_1$. Since $\overline{\psi_j(x_0)}\psi_j(x_1)\in\mathbb{R}$ for all (x_0,x_1), we have

$$\begin{aligned}\langle f_j^*, (\mathcal{T}^{\gamma_j}K)_B f_j\rangle = &\sum_{(x_0,x_1)}\overline{\psi_j(x_0)}\psi_j(x_1)\Big(\frac{1}{\overline{\zeta^{\gamma_j}_{x_1}(x_0)}\zeta^{\gamma_j}_{x_0}(x_1)}+1\Big)\mathcal{T}^{\gamma_j}(x_0,x_1)\\ &-\sum_{(x_0,x_1)}(\mathcal{T}^{\gamma_j}K)(x_0,x_1)\Big(\frac{|\psi_j(x_0)|^2}{\overline{\zeta^{\gamma_j}_{x_1}(x_0)}}+\frac{|\psi_j(x_1)|^2}{\zeta^{\gamma_j}_{x_0}(x_1)}\Big).\end{aligned}$$

By definition of $\mathcal{T}^\gamma$ and $\mathcal{O}_1^\gamma$, this implies

$$\langle f_j^*, (\mathcal{T}^{\gamma_j} K)_B f_j\rangle = \langle \psi_j, K_G \psi_j\rangle - \langle \psi_j, (\mathcal{O}_1^{\gamma_j} K)_G \psi_j\rangle$$

and thus

$$\mathrm{Var}^{\mathrm{I}}_{\eta_0}(K - \langle K\rangle_\gamma) \le \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{T}^\gamma K) + \mathrm{Var}^{\mathrm{I}}_{\eta_0}(\mathcal{O}_1^\gamma K - \langle K\rangle_\gamma).$$

Recall the definition of $\langle K\rangle_\gamma$ in (1.5). We claim that

$$\langle \mathcal{O}_1^\gamma K\rangle_\gamma = \langle K\rangle_\gamma. \tag{10.10}$$

Indeed, we have $\langle K\rangle_\gamma = \sum_{(x_0,x_1)\in B} K(x_0,x_1)\Phi_\gamma(x_0,x_1)$. On the other hand,

$$\langle \mathcal{O}_1^\gamma K\rangle_\gamma = \sum_{(x_0,x_1)\in B} \frac{(\mathcal{T}^\gamma K)(x_0,x_1)\Phi_\gamma(x_1,x_1)}{\zeta^\gamma_{x_0}(x_1)} + \sum_{(x_0,x_1)\in B} \frac{(\mathcal{T}^\gamma K)(x_0,x_1)\Phi_\gamma(x_0,x_0)}{\overline{\zeta^\gamma_{x_1}(x_0)}}.$$

But $\frac{\Phi_\gamma(x_1,x_1)}{\zeta^\gamma_{x_0}(x_1)} + \frac{\Phi_\gamma(x_0,x_0)}{\overline{\zeta^\gamma_{x_1}(x_0)}} = \frac{1+\overline{\zeta^\gamma_{x_1}(x_0)}\zeta^\gamma_{x_0}(x_1)}{\zeta^\gamma_{x_0}(x_1)\overline{\zeta^\gamma_{x_1}(x_0)}}\Phi_\gamma(x_0,x_1)$ by (2.13) and the fact that $\Psi_{\gamma,x}(y) = \Psi_{\gamma,y}(x)$ by (2.8), so that $\Phi_\gamma(x,y) = \Phi_\gamma(y,x)$. Hence,

$$\langle \mathcal{O}_1^\gamma K\rangle_\gamma = \sum_{(x_0,x_1)\in B} \frac{\overline{\zeta^\gamma_{x_1}(x_0)}\zeta^\gamma_{x_0}(x_1)}{\overline{\zeta^\gamma_{x_1}(x_0)}\zeta^\gamma_{x_0}(x_1)+1} K(x_0,x_1) \cdot \frac{1+\overline{\zeta^\gamma_{x_1}(x_0)}\zeta^\gamma_{x_0}(x_1)}{\zeta^\gamma_{x_0}(x_1)\overline{\zeta^\gamma_{x_1}(x_0)}}\Phi_\gamma(x_0,x_1) = \langle K\rangle_\gamma.$$

This proves the proposition for $m=1$. Now let $m\ge 2$. It is easily checked that

$$\langle f_j^*, (\mathcal{U}_m^{\gamma_j} K)_B f_j\rangle = \langle \psi_j, (K - \mathcal{O}_m^{\gamma_j} K + \mathcal{P}_m^{\gamma_j} K)_G \psi_j\rangle,$$

and thus

$$\mathrm{Var}^{\mathrm{I}}_{\eta_0}(K - \langle K\rangle_\gamma) \le \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{U}_m^\gamma K) + \mathrm{Var}^{\mathrm{I}}_{\eta_0}(\mathcal{O}_m^\gamma K - \mathcal{P}_m^\gamma K - \langle K\rangle_\gamma). \tag{10.11}$$

We now note that

$$\langle K\rangle_\gamma = \langle \mathcal{O}_m^\gamma K - \mathcal{P}_m^\gamma K\rangle_\gamma. \tag{10.12}$$

Indeed, we have

$$\begin{aligned}\langle \mathcal{O}_m^\gamma K - \mathcal{P}_m^\gamma K\rangle_\gamma = &\sum_{(x_{-1};x_{m-1})\in B_m} \overline{\zeta^\gamma_{x_0}(x_{-1})} K(x_{-1};x_{m-1})\Phi_\gamma(x_0,x_{m-1}) \\ &+ \sum_{(x_0;x_m)\in B_m} K(x_0;x_m)\zeta^\gamma_{x_{m-1}}(x_m)\Phi_\gamma(x_0,x_{m-1}) \\ &- \sum_{(x_0;x_m)\in B_m} \overline{\zeta^\gamma_{x_1}(x_0)} K(x_0;x_m)\zeta^\gamma_{x_{m-1}}(x_m)\Phi_\gamma(x_1,x_{m-1}),\end{aligned}$$

so (10.12) follows from (2.13). Using (10.11), this completes the proof. □

We introduce one last operator $\mathcal{X}_\gamma : \mathscr{H}_0 \to \mathscr{H}_0$ given by

$$\mathcal{X}_\gamma K = \langle K \rangle_\gamma \mathbf{1}.$$

The following corollary then holds assuming all eigenfunctions ψ_j are real. Note that this assumption is not needed in the special case $m = 0$, corresponding to Theorem 1.1.

COROLLARY 10.3. *Suppose we have shown that*

$$\lim_{\eta_0 \downarrow 0} \limsup_{N\to\infty} \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K) = 0, \quad \lim_{\eta_0 \downarrow 0} \limsup_{N\to\infty} \widetilde{\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}}(\widetilde{\mathcal{F}}_\gamma K) = 0$$

for any $\mathcal{F}_\gamma : \mathscr{H}_m \to \mathscr{H}_k$ that is a polynomial combination of $\mathcal{L}^\gamma d^{-1}\mathcal{S}_{T,\gamma}$, $\mathcal{X}_\gamma$, $\mathcal{U}^\gamma_j$, $\mathcal{T}^\gamma$, $\mathcal{O}^\gamma_j$ and $\mathcal{P}^\gamma_j$ (T fixed), $\widetilde{\mathcal{F}}_\gamma$ being the same combination with $\mathcal{L}^\gamma$ replaced by $\widetilde{\mathcal{L}}^\gamma$. Suppose that

$$\lim_{T\longrightarrow+\infty} \lim_{\eta_0 \downarrow 0} \limsup_{N\to\infty} \mathrm{Var}^{\mathrm{I}}_{\eta_0} \left(\widetilde{\mathcal{S}}_{T,\gamma}(C_\gamma K - \langle C_\gamma K \rangle_\gamma) \right) = 0, \tag{10.13}$$

where $C_\gamma : \mathscr{H}_m \to \mathscr{H}_0$ is any polynomial combination of $\mathcal{U}^\gamma_j$, $\mathcal{T}^\gamma$, $\mathcal{O}^\gamma_j$ and $\mathcal{P}^\gamma_j$. Then it will follow that $\lim_{\eta_0\downarrow 0} \limsup_{N\to\infty} \mathrm{Var}^{\mathrm{I}}_{\eta_0}(K - \langle K \rangle_\gamma) = 0$ for any $K \in \mathscr{H}_m$. In other words, Theorem 1.3 will follow.

Proof. The case $m = 0$ holds by Proposition 10.1 and the triangle inequality $\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K - \langle K\rangle_\gamma) \le \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(K) + \mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{X}_\gamma K)$. Here, $\mathcal{F}_\gamma$ has the form $\mathcal{L}^\gamma d^{-1} S_{T,\gamma}$, $\mathcal{L}^\gamma d^{-1} S_{T,\gamma}\mathcal{X}_\gamma$ and $C_\gamma = I$.

The result for higher m follows by induction using Proposition 10.2. For example, for $m = 2$, the conclusion is obtained by taking $\mathcal{F}_\gamma$ of the form $\mathcal{U}^\gamma_2$, $\mathcal{T}^\gamma\mathcal{O}^\gamma_2$, $\mathcal{L}^\gamma d^{-1}\mathcal{S}_{T,\gamma}\mathcal{O}^\gamma_1\mathcal{O}^\gamma_2$, $\mathcal{L}^\gamma d^{-1}\mathcal{S}_{T,\gamma}\mathcal{X}_\gamma\mathcal{O}^\gamma_1\mathcal{O}^\gamma_2$, $\mathcal{L}^\gamma d^{-1}\mathcal{S}_{T,\gamma}\mathcal{P}^\gamma_2$, $\mathcal{L}^\gamma d^{-1}\mathcal{S}_{T,\gamma}\mathcal{X}_\gamma\mathcal{P}^\gamma_2$, and C_γ of the form $\mathcal{O}^\gamma_1\mathcal{O}^\gamma_2$ and $\mathcal{P}^\gamma_2$. □

Remark 10.4. All the operators in Corollary 10.3 satisfy the assumptions (Hol) from Definition 3.2. Indeed, the first two points of (Hol) are clear. (The derivative of any Green function such as ζ^z or G^z may be assessed, for example, using the resolvent equation, yielding $|\partial_z \zeta^z| \le (\mathrm{Im}\, z)^{-2}$.)

For the third point, we should estimate $\frac{1}{N}\sum_{\omega\in B_k} |\mathcal{F}_\gamma K(\omega)|^s$. Assume first that $\mathcal{X}_\gamma$ is not contained in $\mathcal{F}_\gamma$. Then assuming $\|K\|_\infty \le 1$, we write

$$|\mathcal{F}_\gamma K(\omega)| = \left| \sum_{\omega'\in B_m} \mathcal{F}_\gamma(\omega,\omega')K(\omega') \right| \le \sum_{\omega'\in B_m} |\mathcal{F}_\gamma(\omega,\omega')|.$$

Now $\mathcal{F}_\gamma = A^{(1)}\cdots A^{(\ell)}$ is a composition of operators $A^{(r)}$, each of which is either a multiplication or of nearest-neighbor type (with $\mathcal{S}_{T,\gamma}$ a composition of Laplacians). So the sum $\sum_{\omega'} A^{(r)}(\omega,\omega')$ reduces to $\sum_{\omega'\approx\omega} A^{(r)}(\omega,\omega')$, where depending on the operator, $\omega'\approx\omega$ means $\omega'=\omega$, $\omega'\sim\omega$, $\omega'\in\{o_\omega, t_\omega\}$ (origin and terminus of ω), $\omega' \in \{(x,\omega),(\omega,y) : x \sim o_\omega, y \sim t_\omega\}$ or $\omega' \in \{(x,\omega,y) : x \sim o_\omega, y \sim t_\omega\}$. In any case, $\#\{\omega' \approx \omega\} \le 2D$. So

$$\mathcal{F}_\gamma(\omega,\omega') = \sum_{\omega_1 \approx \omega} \cdots \sum_{\omega_{\ell-1} \approx \omega_{\ell-2}} A^{(1)}(\omega,\omega_1) \dots A^{(\ell)}(\omega_{\ell-1},\omega')$$

and thus

$$\sum_{\omega' \in B_m} |\mathcal{F}_\gamma(\omega,\omega')| \le \sum_{\omega_1 \approx \omega} \cdots \sum_{\omega_\ell \approx \omega_{\ell-1}} |A^{(1)}(\omega,\omega_1) \dots A^{(\ell)}(\omega_{\ell-1},\omega_\ell)|.$$

It follows that

$$|\mathcal{F}_\gamma K(\omega)|^s \le (2\ell D)^{s-1} \sum_{\omega_1 \approx \omega} \cdots \sum_{\omega_\ell \approx \omega_{\ell-1}} |A^{(1)}(\omega,\omega_1) \dots A^{(\ell)}(\omega_{\ell-1},\omega_\ell)|^s.$$

Using Hölder's inequality, if $\sum_{r=1}^{\ell} \frac{1}{p_r} = 1$, using Remark A.3 we get that

(10.14)

$$\frac{1}{N} \sum_{\omega \in B_k} |\mathcal{F}_\gamma K(\omega)|^s \le C_{D,\ell,k,s} \prod_{r=1}^{\ell} \left(\frac{1}{N} \sum_{\omega_{r-1}} \sum_{\omega_r \approx \omega_{r-1}} |A^{(r)}(\omega_{r-1},\omega_r)|^{sp_r} \right)^{1/p_r}$$
$$\underset{N \longrightarrow +\infty}{\longrightarrow} C_{D,\ell,k,s} \prod_{r=1}^{\ell} \mathbb{E} \left[\sum_{\omega_{r-1} : o_{\omega_{r-1}} = o} \sum_{\omega_r \approx \omega_{r-1}} |\hat{A}^{(r)}(\omega_{r-1},\omega_r)|^{sp_r} \right]^{1/p_r}$$

uniformly in λ. Here, ℓ may depend on T. By definition, all $\hat{A}^{(r)}(\omega,\omega')$ are well-behaved functions of $\hat{\zeta}$ and $\mathcal{G}^z$, so the previous expression is finite using Remark A.4. For example, if $\mathcal{F}^\gamma = \mathcal{T}^\gamma$, we are reduced to estimating $\mathbb{E}\left(\sum_{o' \sim o} |\frac{\overline{\hat{\zeta}^\gamma_{o'}(o)}\hat{\zeta}^\gamma_o(o')}{\overline{\hat{\zeta}^\gamma_{o'}(o)}\hat{\zeta}^\gamma_o(o')+1}|^s \right)$. Using (2.7), we observe that

$$\frac{|\hat{\zeta}^\gamma_o(o')|}{\left| \hat{\zeta}^\gamma_o(o') + \overline{\hat{\zeta}^\gamma_{o'}(o)}^{-1} \right|} = \frac{|\hat{\zeta}^\gamma_o(o')|}{|2 \operatorname{Re} \hat{\zeta}^\gamma_o(o') + \overline{2\hat{m}^\gamma_o}|} \le \frac{|\hat{\zeta}^\gamma_o(o')|}{2 \operatorname{Im} \hat{m}^\gamma_o},$$

and we know from Remark A.4 that $\sup_\gamma \mathbb{E}\left(\sum_{o' \sim o} \frac{|\hat{\zeta}^\gamma_o(o')|^s}{(2 \operatorname{Im} \hat{m}^\gamma_o)^s} \right) < \infty$. Similarly, if $\mathcal{F}^\gamma = \mathcal{L}^\gamma d^{-1} \mathcal{S}_{T,\gamma}$, then

$$|(\mathcal{F}^\gamma K)(e)|$$
$$\le \frac{|\zeta^\gamma_{o_e}(t_e)|^2}{|m^\gamma_{o_e}|^2} \frac{1}{N_\gamma(o_e) N_\gamma(t_e)} \sum_{r=0}^{T-1} \left[|(P^r d^{-1} N_\gamma K)(o_e)| + \frac{|(P^r d^{-1} N_\gamma K)(t_e)|}{|\zeta^\gamma_{o_e}(t_e) \zeta^\gamma_{t_e}(o_e)|} \right],$$

so (10.14) reduces to

$$C_{D,T,s}\,\mathbb{E}\left(\sum_{o'\sim o}\Big(\frac{|\hat{\zeta}_o^{\gamma}(o')|^2}{|\hat{m}_o^{\gamma}|^2\hat{N}_\gamma(o)\hat{N}_\gamma(o')}\Big)^{p_1 s}\right)^{1/p_1}\mathbb{E}\left(\hat{N}_\gamma(o)^{p_2 s}\right)^{1/p_2}$$
$$+\,C_{D,T,s}\,\mathbb{E}\left(\sum_{o'\sim o}\Big(\frac{|\hat{\zeta}_o^{\gamma}(o')|}{|\hat{m}_o^{\gamma}|^2\hat{N}_\gamma(o)\hat{N}_\gamma(o')|\hat{\zeta}_{o'}^{\gamma}(o)|}\Big)^{p_1 s}\right)^{1/p_1}\mathbb{E}\left(\hat{N}_\gamma(o)^{p_2 s}\right)^{1/p_2}$$

for some p_1, p_2.

The previous discussion was under the assumption $A^{(r)} \neq \mathcal{X}_\gamma$. If $\mathcal{F}_\gamma = F_1^\gamma \mathcal{X}_\gamma F_2^\gamma$ with F_1^γ and F_2^γ as in the previous paragraph, we write

$$\mathcal{F}_\gamma K(\omega) = \sum_{\omega'} F_1^\gamma(\omega,\omega')\langle F_2^\gamma K\rangle_\gamma,$$

with

$$\begin{aligned}|\langle F_2^\gamma K\rangle_\gamma| &= \left|\frac{\sum_x N_\gamma(x)(F_2^\gamma K)(x)}{\sum_x N_\gamma(x)}\right| \\ &= \left|\frac{\sum_x\sum_w N_\gamma(x)F_2^\gamma(x,w)K(w)}{\sum_x N_\gamma(x)}\right| \le \frac{\sum_x\sum_w N_\gamma(x)|F_2^\gamma(x,w)|}{\sum_x N_\gamma(x)}.\end{aligned}$$

Hence,

$$|\mathcal{F}_\gamma K(\omega)| \le \sum_{\omega'}|F_1^\gamma(\omega,\omega')|\cdot\frac{N}{\sum_x N_\gamma(x)}\cdot\frac{1}{N}\sum_x\sum_w N_\gamma(x)|F_2^\gamma(x,w)|.$$

Applying Hölder's inequality to

$$\frac{1}{N}\sum_{\omega\in B_k}\left(\sum_{\omega'}|F_1^\gamma(\omega,\omega')|\right)^s \quad\text{and}\quad \left(\frac{1}{N}\sum_x N_\gamma(x)\sum_w|F_2^\gamma(x,w)|\right)^s$$

and taking the limit, we obtain a uniform control as before. Thus, all points of (Hol) are satisfied.

In view of Remark 10.4, we may use Theorem 3.3 to conclude that for the $\mathcal{F}_\gamma$ in Corollary 10.3, we have $\lim_{\eta_0\downarrow 0}\limsup_{N\to\infty}\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K) = 0$.

Since $\widetilde{\mathrm{Var}}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\widetilde{\mathcal{F}}_\gamma K)$ is defined exactly like $\mathrm{Var}^{\mathrm{I}}_{\mathrm{nb},\eta_0}(\mathcal{F}_\gamma K)$ except that ζ is replaced by $\overline{\zeta}$, it is clear that it can be shown to vanish asymptotically using the same arguments, simply replacing ζ by $\overline{\zeta}$ when necessary. By Corollary 10.3, to finish the proof of Theorem 1.3, it suffices to show (10.13). This is what we do now.

Recall that we introduced $\|K\|_\gamma$ for $K \in \mathscr{H}_k$, $k \ge 1$, in (4.1). For $K \in \mathscr{H}_0$, we let

$$\|K\|_\gamma^2 = \|N_\gamma K\|_{\mathscr{H}_0}^2 = \frac{1}{N}\sum_{x\in V} N_\gamma^2(x)|K(x)|^2.$$

We also define

$$(Y_\gamma K)(x) = \frac{d(x)}{N_\gamma(x)} \cdot \frac{\sum_{y\in V} N_\gamma(y)K(y)}{\sum_{y\in V} d(y)}.$$

Denoting $\langle J\rangle_U := \frac{1}{N}\sum_{x\in V} J(y)$ as the uniform average of J, we have $Y_\gamma K = \frac{\langle N_\gamma K\rangle_U}{\langle d\rangle_U} \cdot \frac{d}{N_\gamma}$. Fix $I = (a,b) \subset I_1$ as in Section 4.

PROPOSITION 10.5. *Under assumptions* (BSCT) *and* (Green), *if* $K^\gamma \in \mathscr{H}_0$ *satisfies the set of assumptions* (Hol), *then for any interval* $I = (a,b)$ *as above*,

$$\lim_{\eta_0\downarrow 0} \limsup_{N\to+\infty} \mathrm{Var}^{\mathrm{I}}_{\eta_0}(\widetilde{\mathcal{S}}_{T,\gamma}K^\gamma - Y_\gamma K^\gamma)^2 \le \frac{D\,|I|}{\beta^2T^2} \lim_{\eta_0\downarrow 0}\lim_{\eta\downarrow 0}\limsup_{N\to\infty} \int_{a-2\eta}^{b+2\eta} \|K^{\lambda+i(\eta^4+\eta_0)} - Y_{\lambda+i(\eta^4+\eta_0)}K^{\lambda+i(\eta^4+\eta_0)}\|^2_{\lambda+i(\eta^4+\eta_0)}\,\mathrm{d}\lambda.$$

Proof. We follow the steps in the proof of Theorem 4.1. Let

$$J^\gamma = (\widetilde{\mathcal{S}}_{T,\gamma} - Y_\gamma)K^\gamma \text{ and } \alpha_{\gamma_j}(x) = N_{\gamma_j}^{1/2}(x).$$

Then

$$\mathrm{Var}^{\mathrm{I}}_{\eta_0}(J^\gamma)^2 \le \left(\frac{1}{N}\sum_{\lambda_j\in I} \|\alpha_{\gamma_j}^{-1}\psi_j\|^2\right)\left(\frac{1}{N}\sum_{\lambda_j\in I}\|\alpha_{\gamma_j} J_G^{\gamma_j}\psi_j\|^2\right).$$

As in the proof of (4.3),

$$\frac{1}{N}\sum_{\lambda_j\in I}\|\alpha_{\gamma_j}^{-1}\psi_j\|^2 \lesssim \frac{3}{\pi N}\int_{a-2\eta}^{b+2\eta} \sum_{\rho_G(x)\ge d_{R,\eta}} \frac{\Psi_{z+i\eta_0,\tilde{x}}(\tilde{x})}{N_{\lambda+i\eta_0}(x)}\,\mathrm{d}\lambda \le \frac{3(|I|+4\eta)}{\pi}$$

for any small $\eta > 0$, since $N_\gamma(x) = \Psi_{\gamma,\tilde{x}}(\tilde{x})$.

Hence,

$$\lim_{\eta_0\downarrow 0}\limsup_{N\to\infty} \mathrm{Var}^{\mathrm{I}}_{\eta_0}(J^\gamma)^2 \le \frac{3|I|}{\pi}\lim_{\eta_0\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\sum_{\lambda_j\in I}\|\alpha_{\gamma_j}J_G^{\gamma_j}\psi_j\|^2.$$

Now

$$\|\alpha_{\gamma_j}J_G^{\gamma_j}\psi_j\|^2 = \sum_{x\in V} N_{\gamma_j}(x)|J^{\gamma_j}(x)|^2|\psi_j(x)|^2.$$

Arguing as in Section 4, we get

$$\frac{1}{N}\sum_{\lambda_j\in I}\|\alpha_{\gamma_j}J_G^{\gamma_j}\psi_j\|^2 \le \frac{3}{\pi N}\int_{a-2\eta}^{b+2\eta}\sum_{\rho_G(x)\ge d_{R,\eta}} \chi(\lambda)N_{z+i\eta_0}(x)|J^{z+i\eta_0}(x)|^2\Psi_{z+i\eta_0,\tilde{x}}(\tilde{x})\,\mathrm{d}\lambda,$$

where $z := \lambda + i\eta^4$. This is bounded by $\frac{3}{\pi}\int_{a-2\eta}^{b+2\eta}\|J^{z+i\eta_0}\|^2_{z+i\eta_0}\,\mathrm{d}\lambda$, since $\Psi_{\gamma,\tilde{x}}(\tilde{x}) = N_\gamma(x)$ and $\chi(\lambda) \le 1$ on $\mathbb{R}$.

Summarizing, we have

$$\lim_{\eta_0\downarrow 0}\limsup_{N\to\infty}\mathrm{Var}^{\mathrm{I}}_{\eta_0}(J^\gamma)^2 \le \frac{9|I|}{\pi^2}\int_{a-2\eta}^{b+2\eta}\|J^{z+i\eta_0}\|^2_{z+i\eta_0}\,\mathrm{d}\lambda.$$

Now recall that $\widetilde{\mathcal{S}}_{T,\gamma} = \frac{1}{T}\sum_{s=1}^T P_\gamma^s$ and $P_\gamma = \frac{d}{N_\gamma}P\frac{N_\gamma}{d}$, so that $P_\gamma^s = \frac{d}{N_\gamma}P^s\frac{N_\gamma}{d}$. Moreover, $Y_\gamma K = \frac{d}{N_\gamma}\frac{\langle N_\gamma K\rangle_U}{\langle d\rangle_U}$. So denoting $\gamma = z+i\eta_0$, $\|K\|_d^2 = \frac{1}{N}\sum_{x\in V} d(x)|K(x)|^2$, we have

$$\begin{aligned}\|J^\gamma\|_\gamma^2 = \|N_\gamma J^\gamma\|^2_{\mathscr{H}_0} &= \frac{1}{N}\sum_{x\in V}\left|\frac{1}{T}\sum_{s=1}^T d(x)\Big(P^s\frac{N_\gamma K^\gamma}{d}\Big)(x) - \frac{\langle N_\gamma K^\gamma\rangle_U}{\langle d\rangle_U}d(x)\right|^2\\
&\le D\cdot\left\|\frac{1}{T}\sum_{s=1}^T P^s\Big(\frac{N_\gamma K^\gamma}{d} - \frac{\langle N_\gamma K^\gamma\rangle_U}{\langle d\rangle_U}\mathbf{1}\Big)\right\|_d^2\\
&\le \frac{D}{T^2}\left(\sum_{s=1}^T(1-\beta)^s\left\|\frac{N_\gamma K^\gamma}{d} - \frac{\langle N_\gamma K^\gamma\rangle_U}{\langle d\rangle_U}\mathbf{1}\right\|_d\right)^2\\
&\le \frac{D}{\beta^2T^2}\left\|\frac{N_\gamma K^\gamma}{d} - \frac{\langle N_\gamma K^\gamma\rangle_U}{\langle d\rangle_U}\mathbf{1}\right\|_d^2.\end{aligned}$$

Here we used (EXP) and the fact that $\frac{N_\gamma K^\gamma}{d} - \frac{\langle N_\gamma K^\gamma\rangle_U}{\langle d\rangle_U}\mathbf{1}$ is orthogonal to the constants in $\ell^2(V,d)$. Indeed, the orthogonal projector onto $\mathbf{1}$ in $\ell^2(V,d)$ is $P_{\mathbf{1},d}J = \frac{\langle \mathbf{1},J\rangle_d}{\langle\mathbf{1},\mathbf{1}\rangle_d}\mathbf{1} = \frac{\langle dJ\rangle_U}{\langle d\rangle_U}\mathbf{1}$. Since $\frac{\langle N_\gamma K^\gamma\rangle_U}{\langle d\rangle_U}\mathbf{1} = \frac{N_\gamma Y_\gamma K^\gamma}{d}$ and $\frac{1}{d}\le 1$, the proposition follows. □

Corollary 10.6. *For any $C_\gamma : \mathscr{H}_m \to \mathscr{H}_0$ as in Corollary 10.3 and $\bar I \subset I_1$, $\|K\|_\infty \le 1$,*

$$\lim_{\eta_0\downarrow 0}\limsup_{N\to+\infty}\mathrm{Var}^{\mathrm{I}}_{\eta_0}\left(\widetilde{\mathcal{S}}_{T,\gamma}\left(C_\gamma K - \langle C_\gamma K\rangle_\gamma\right)\right)^2 \le \frac{c\,|I|^2}{\beta^2T^2}.$$

Proof. Let $K'_\gamma = C_\gamma K - \langle C_\gamma K\rangle_\gamma\mathbf{1}$. Then $Y_\gamma K'_\gamma = 0$, since $Y_\gamma C_\gamma K = \frac{d}{N_\gamma}\frac{\langle N_\gamma C_\gamma K\rangle_U}{\langle d\rangle_U}$ and $\langle C_\gamma K\rangle_\gamma Y_\gamma\mathbf{1} = \frac{\langle N_\gamma C_\gamma K\rangle_U}{\langle N_\gamma\rangle_U}\frac{d}{N_\gamma}\frac{\langle N_\gamma\rangle_U}{\langle d\rangle_U}$. Hence, denoting $z = \lambda + i(\eta^4+\eta_0)$,

$$\begin{aligned}&\lim_{\eta_0\downarrow 0}\limsup_{N\to+\infty}\mathrm{Var}^{\mathrm{I}}_{\eta_0}\left(\widetilde{\mathcal{S}}_{T,\gamma}\left(C_\gamma K - \langle C_\gamma K\rangle_\gamma\right)\right)^2\\
&\qquad\le \frac{D\,|I|}{\beta^2T^2}\lim_{\eta_0\downarrow 0}\lim_{\eta\downarrow 0}\limsup_{N\to\infty}\int_{a-2\eta}^{b+2\eta}\|C_zK - \langle C_zK\rangle_z\|_z^2\,\mathrm{d}\lambda.\end{aligned}$$

Now

$$\|C_zK\|_z^2 = \frac{1}{N}\sum_{x\in V}N_z^2(x)|(C_zK)(x)|^2 \le \frac{1}{N}\sum_{x\in V}N_z^2(x)\Big[\sum_{w\in B_m}|C_z(x,w)|\Big]^2.$$

Similarly, $|\langle C_zK\rangle_z| \le \frac{1}{\sum_x N_z(x)}\sum_x N_z(x)\sum_w|C_z(x,w)|$. For our operators C_z, we thus get $\|C_zK\|_z^2 = O(1)_{N\longrightarrow+\infty,z}$ and $|\langle C_zK\rangle_z| = O(1)_{N\longrightarrow+\infty,z}$, as in Remark 10.4. □

This proves (10.13) and ends the proof of Theorem 1.3 on the interval I.

Suppose further that $\rho(\partial I_1) = 0$. As I_1 is open, we have $I_1 = \cup_{j\in\mathbb{N}} J_j$ for open intervals $J_j = (a_j, b_j)$. Let $J_j^\varsigma = (a_j + \varsigma, b_j - \varsigma)$ with $\varsigma > 0$ small. Then $\overline{J_j^\varsigma} \subset I_1$, so using (9.7) and Corollary 10.6, we get

$$\lim_{\eta_0\downarrow 0}\limsup_{N\to\infty} \mathrm{Var}_{\eta_0}^{\mathrm{J_j^\varsigma}}(K - \langle K\rangle_\gamma) = 0.$$

Now $\mathrm{Var}_{\eta_0}^{\mathrm{I_1}}(K') = \sum_{j=1}^M \mathrm{Var}_{\eta_0}^{\mathrm{J_j^\varsigma}}(K') + \mathrm{Var}_{\eta_0}^{\mathrm{I_1\setminus\cup_{j=1}^M J_j^\varsigma}}(K')$ for any given M. By (A.14) and (Green), we have

$$\mathrm{Var}_{\eta_0}^{\mathrm{I_1\setminus\cup_{j=1}^M J_j^\varsigma}}(K - \langle K\rangle_\gamma) \le \frac{\sharp\{\lambda_j \in I_1 \setminus \cup_{k=1}^M J_k^\varsigma\}}{N}\, O(1)_{N\longrightarrow+\infty,\gamma}.$$

By the convergence of empirical spectral measures (Remark A.3), and using the fact that $\rho(\partial I_1) = 0$, we have $\frac{\sharp\{\lambda_j\in I_1\setminus\cup_{k=1}^M J_k^\varsigma\}}{N} \to \rho(I_1 \setminus \cup_{k=1}^M J_k^\varsigma)$. Finally, $\rho(I_1 \setminus \cup_{k=1}^M J_k^\varsigma) \to 0$ as $\varsigma \downarrow 0$ and $M \longrightarrow +\infty$. The conclusion of Theorem 1.3 thus holds with I replaced by I_1.

Finally, if (Green) holds on $\overline{I_1}$, then

$$\rho(\{\lambda\}) = \lim_{\eta\downarrow 0} \eta \,\mathrm{Im}\,\mathbb{E}(\mathcal{G}^{\lambda+i\eta}(o,o)) = 0$$

for any $\lambda \in \overline{I_1}$, since $\sup_{\eta>0} \mathrm{Im}\,\mathbb{E}(\mathcal{G}^{\lambda+i\eta}(o,o)) < \infty$. In particular, $\rho(\partial I_1) = 0$.

Appendix A. Benjamini-Schramm topology

A.1. *Generalities.* Here we collect known facts on the Benjamini-Schramm convergence; we refer the reader to [1], [4], [15], [14], [37] for details.

A *colored rooted graph* (G, o, W) is a graph $G = (V, E)$ with a marked vertex $o \in V$ called the *root*, and a map $W : V \to \mathbb{R}$ that we see as a "coloring"; it can also be regarded as a potential on $\ell^2(V)$. This is a special case of what is called a *network* in [4]. All graphs are assumed to be *locally finite*; i.e., each vertex has a finite degree.

If G is connected, we denote by $B_G(x, r)$ the *r-ball* $\{y \in V : d_G(x, y) \le r\}$, where d_G is the length of the shortest path between x and y in G.

As in [4], we define a distance between colored connected graphs by

$$d_{\mathrm{loc}}\big((G, o, W), (G', o', W')\big) = \frac{1}{1+\alpha}, \tag{A.1}$$

$$\alpha := \sup\Big\{r > 0 : \exists \text{ graph isomorphism } \phi : B_G(o, \lfloor r\rfloor) \to B_{G'}(o', \lfloor r\rfloor) \text{ with } \phi(o) = o' \text{ and } |W'(\phi(v)) - W(v)| < 1/r \ \forall v \in B_G(o, \lfloor r\rfloor)\Big\}.$$

Two colored rooted graphs (G, o, W) and (G', o', W') are *equivalent* if there is a graph isomorphism $\phi : G \to G'$ such that $\phi(o) = o'$ and $W' \circ \phi = W$. We denote the equivalence class of (G, o, W) by $[G, o, W]$.

Let $\mathscr{G}_*$ be the set of equivalence classes of connected colored rooted graphs. Then d_{loc} turns $\mathscr{G}_*$ into a separable complete metric space. We may thus consider the set of probability measures on $\mathscr{G}_*$, denoted by $\mathcal{P}(\mathscr{G}_*)$.

Any finite connected colored graph (G, W), $G = (V, E)$, defines a probability measure $U_{(G,W)} \in \mathcal{P}(\mathscr{G}_*)$ by choosing the root x uniformly at random in V:

$$U_{(G,W)} = \frac{1}{|V|} \sum_{x \in V} \delta_{[G,x,V]}. \tag{A.2}$$

If (G_n, W_n) is a sequence of finite colored graphs, we say that $\mathbb{P} \in \mathcal{P}(\mathscr{G}_*)$ is the *local weak limit* of (G_n, W_n) if $U_{(G_n,W_n)}$ converges weakly-$*$ to $\mathbb{P}$ in $\mathcal{P}(\mathscr{G}_*)$. This notion of convergence was introduced in [15] and generalized in [4]. In this case, we also say that (G_n, W_n) converges in the sense of Benjamini-Schramm.

The subset $\mathscr{G}_*^{D,A} \subset \mathscr{G}_*$ of equivalence classes $[G, o, W]$ such that G is of degree bounded by D, and W takes values in $[-A, A]$, is compact. It follows that $\mathcal{P}(\mathscr{G}_*^{D,A})$ is compact in the weak-$*$ topology. Hence, if $\mathcal{C}_{\mathrm{fin}}^{D,A}$ denotes the set of finite colored graphs (G, W), $G = (V, E)$, of degree bounded by D and coloring $W : V \to [-A, A]$, then any sequence $(G_n, W_n) \subset \mathcal{C}_{\mathrm{fin}}^{D,A}$ has a subsequence that converges in the sense of Benjamini-Schramm.

Let $C(\mathscr{G}_*^{D,A})$ be the set of continuous functions $f : \mathscr{G}_*^{D,A} \to \mathbb{R}$. Then a sequence $(G_n, W_n) \subset \mathcal{C}_{\mathrm{fin}}^{D,A}$ has a local weak limit $\mathbb{P}$ if and only if there is an algebra $\mathscr{A} \subset C(\mathscr{G}_*^{D,A})$ that separates points such that for all $f \in \mathscr{A}$,

$$\lim_{n\to\infty} \frac{1}{|V_n|} \sum_{x \in V_n} f\left([G_n, x, W_n]\right) = \int_{\mathscr{G}_*^{D,A}} f\left([G, o, W]\right) \mathrm{d}\,\mathbb{P}\left([G, o, W]\right). \tag{A.3}$$

This follows from the compactness of $\mathscr{G}_*^{D,A}$; see [33, Ch. 13].

It may not be very clear what a continuous function on $\mathscr{G}_*^{D,A}$ looks like, so we give a basic example. If $B_F(o, r)$ is an r-ball, the sets $\mathscr{C}_F = \{[G, x, W] : B_G(x, r) \cong B_F(o, r)\}$ turn out to be clopen in $\mathscr{G}_*^{D,A}$, so the characteristic function $\chi_{\mathscr{C}_F}$ is continuous. Here $B_G(x, r) \cong B_F(o, r)$ means there exists a graph isomorphism $\phi : B_G(x, r) \to B_F(o, r)$ with $\phi(x) = o$. Using (A.3), it can be shown that in the special case where there is no coloring, $(G_n) \subset \mathcal{C}_{\mathrm{fin}}^{D,A}$ has a local weak limit $\mathbb{P}$ if and only if

$$\lim_{n\to\infty} \frac{\#\{x : B_{G_n}(x, r) \cong B_F(o, r)\}}{|V_n|} = \mathbb{P}(\{[G, x] : B_G(x, r) \cong B_F(o, r)\})$$

for any $B_F(o, r)$. This was in fact the original criterion in [15]. Using it, one readily checks that a sequence of $(q+1)$-regular graphs (G_n) satisfies (BST) if and only if it converges to the $(q+1)$-regular tree $\mathbb{T}_q$ in the sense of Benjamini-Schramm, i.e., if and only if (G_n) has the local weak limit $\delta_{[\mathbb{T}_q, o]}$, with $o \in \mathbb{T}_q$ arbitrary. More generally, by considering the clopen sets $\mathscr{C}_r = \{[G, x, W] : B_G(x, r) \text{ is not a tree}\}$, one sees that if $(G_n, W_n) \subset \mathcal{C}_{\mathrm{fin}}^{D,A}$ has a local weak limit $\mathbb{P}$ that is concentrated on the subset $\mathscr{T}_*^{D,A} \subset \mathscr{G}_*^{D,A}$ of colored rooted trees, then (G_n) satisfies (BST). Conversely, if (G_n) satisfies (BST) and if a subsequence of (G_n, W_n) has a local weak limit $\mathbb{P}$, then $\mathbb{P}$ must be concentrated on $\mathscr{T}_*^{D,A}$.

A.2. *Convergence of empirical spectral measures.* We now show Benjamini-Schramm convergence implies convergence of the empirical spectral measures. This is already known in some settings [1], [37], [40]. In this paper we need the variant stated as Corollary A.2.

Given $[G, o, W] \in \mathscr{G}_*^{D,A}$, $\gamma \in \mathbb{C}^+ = \{z, \operatorname{Im} z > 0\}$ and $x \sim y \in G$, we define $\zeta_x^\gamma(y)$ as in Section 2.2. Like in Section 2.1, B_k is the set of non-backtracking paths of length k on G.

Fix $s \in \mathbb{N}$. Let $F : (\mathbb{C} \setminus \{0\})^{2s} \to \mathbb{C}$ be a continuous function and $\gamma \in \mathbb{C}^+$. Let

(A.4)
$$F_\gamma([G, o, W]) = \sum_{(x_0;x_s)\in B_s\,:\, x_0=o} F\left(\zeta_{x_0}^\gamma(x_1), \zeta_{x_1}^\gamma(x_0), \dots, \zeta_{x_{s-1}}^\gamma(x_s), \zeta_{x_s}^\gamma(x_{s-1})\right).$$

For $s = 1$, the sum reduces to $\sum_{x_1 : x_1 \sim o}$. One can remark that $F_\gamma([G, o, W]) = F_\gamma([\widetilde{G}, \tilde{o}, \widetilde{W}])$, where $\widetilde{G}$ is the universal cover of G and $\tilde{o}, \widetilde{W}$ are lifts of o, W.

Next, given Borel $J \subseteq \mathbb{R}$, we define the measure

$$\mu_{o,F,\gamma}^{(G,W)}(J) = F_\gamma([G, o, W])\langle \delta_o, \chi_J(H_{G,W})\delta_o\rangle.$$

Fix a compact $I \subset \mathbb{R}$, and fix $\eta \in (0, 1)$.

LEMMA A.1. *Suppose that $(\lambda_n, [G_n, o_n, W_n]) \subset I \times \mathscr{G}_*^{D,A}$ converges to $(\lambda, [G, o, W])$ in $I \times \mathscr{G}_*^{D,A}$. Then $\mu_{o_n,F,\lambda_n+i\eta}^{(G_n,W_n)}$ converges weakly-$*$ to $\mu_{o,F,\lambda+i\eta}^{(G,W)}$.*

Proof. Since all operators $H_n = H_{(G_n,W_n)}$ and $H = H_{(G,W)}$ are uniformly bounded by $D + A$, the supports of the spectral measures is compact, so it suffices to show that for any $k \in \mathbb{N}$, $\mu_{o_n,F,\lambda_n+i\eta}^{(G_n,W_n)}(t^k) \to \mu_{o,F,\lambda+i\eta}^{(G,W)}(t^k)$; see [33, Ch. 13].

Let $k \in \mathbb{N}$. Denote $\gamma_n = \lambda_n + i\eta$, $\gamma = \lambda + i\eta$. We have

$$\begin{aligned}&\left|\mu_{o_n,F,\gamma_n}^{(G_n,W_n)}(t^k) - \mu_{o,F,\gamma}^{(G,W)}(t^k)\right| \\ &\qquad = \left|F_{\gamma_n}([G_n, o_n, W_n])\langle \delta_{o_n}, H_n^k \delta_{o_n}\rangle - F_\gamma([G, o, W])\langle \delta_o, H^k \delta_o\rangle\right|.\end{aligned}$$

We first approximate F by a polynomial.

We have $|\zeta_x^{\lambda+i\eta}(y)| \le \eta^{-1}$ and $|\operatorname{Im} \zeta_x^{\lambda+i\eta}(y)| = \eta\, \|(\widetilde{H}^{(\tilde{y}|\tilde{x})} - \lambda - i\eta)^{-1}\delta_{\tilde{y}}\|_{\ell^2(\widetilde{G})}^2$. Since $\|\widetilde{H}^{(x|y)} - \lambda - i\eta\|_{\ell^2 \to \ell^2} \le A + D + c_I + 1 =: c$ for all $\lambda \in I$ and $\eta \in (0, 1)$, we get $|\operatorname{Im} \zeta_x^{\lambda+i\eta}(y)| \ge \eta c^{-2}$.

So let $\mathcal{O} \subset \mathbb{C}$ be the compact region $\{\eta c^{-2} \le |z| \le \eta^{-1}\}$. If F is continuous on $\mathcal{O}^{2s} \subset \mathbb{C}^{2s}$, by Stone-Weierstrass, given $R \in \mathbb{N}^*$, there is a polynomial P_R

of $4s$ variables such that

$$\sup_{(z_1;z_{2s})\in\mathcal{O}^{2s}} |F(z_1,\dots,z_{2s}) - P_R(z_1,\bar z_1,\dots,z_{2s},\bar z_{2s})| \le \frac{1}{2R}.$$

Hence, for any $\lambda \in I$ and $(x_0;x_s)$, if $\gamma = \lambda + i\eta$, then

$$\Big| F\left(\zeta_{x_0}^{\gamma}(x_1), \zeta_{x_1}^{\gamma}(x_0), \dots, \zeta_{x_s}^{\gamma}(x_{s-1})\right) - P_R\left(\zeta_{x_1}^{\gamma}(x_0), \overline{\zeta_{x_1}^{\gamma}(x_0)}, \dots, \overline{\zeta_{x_s}^{\gamma}(x_{s-1})}\right)\Big| \le \frac{1}{2R}. \tag{A.5}$$

Let $h_\eta(t) = -(t-i\eta)^{-1}$. Given $\epsilon > 0$, we may choose a polynomial $Q_\epsilon = Q_\epsilon^\eta$ such that $\|h_\eta - Q_\epsilon\|_\infty < \epsilon$. It follows that $\|h_\eta(H_{\tilde G}^{(\tilde x|\tilde y)} - \lambda) - Q_\epsilon(H_{\tilde G}^{(\tilde x|\tilde y)} - \lambda)\| < \epsilon$. In particular, if $Z_\epsilon^\gamma(x,y) := Q_\epsilon(H_{\tilde G}^{(\tilde y|\tilde x)} - \lambda)(\tilde y,\tilde y)$, we have for any $\lambda \in I$ and $(x,y)\in B$,

$$|\zeta_x^\gamma(y) - Z_\epsilon^\gamma(x,y)| < \epsilon. \tag{A.6}$$

As P_R is Lipschitz-continuous on $\mathcal{O}^{2s}$, we may thus find $C_{R,\eta^{-1}}$ such that

$$\Big| P_R\left(\zeta_{x_0}^{\gamma}(x_1), \dots, \overline{\zeta_{x_s}^{\gamma}(x_{s-1})}\right) - P_R\left(Z_\epsilon^\gamma(x_0,x_1), \dots, \overline{Z_\epsilon^\gamma(x_s,x_{s-1})}\right)\Big| \le C_{R,\eta^{-1}}\cdot\epsilon = \frac{1}{2R}$$

by choosing $\epsilon = \frac{1}{2R}\frac{1}{C_{R,\eta^{-1}}}$. Using (A.5), we thus get uniformly in $\lambda \in I$, $(x_0;x_s)$,

$$\Big| F\left(\zeta_{x_0}^{\gamma}(x_1), \zeta_{x_1}^{\gamma}(x_0), \dots, \zeta_{x_s}^{\gamma}(x_{s-1})\right) - P_R\left(Z_R^\gamma(x_0,x_1), \dots, \overline{Z_R^\gamma(x_s,x_{s-1})}\right)\Big| \le \frac{1}{R}, \tag{A.7}$$

where we now denote Z_R because ϵ is a function of R. Define

$$P_\gamma([G,o,W]) = \sum_{(x_1;x_s),x_0=o} P_R\left(Z_R^\gamma(x_0,x_1),\dots,\overline{Z_R^\gamma(x_s,x_{s-1})}\right).$$

Then up to an error $\frac{C_{D,s,A,k}}{R}$, it suffices to consider

$$\Big| P_{\gamma_n}([G_n,o_n,W_n])\langle \delta_{o_n}, H_n^k\delta_{o_n}\rangle - P_\gamma([G,o,W])\langle\delta_o, H^k\delta_o\rangle\Big|.$$

Let d_R be the degree of Q_R, and choose an arbitrary integer $r \ge d_R + s + k =: d_{R,s,k}$. Then we may find n_r such that for $n \ge n_r$, there exists $\varphi_r : B_{G_n}(o_n,r) \xrightarrow{\sim} B_G(o,r)$ with $\|W\circ\varphi_r - W_n\|_{B_{G_n}(o,r)} < 1/r$. Now

$$\langle\delta_{o_n}, H_n^k\delta_{o_n}\rangle = \sum_{u_0,\dots,u_{k-1}} H_n(o_n,u_0)H_n(u_0,u_1)\dots H_n(u_{k-1},o_n)$$

and

$$H_n(v,w) = \mathcal{A}_n(v,w) + W_n(v)\delta_w(v).$$

This only depends on $B_{G_n}(o_n, k)$ and its coloring. Similarly, the quantity $Z_R^\gamma(x,y)$ corresponding to (G_n, o_n, W_n) only depends on $B_{G_n}(y, d_R)$ and its coloring. Since $r \geq d_{R,s,k}$ and $\varphi_r : B_{G_n}(o_n, r) \xrightarrow{\sim} B_G(o, r)$, if we let $\mathcal{H}_n = \mathcal{A}_G + W_n \circ \varphi_r^{-1}$ on G, we get $\langle \delta_{o_n}, H_n^k \delta_{o_n}\rangle = \langle \delta_o, \mathcal{H}_n^k \delta_o\rangle$. Similarly, $P_{\gamma_n}([G_n, o_n, W_n]) = P_{\gamma_n}([G, o, W_n \circ \varphi_r^{-1}])$. Let $W_n' = W_n \circ \varphi_r^{-1}$. Then for $n \geq n_r$,

$$\begin{aligned}&\left|\mu_{o_n,F,\gamma_n}^{(G_n,W_n)}(t^k) - \mu_{o,F,\gamma}^{(G,W)}(t^k)\right| \\ &\qquad\leq \frac{C}{R} + \left|P_{\gamma_n}([G,o,W_n'])\langle \delta_o, \mathcal{H}_n^k \delta_o\rangle - P_\gamma([G,o,W])\langle \delta_o, H^k \delta_o\rangle\right|.\end{aligned}$$

Writing $\mathcal{H}_n^k - H^k = \sum_{i=1}^k \mathcal{H}_n^{k-i}(\mathcal{H}_n - H)H^{i-1}$, we have

$$|\langle \delta_o, (\mathcal{H}_n^k - H^k)\delta_o\rangle| \leq C_{k,D,A}' \|W_n \circ \varphi_r^{-1} - W\|_{B_G(o,r)} \leq \frac{C_{k,D,A}'}{r}.$$

A similar argument yields

$$|P_\gamma([G,o,W_n']) - P_\gamma([G,o,W])| \leq \frac{C_{R,D,s,A}}{r}$$

and

$$|P_{\gamma_n}([G,o,W_n']) - P_\gamma([G,o,W_n'])| \leq C_{R,D,s,A,I}|\lambda_n - \lambda| \leq \frac{C_{R,D,s,A,I}}{r}$$

for $n \geq n_r'$. We thus showed that for any $r \geq d_{R,s,k}$, there exists n_r'' such that if $n \geq n_r''$, then

$$|\mu_{o_n,F,\gamma_n}^{(G_n,W_n)}(t^k) - \mu_{o,F,\gamma}^{(G,W)}(t^k)| \leq \frac{C_{D,s,A,k}}{R} + \frac{C_{k,D,A}' + C_{R,D,s,A} + C_{R,D,s,A,I}}{r}.$$

It follows that $\limsup_{n\to\infty} |\mu_{o_n,F,\gamma_n}^{(G_n,W_n)}(t^k) - \mu_{o,F,\gamma}^{(G,W)}(t^k)| \leq \frac{C_{D,s,A,k}}{R}$. Since R is arbitrary, the proof is complete. □

If $(G,W) \in \mathcal{C}_{\text{fin}}^{D,A}$, we now define, for $\gamma \in \mathbb{C}^+$,

$$\mu_{F,\gamma}^{(G,W)} = \frac{1}{|V|}\sum_{x\in V} \mu_{x,F,\gamma}^{(G,W)}.$$

COROLLARY A.2. *Suppose* $(G_n, W_n) \subset \mathcal{C}_{\text{fin}}^{D,A}$ *has a local weak limit* $\mathbb{P}$. *Fix a compact* $I \subset \mathbb{R}$ *and* $\eta \in (0,1)$. *Then* $\mu_{F,\lambda+i\eta}^{(G_n,W_n)}$ *converges weakly to* $\int_{\mathscr{G}_*^{D,A}} \mu_{o,F,\lambda+i\eta}^{(G,W)}\,\mathrm{d}\,\mathbb{P}([G,o,W])$, *uniformly in* $\lambda \in I$. *In other words, for any continuous* $\varphi : \mathbb{R} \to \mathbb{R}$, *we have uniformly in* $\lambda \in I$,

$$\begin{aligned}&\frac{1}{|V_n|}\sum_{x\in V_n} F_{\lambda+i\eta}([G_n,x,W_n])\langle \delta_x, \varphi(H_{(G_n,W_n)})\delta_x\rangle \\ &\qquad\underset{N\longrightarrow+\infty}{\longrightarrow} \int_{\mathscr{G}_*^{D,A}} F_{\lambda+i\eta}([G,o,W])\langle \delta_o, \varphi(H_{(G,W)})\delta_o\rangle\,\mathrm{d}\,\mathbb{P}([G,o,W]).\end{aligned}$$

Proof. Given continuous $\varphi : \mathbb{R} \to \mathbb{R}$, define $\widehat{\varphi} : I \times \mathscr{G}_*^{D,A} \to \mathbb{R}$ by $\widehat{\varphi}(\lambda, [G,o,W]) = \int \varphi(t)\,\mathrm{d}\mu_{o,F,\lambda+i\eta}^{(G,W)}(t)$. Lemma A.1 states $\widehat{\varphi}$ is continuous on $I \times \mathscr{G}_*^{D,A}$ — hence, uniformly continuous. Let $\widehat{\varphi}_\lambda([G,o,W]) = \widehat{\varphi}(\lambda,[G,o,W])$. Local convergence means that the measures $U_{(G_n,W_n)}$ (defined in (A.2)) converge weakly to $\mathbb{P}$. Thus, for any $\lambda \in I$, $\int \widehat{\varphi}_\lambda\,\mathrm{d}U_{(G_n,W_n)} \to \int \widehat{\varphi}_\lambda\,\mathrm{d}\rho$, i.e., $\frac{1}{|V_n|}\sum_{x\in V_n} \widehat{\varphi}_\lambda([G_n,x,W_n]) \to \int \widehat{\varphi}_\lambda([G,o,W])\,\mathrm{d}\,\mathbb{P}([G,o,W])$, which is the statement of the lemma for fixed $\lambda \in I$.

Uniformity in λ comes from the uniform continuity of $\widehat{\varphi}$, which implies that the maps $\lambda \mapsto \int \widehat{\varphi}_\lambda\,\mathrm{d}U_{(G_n,W_n)}$ form a uniformly equicontinuous family. □

Remark A.3. Taking $F \equiv 1$ we get, in particular, the convergence of empirical spectral measures. On the other hand, when $\varphi \equiv 1$ we get, in particular, that under assumption (BSCT), if $I \subset \mathbb{R}$ is compact and $\eta \in (0,1)$ is fixed, then uniformly in $\lambda \in I$,

(A.8)
$$\frac{1}{N}\sum_{(x_0;x_s)\in B_s} F\left(\zeta_{x_0}^{\lambda+i\eta}(x_1), \zeta_{x_1}^{\lambda+i\eta}(x_0), \dots, \zeta_{x_{s-1}}^{\lambda+i\eta}(x_s), \zeta_{x_s}^{\lambda+i\eta}(x_{s-1})\right)$$
$$\underset{N\longrightarrow+\infty}{\longrightarrow} \mathbb{E}\left[\sum_{(v_0;v_s)\in B_s : v_0=o} F\left(\hat{\zeta}_{v_0}^{\lambda+i\eta}(v_1), \hat{\zeta}_{v_1}^{\lambda+i\eta}(v_0), \dots, \hat{\zeta}_{v_{s-1}}^{\lambda+i\eta}(v_s), \hat{\zeta}_{v_s}^{\lambda+i\eta}(v_{s-1})\right)\right].$$

In this article, we often encounter expressions of the form

$$\vartheta_\gamma(x_0,x_1) = F(\zeta_{x_0}^\gamma(x_1), \zeta_{x_1}^\gamma(x_0))$$

in the left-hand side of (A.8). In this case, we write

$$\hat{\vartheta}_\gamma(v_0,v_1) := F(\hat{\zeta}_{v_0}^\gamma(v_1), \hat{\zeta}_{v_1}^\gamma(v_0))$$

for the object defined similarly at the limit. For instance, $\hat{\mu}_1^\gamma$ is defined like μ_1^γ but on the limiting tree $(\mathcal{T},\mathcal{W})$. In the particular case of m^γ, we have $\hat{m}_o^\gamma = \frac{-1}{2\mathcal{G}^\gamma(o,o)}$.

It is worth noting that $\mathbb{E}[\sum_{o'\sim o} F(\hat{\zeta}_o^\gamma(o'))] = \mathbb{E}[\sum_{o'\sim o} F(\hat{\zeta}_{o'}^\gamma(o))]$. This holds because $\frac{1}{N}\sum_{(x_0,x_1)} F(\zeta_{x_0}^\gamma(x_1)) = \frac{1}{N}\sum_{(x_0,x_1)} F(\zeta_{x_1}^\gamma(x_0))$.

Remark A.4. Using (2.4b), we have $|\hat{\zeta}_{o'}^\gamma(o)|^s \le |\operatorname{Im}\hat{\zeta}_o^\gamma(u)|^{-s}$ for any $u \in \mathcal{N}_o \setminus \{o'\}$. In particular, $|\hat{\zeta}_{o'}^\gamma(o)|^s \le \sum_{o''\sim o} |\operatorname{Im}\hat{\zeta}_o^\gamma(o'')|^{-s}$. We thus see by (Green) that for any $s > 0$,

(A.9)
$$\sup_{\lambda\in I_1,\eta\in(0,1)} \mathbb{E}(|\operatorname{Im}\mathcal{G}^{\lambda+i\eta}(o,o)|^{-s}) < \infty, \qquad \sup_{\lambda\in I_1,\eta\in(0,1)} \mathbb{E}(|\mathcal{G}^{\lambda+i\eta}(o,o)|^{s}) < \infty,$$

(A.10)
$$\sup_{\lambda\in I_1,\eta\in(0,1)} \mathbb{E}\Big(\sum_{y\sim o} |\hat{\zeta}_y^{\lambda+i\eta}(o)|^s\Big) < \infty, \qquad \sup_{\lambda\in I_1,\eta\in(0,1)} \mathbb{E}\Big(\sum_{y\sim o} |\hat{\zeta}_o^{\lambda+i\eta}(y)|^s\Big) < \infty,$$

$$\sup_{\lambda\in I_1,\eta\in(0,1)} \mathbb{E}\Big(\sum_{y\sim o}|\operatorname{Im}\hat{\zeta}_y^{\lambda+i\eta}(o)|^{-s}\Big)<\infty.$$

We also have

$$\sup_{\lambda\in I_1,\eta\in(0,1)} \mathbb{E}\left[\sum_{(v_0;v_t)\in B_t:v_0=o}\left|\hat{\zeta}_{v_0}^{\lambda+i\eta}(v_1)\hat{\zeta}_{v_1}^{\lambda+i\eta}(v_0)\cdots\hat{\zeta}_{v_{t-1}}^{\lambda+i\eta}(v_t)\hat{\zeta}_{v_t}^{\lambda+i\eta}(v_{t-1})\right|^s\right]<\infty.$$

To see this, consider for simplicity $\mathbb{E}[\sum_{(v_0;v_2),v_0=o}|\hat{\zeta}_{v_0}^{\gamma}(v_1)\hat{\zeta}_{v_1}^{\gamma}(v_2)|^s]$. This is the limit of $\frac{1}{N}\sum_{(x_0;x_2)\in B_2}|\zeta_{x_0}^{\gamma}(x_1)\zeta_{x_1}^{\gamma}(x_2)|^s$. This sum is bounded by

$$\left(\frac{1}{N}\sum_{(x_0;x_2)\in B_2}|\zeta_{x_0}^{\gamma}(x_1)|^{2s}\right)^{1/2}\cdot\left(\frac{1}{N}\sum_{(x_0;x_2)\in B_2}|\zeta_{x_1}^{\gamma}(x_2)|^{2s}\right)^{1/2}$$

for any N. Using $|\mathcal{N}_{x_1}|-1\le D$ and taking $N\to\infty$, we see the limit is bounded by

$$D\,\mathbb{E}\left(\sum_{o'\sim o}|\hat{\zeta}_o^{\gamma}(o')|^{2s}\right)^{1/2}\mathbb{E}\left(\sum_{o'\sim o}|\hat{\zeta}_o^{\gamma}(o')|^{2s}\right)^{1/2}\le DC_s$$

by (A.10), for any $\lambda\in I_1$ and $\eta>0$. Hence,

$$\sup_{\lambda\in I_1,\eta>0}\mathbb{E}\left[\sum_{(v_0;v_2),v_0=o}|\hat{\zeta}_{v_0}^{\gamma}(v_1)\hat{\zeta}_{v_1}^{\gamma}(v_2)|^s\right]\le DC_s.$$

Remark A.5. Let us now look at the quantity

$$\frac{1}{N}\sum_{(x_0,x_1)}\sum_{(x_2;x_k),(y_2;y_k)}|\tilde{g}^{\gamma}(\tilde{x}_k,\tilde{y}_k)|^s,$$

which we had to control in Section 4.

Let $x_k\wedge y_k$ be the vertex of maximal length in $(x_0;x_k)\cap(x_0;y_k)$, so $x_k\wedge y_k=x_t$ for some $1\le t\le k$. Then

$$\tilde{g}^{\gamma}(\tilde{x}_k,\tilde{y}_k)=\frac{-\prod_{l=0}^{k-t-1}\zeta_{x_{k-l}}^{\gamma}(x_{k-l-1})\cdot\zeta_{x_t}^{\gamma}(y_{t+1})\prod_{l=t+1}^{k-1}\zeta_{y_l}^{\gamma}(y_{l+1})}{2m_{x_k}^{\gamma}}.$$

We then write

$$\frac{1}{N}\sum_{(x_0,x_1)}\sum_{(x_2;x_k),(y_2;y_k)}=\frac{1}{N}\sum_{(x_0,x_1)}\sum_{t=1}^{k}\sum_{(x_2;x_k),(y_2;y_k),x_k\wedge y_k=x_t},$$

use Hölder's inequality, and take $N\to\infty$ to get a uniform bound involving $\mathbb{E}[\sum_{o'\sim o}|\hat{\zeta}_o^{\gamma}(o')|^{s_2}]$ and $\mathbb{E}[|2\hat{m}_o|^{-s_1}]$, both of which are finite. Hence,

$$\frac{1}{N}\sum_{(x_0,x_1)}\sum_{(x_2;x_k),(y_2;y_k)}|\tilde{g}^{\gamma}(\tilde{x}_k,\tilde{y}_k)|^s$$

is uniformly bounded as $N\to\infty$.

A.3. *Proofs of auxiliary results.* We now turn to the proofs of some claims in Section 1. In what follows, $\eta_0 \in (0,1)$ is fixed.

Claim (1.8). Let $\chi : \mathscr{G}_*^{D,A} \to \mathbb{R}$ and $F : \mathbb{C} \to \mathbb{R}$ be continuous. Then under (BSCT),

$$\begin{aligned}(\mathrm{A.11}) \qquad &\frac{1}{N}\sum_{x\in V_N}\chi([G_N,x])\sum_{y,d(y,x)=k}F(\tilde{g}_N^{\lambda+i\eta_0}(\tilde{x},\tilde{y}))\\ &\underset{N\longrightarrow+\infty}{\longrightarrow}\mathbb{E}\Big(\chi([\mathcal{T},o])\sum_{v,d(v,o)=k}F(\mathcal{G}^{\lambda+i\eta_0}(o,v))\Big)\end{aligned}$$

uniformly in $\lambda \in I_0$. This is a variant of Corollary A.2 when one considers $F_{\gamma,\chi}$: $(\lambda,[G,x,W]) \mapsto \chi([G,x])\sum_{y,d(y,x)=k}F(\tilde{g}^{\gamma}(x,y))$ instead of F_γ. In particular, taking $k=0$ and $\chi=1$, we obtain (1.8).

Claim (1.9). We may assume F is compactly supported (cf. Lemma A.1), hence uniformly continuous. Let

$$\begin{aligned}h_N(t) &= \frac{1}{N}\sum_{x\in V_N}\chi([G_N,x])\sum_{y,d(y,x)=k}F(t\,\mathrm{Im}\,\tilde{g}_N^{\lambda+i\eta_0}(x,y)),\\ h(t) &= \mathbb{E}\Big(\chi([\mathcal{T},o])\sum_{v,d(v,o)=k}F(t\,\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,v))\Big),\end{aligned}$$

and let

$$c_N(\lambda) = \frac{N}{\sum_{\tilde{x}\in\mathcal{D}_N}\mathrm{Im}\,\tilde{g}_N^{\lambda+i\eta_0}(\tilde{x},\tilde{x})} \quad \text{and} \quad c(\lambda) = \frac{1}{\mathbb{E}(\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o))}.$$

The family h_N is uniformly equicontinuous, and as in (A.11) it converges uniformly to h. By (1.8), $c_N(\lambda) \to c(\lambda)$ uniformly in λ. So $|h_N(c_N(\lambda)) - h(c(\lambda))| \to 0$ uniformly in λ. This proves (1.9).

We now turn to the proof of Claim (1.7). Consider the set of (double)-colored rooted graphs (G,o,W,a), where now $W : V \longrightarrow \mathbb{R}$ and $a : V \to \{0,1\}$. We say (G,o,W,a) and (G',o',W',a') are equivalent if there is $\phi : G \to G'$ with $\phi(o)=o'$, $W'\circ\phi = W$ and $a'\circ\phi = a$. We let $\widehat{\mathscr{G}}_*^{D,A}$ be the corresponding set of equivalence classes and endow it with a metric d_{loc} defined similarly to (A.1). This amounts to the same definition as before, except that the colorings now take values in $\mathbb{R}\times\{0,1\}$ instead of $\mathbb{R}$. The notion of local weak limit may obviously be extended to this situation.

Assuming that (BSCT) holds, then up to passing to a subsequence,

$$(G_N, W_N, \mathbb{1}_{\Lambda_N})$$

will have a local weak limit $\hat{\mathbb{P}}$ concentrated on $\{[\mathcal{T},o,\mathcal{W},a]\}$, whose marginals on $\mathscr{T}_*^{D,A}$ coincides with $\mathbb{P}$. The fact that $|\Lambda_N| \ge \alpha N$ implies $\hat{\mathbb{P}}(a(o)=1) \ge \alpha$, since $\{a(o)=1\}$ is clopen in $\widehat{\mathscr{G}}_*^{D,A}$. We claim that

$$\lim_{N\longrightarrow+\infty}\langle \mathbb{1}_{\Lambda_N}\rangle_{\lambda+i\eta_0}=\frac{\hat{\mathbb{E}}\left(a(o)\,\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\right)}{\mathbb{E}\left(\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\right)} \tag{A.12}$$

uniformly in $\lambda\in I_0$. Indeed, as in Lemma A.1, if $F: I_0\times\widehat{\mathscr{G}}_*^{D,A}\to\mathbb{C}$ is given by $F(\lambda,[G,x,W,a])=a(x)\,\mathrm{Im}\,\tilde{g}^{\lambda+i\eta_0}(x,x)$, then F is continuous. So $\int F_\lambda\,\mathrm{d}U_{G_N,W_N,\mathbb{1}_{\Lambda_N}}\to\int F_\lambda\,\mathrm{d}\hat{\mathbb{P}}$ uniformly in λ as in Corollary A.2. Combined with (1.8), this yields (A.12). We next note that for any $\alpha>0$,

$$\inf_{\lambda\in I_1,\eta_0\in(0,1)}\ \inf_{a,\hat{\mathbb{P}}(a(o)=1)\geq\alpha}\frac{\hat{\mathbb{E}}\left(a(o)\,\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\right)}{\mathbb{E}\left(\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\right)}>0. \tag{A.13}$$

In fact, suppose on the contrary that for all $\epsilon>0$, we can find $\lambda\in I_1,\eta_0\in(0,1)$ and a such that $\hat{\mathbb{P}}(a(o)=1)\geq\alpha$ and $\hat{\mathbb{E}}\left(a(o)\,\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\right)\leq\epsilon$. The latter implies

$$\hat{\mathbb{P}}\left(a(o)=1,\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\geq\epsilon^{1/2}\right)\leq\epsilon^{1/2}.$$

On the other hand, since a takes only the values 0 and 1,

$$\hat{\mathbb{P}}\left(a(o)=1,\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\geq\epsilon^{1/2}\right)\geq\hat{\mathbb{P}}(\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\geq\epsilon^{1/2})-\hat{\mathbb{P}}(a(o)=0).$$

Thus,

$$\hat{\mathbb{P}}(\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\geq\epsilon^{1/2})-\hat{\mathbb{P}}(a(o)=0)\leq\epsilon^{1/2}.$$

Equation (A.9) with $s=2$ implies that $\hat{\mathbb{P}}(\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)<\epsilon^{1/2})\leq C\epsilon$ for some constant $C<\infty$ independent of λ,η_0. So $\hat{\mathbb{P}}(\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\geq\epsilon^{1/2})\geq 1-C\epsilon$. By assumption, $\hat{\mathbb{P}}(a(o)=0)\leq 1-\alpha$. Taking $\epsilon\to 0$ we would obtain $\alpha\leq 0$, a contradiction. We thus proved (A.13). Since (A.12) holds uniformly in λ, we get (1.7).

Finally, as in the proof of (A.12), we may consider the set of double-colored rooted graphs (G,o,W,K), where K is a coloring of pairs of vertices $x,y\in G$, $d_G(x,y)\leq R$, with values in $\{|z|\leq 1\}\subset\mathbb{C}$. Assuming (BSCT) holds, up to passing to a subsequence, (G_N,W_N,K_N) will have a local weak limit $\hat{\mathbb{P}}$ concentrated on $\{[\mathcal{T},o,\mathcal{W},\mathcal{K}]\}$ whose marginals on $\mathscr{T}_*^{D,A}$ coincides with $\mathbb{P}$. We then deduce as before that uniformly in $\lambda\in I_0$,

$$\lim_{N\longrightarrow+\infty}\langle\mathbf{K}_N\rangle_{\lambda+i\eta_0}=\frac{\hat{\mathbb{E}}\left(\sum_{y:d(y,o)\leq R}\mathcal{K}(o,y)\,\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,y)\right)}{\mathbb{E}\left(\mathrm{Im}\,\mathcal{G}^{\lambda+i\eta_0}(o,o)\right)}. \tag{A.14}$$

Acknowledgements. This material is based upon work supported by the Agence Nationale de la Recherche under grant No.ANR-13-BS01-0007-01, by the Labex IRMIA and the Institute of Advance Study of Université de Strasbourg, and by Institut Universitaire de France.

References

[1] M. Abért, A. Thom, and B. Virág, Benjamini-Schramm convergence and pointwise convergence of the spectral measure, author homepage.

[2] M. Aizenman, M. Shamis, and S. Warzel, Resonances and partial delocalization on the complete graph, *Ann. Henri Poincaré* **16** no. 9 (2015), 1969–2003. MR 3383321. Zbl 1329.81192. https://doi.org/10.1007/s00023-014-0366-9.

[3] M. Aizenman and S. Warzel, Absolutely continuous spectrum implies ballistic transport for quantum particles in a random potential on tree graphs, *J. Math. Phys.* **53** no. 9 (2012), 095205, 15. MR 2905787. Zbl 1278.81090. https://doi.org/10.1063/1.4714617.

[4] D. Aldous and R. Lyons, Processes on unimodular random networks, *Electron. J. Probab.* **12** (2007), no. 54, 1454–1508. MR 2354165. Zbl 1131.60003. https://doi.org/10.1214/EJP.v12-463.

[5] N. Anantharaman, Quantum ergodicity on regular graphs, *Comm. Math. Phys.* **353** no. 2 (2017), 633–690. MR 3649482. Zbl 1368.58015. https://doi.org/10.1007/s00220-017-2879-9.

[6] N. Anantharaman, Some relations between the spectra of simple and non-backtracking random walks, 2017. arXiv 1703.03852.

[7] N. Anantharaman and E. Le Masson, Quantum ergodicity on large regular graphs, *Duke Math. J.* **164** no. 4 (2015), 723–765. MR 3322309. Zbl 1386.58015. https://doi.org/10.1215/00127094-2881592.

[8] N. Anantharaman and M. Sabri, Quantum ergodicity for the Anderson model on regular graphs, *J. Math. Phys.* **58** no. 9 (2017), 091901, 10. MR 3707062. Zbl 1376.82091. https://doi.org/10.1063/1.5000962.

[9] N. Anantharaman and M. Sabri, Poisson kernel expansions for Schrödinger operators on trees, *J. Spectr. Theory* **9** no. 1 (2019), 243–268. MR 3900786. https://doi.org/10.4171/JST/247.

[10] A. Backhausz and B. Szegedy, On the almost eigenvectors of random regular graphs, 2016, to appear in *Ann. Prob.* arXiv 1607.04785.

[11] R. Bauerschmidt, J. Huang, A. Knowles, and H.-T. Yau, Bulk eigenvalue statistics for random regular graphs, *Ann. Probab.* **45** no. 6A (2017), 3626–3663. MR 3729611. Zbl 1379.05098. https://doi.org/10.1214/16-AOP1145.

[12] R. Bauerschmidt, J. Huang, and H. Yau, Local Kesten–McKay law for random regular graphs, *Comm. Math. Phys.*, published online 28 February 2019. https://doi.org/10.1007/s00220-019-03345-3.

[13] R. Bauerschmidt, A. Knowles, and H.-T. Yau, Local semicircle law for random regular graphs, *Comm. Pure Appl. Math.* **70** no. 10 (2017), 1898–1960. MR 3688032. Zbl 1372.05194. https://doi.org/10.1002/cpa.21709.

[14] I. Benjamini, *Coarse Geometry and Randomness*, *Lecture Notes in Math.* **2100**, Springer, Cham, 2013, lecture notes from the 41st Probability Summer School held in Saint-Flour, 2011, Chapter 5 is due to Nicolas Curien, Chapter 12 was written by Ariel Yadin, and Chapter 13 is joint work with Gady Kozma, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. MR 3156647. Zbl 1282.05001. https://doi.org/10.1007/978-3-319-02576-6.

[15] I. Benjamini and O. Schramm, Recurrence of distributional limits of finite planar graphs, *Electron. J. Probab.* **6** (2001), no. 23, 13. MR 1873300. Zbl 1010.82021. https://doi.org/10.1214/EJP.v6-96.

[16] M. Berry and M. Tabor, Level clustering in the regular spectrum, *Proc. Royal Soc. A* **356** (1977), 375–394. Zbl 1119.81395. https://doi.org/10.1098/rspa.1977.0140.

[17] O. Bohigas, M.-J. Giannoni, and C. Schmit, Characterization of chaotic quantum spectra and universality of level fluctuation laws, *Phys. Rev. Lett.* **52** no. 1 (1984), 1–4. MR 0730191. Zbl 1119.81326. https://doi.org/10.1103/PhysRevLett.52.1.

[18] O. Bohigas, M.-J. Giannoni, and C. Schmit, Spectral fluctuations, random matrix theories and chaotic motion, in *Stochastic Processes in Classical and Quantum Systems* (Ascona, 1985), *Lecture Notes in Phys.* **262**, Springer, Berlin, 1986, pp. 118–138. MR 0870168. https://doi.org/10.1007/3540171665_59.

[19] C. Bordenave, On quantum percolation in finite regular graphs, *Ann. Henri Poincaré* **16** no. 11 (2015), 2465–2497. MR 3411739. Zbl 1332.82054. https://doi.org/10.1007/s00023-014-0382-9.

[20] P. Bourgade and H.-T. Yau, The eigenvector moment flow and local quantum unique ergodicity, *Comm. Math. Phys.* **350** no. 1 (2017), 231–278. MR 3606475. Zbl 1379.58014. https://doi.org/10.1007/s00220-016-2627-6.

[21] S. Brooks, E. Le Masson, and E. Lindenstrauss, Quantum ergodicity and averaging operators on the sphere, *Int. Math. Res. Not. IMRN* no. 19 (2016), 6034–6064. MR 3567266. Zbl 1404.35307. https://doi.org/10.1093/imrn/rnv337.

[22] S. Brooks and E. Lindenstrauss, Non-localization of eigenfunctions on large regular graphs, *Israel J. Math.* **193** no. 1 (2013), 1–14. MR 3038543. Zbl 1317.05110. https://doi.org/10.1007/s11856-012-0096-y.

[23] A. De Luca, B. L. Altshuler, V. E. Kravtsov, and A. Scardicchio, Support set of random wave-functions on the Bethe lattice, 2013, preprint. arXiv 1401.0019.

[24] A. De Luca, B. L. Altshuler, V. E. Kravtsov, and A. Scardicchio, Anderson localization on the Bethe lattice: Nonergodicity of extended states, *Phys. Rev. Lett.* **113** (2014), 046806. https://doi.org/10.1103/PhysRevLett.113.046806.

[25] P. Diaconis and D. Stroock, Geometric bounds for eigenvalues of Markov chains, *Ann. Appl. Probab.* **1** no. 1 (1991), 36–61. MR 1097463. Zbl 0731.60061. https://doi.org/10.1214/aoap/1177005980.

[26] I. Dumitriu and S. Pal, Sparse regular random graphs: spectral density and eigenvectors, *Ann. Probab.* **40** no. 5 (2012), 2197–2235. MR 3025715. Zbl 1255.05173. https://doi.org/10.1214/11-AOP673.

[27] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, Spectral statistics of Erdős-Rényi graphs I: Local semicircle law, *Ann. Probab.* **41** no. 3B (2013), 2279–2375. MR 3098073. Zbl 1272.05111. https://doi.org/10.1214/11-AOP734.

[28] L. Erdős, B. Schlein, and H.-T. Yau, Local semicircle law and complete delocalization for Wigner random matrices, *Comm. Math. Phys.* **287**

no. 2 (2009), 641–655. MR 2481753. Zbl 1186.60005. https://doi.org/10.1007/s00220-008-0636-9.

[29] L. Erdős, B. Schlein, and H.-T. Yau, Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices, *Ann. Probab.* **37** no. 3 (2009), 815–852. MR 2537522. Zbl 1175.15028. https://doi.org/10.1214/08-AOP421.

[30] L. Geisinger, Convergence of the density of states and delocalization of eigenvectors on random regular graphs, *J. Spectr. Theory* **5** no. 4 (2015), 783–827. MR 3433288. Zbl 1384.60024. https://doi.org/10.4171/JST/114.

[31] J. P. Keating, Quantum graphs and quantum chaos, in *Analysis on Graphs and its Applications*, *Proc. Sympos. Pure Math.* **77**, Amer. Math. Soc., Providence, RI, 2008, pp. 279–290. MR 2459875. Zbl 1153.81501. https://doi.org/10.1090/pspum/077/2459875.

[32] A. Klein, Extended states in the Anderson model on the Bethe lattice, *Adv. Math.* **133** no. 1 (1998), 163–184. MR 1492789. Zbl 0899.60088. https://doi.org/10.1006/aima.1997.1688.

[33] A. Klenke, *Probability Theory. A Comprehensive Course*, *Universitext*, Springer, London, 2014. MR 3112259. Zbl 1295.60001. https://doi.org/10.1007/978-1-4471-5361-0.

[34] J. Marklof, Pair correlation densities of inhomogeneous quadratic forms, *Ann. of Math.* (2) **158** no. 2 (2003), 419–471. MR 2018926. Zbl 1106.11018. https://doi.org/10.4007/annals.2003.158.419.

[35] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. MR 1777382. Zbl 0962.15001.

[36] W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, *Astérisque* **187-188**, Math. Soc. France, 1990. MR 1085356. Zbl 0726.58003.

[37] J. Salez, Some implications of local weak convergence for sparse random graphs, Hal Id: tel-00637130.

[38] P. Sarnak, Arithmetic quantum chaos, in *The Schur Lectures* (1992) (Tel Aviv)), *Israel Math. Conf. Proc.* **8**, Bar-Ilan Univ., Ramat Gan, 1995, pp. 183–236. MR 1321639. Zbl 0831.58045.

[39] P. Sarnak, Values at integers of binary quadratic forms, in *Harmonic Analysis and Number Theory* (Montreal, PQ, 1996), *CMS Conf. Proc.* **21**, Amer. Math. Soc., Providence, RI, 1997, pp. 181–203. MR 1472786. Zbl 0911.11032.

[40] C. Schumacher and F. Schwarzenberger, Approximation of the integrated density of states on sofic groups, *Ann. Henri Poincaré* **16** no. 4 (2015), 1067–1101. MR 3317792. Zbl 1316.82015. https://doi.org/10.1007/s00023-014-0342-4.

[41] U. Smilansky, Quantum chaos on discrete graphs, *J. Phys. A* **40** no. 27 (2007), F621–F630. MR 2369953. Zbl 1124.81024. https://doi.org/10.1088/1751-8113/40/27/F07.

[42] U. Smilansky, Discrete graphs—a paradigm model for quantum chaos, in *Chaos* (Poincaré seminar 2010. Proceedings, Paris, France, June 5, 2010), *Prog. Math.*

Phys. **66**, Birkhäuser/Springer, Basel, 2013, pp. 97–124. MR 3204183. Zbl 1319.81047. https://doi.org/10.1007/978-3-0348-0697-8_3.

[43] K. S. Tikhonov, A. D. Mirlin, and M. A. Skvortsov, Anderson localization and ergodicity on random regular graphs, *Phys. Rev. B* **94** no. 22 (2016), 220203. https://doi.org/10.1103/PhysRevB.94.220203.

[44] L. V. Tran, V. H. Vu, and K. Wang, Sparse random graphs: eigenvalues and eigenvectors, *Random Structures Algorithms* **42** no. 1 (2013), 110–134. MR 2999215. Zbl 1257.05089. https://doi.org/10.1002/rsa.20406.

[45] Y. Colin de Verdière, Ergodicité et fonctions propres du laplacien, *Comm. Math. Phys.* **102** no. 3 (1985), 497–502. MR 0818831. Zbl 0592.58050. https://doi.org/10.1007/BF01209296.

[46] A. I. Šnirel′man, Ergodic properties of eigenfunctions, *Uspehi Mat. Nauk* **29** no. 6(180) (1974), 181–182. MR 0402834. Zbl 0324.58020. Available at http://mi.mathnet.ru/eng/umn/v29/i6/p181.

[47] S. Zelditch, Uniform distribution of eigenfunctions on compact hyperbolic surfaces, *Duke Math. J.* **55** no. 4 (1987), 919–941. MR 0916129. Zbl 0643.58029. https://doi.org/10.1215/S0012-7094-87-05546-3.

[48] S. Zelditch, Quantum ergodicity of C^* dynamical systems, *Comm. Math. Phys.* **177** no. 2 (1996), 507–528. MR 1384146. Zbl 0856.58019. https://doi.org/10.1007/BF02101904.

(Received: April 3, 2017)
(Revised: November 6, 2018)

Université de Strasbourg, CNRS, IRMA UMR 7501, Strasbourg, France
E-mail: anantharaman@math.unistra.fr

Department of Mathematics, Faculty of Science, Cairo University, Cairo, Egypt and Université de Strasbourg, CNRS, IRMA UMR 7501, Strasbourg, France
E-mail: mmsabri@sci.cu.edu.eg