Convergence of the density of states and delocalization of eigenvectors on random regular graphs

Leander Geisinger

Abstract. Consider a random regular graph of fixed degree d with n vertices. We study spectral properties of the adjacency matrix and of random Schrödinger operators on such a graph as n tends to infinity.

We prove that the integrated density of states on the graph converges to the integrated density of states on the infinite regular tree and we give uniform bounds on the rate of convergence. This allows to estimate the number of eigenvalues in intervals of size comparable to $\log_{d-1}^{-1}(n)$. Based on related estimates for the Green function we derive results about delocalization of eigenvectors.

Contents

1. Introduction

In his seminal work from 1955, Wigner showed that the density of states of random matrices with independent identically distributed entries converges – as the size of the matrix tends to infinity – to a universal deterministic probability measure, the semicircle law [49]. This result was gradually improved and today it is known that universality of spectral properties of Wigner matrices goes far beyond that. Even the local eigenvalue statistics, studied via the eigenvalue gap distribution, is given by universal laws [20–22, 44]. These laws can be calculated explicitly from Gaussian ensembles and are characterized by local repulsion of the eigenvalues. We refer to the textbooks [7, 8, 36] and the review [13] for general results about random matrices and further references.

One field where random matrices arise is the study of random graphs. Along with a graph of n labeled vertices one considers the adjacency matrix which is the symmetric $n \times n$ matrix with entry ij equal to 1, if vertex i is connected by an edge to vertex j, and equal to 0 otherwise. The spectrum of this matrix bears information about the geometry of the graph, however determining spectral properties is often a difficult task.

For Erdős–Rényi graphs, where every edge is chosen independently with probability p (see [10] for details concerning random graph models), universality of the local eigenvalue statistics for the corresponding adjacency matrix has recently been proved under suitable conditions on p [18,19]. Two of the steps towards universality are proving local convergence of the density of states and delocalization of eigenvectors. The former refers to the fact that the average number of eigenvalues in small intervals converges to a universal law. (Ideally this holds for intervals of size comparable to $1/n$, up to logarithmic corrections.) The latter means that eigenvectors are typically uniformly distributed over the whole graph. Let us emphasize that the mentioned results, for Wigner matrices as well as for Erdős–Rényi graphs, rely on independence of the entries.

In this note we investigate the spectrum on random regular graphs with fixed degree d. In a regular graph of degree d each vertex is connected to d other vertices. Thus in the corresponding adjacency matrix each row and each column contains d entries that are 1 and $n - d$ entries that are 0. Such a graph, or equivalently such a matrix, is chosen at random with uniform probability. These random matrices are sparse: only few entries are non-zero so that moments of the distribution of the entries decay slowly. More importantly the entries lack independence. Therefore it is not clear how to apply the methods developed to study Wigner matrices and Erdős–Rényi graphs.

While there are results concerning extreme eigenvalues at the edge of the spectrum [24, 42] not much seems to been known about eigenvalues in the bulk of the spectrum. However, it is conjectured that the local statistics is universal and governed by repulsion of eigenvalues, see for example [17, 18, 27, 37].

In this article a small step is made in this direction by studying convergence of the density of states and delocalization of eigenvectors. First we analyze the density of states. The analogue of Wigner's result for random regular graphs was proved by McKay [33]: The density of states of the adjacency matrix of a random regular graph converges in distribution to a probability measure, known as the Kesten–McKay law.

In Theorem 2.1 we refine this result by giving uniform bounds on the rate of convergence. This allows us to deduce local convergence of the number of eigenvalues in intervals of size comparable to $\log_{d-1}^{-1}(n)$. Our approach is based on the fact that a random regular graph coincides locally with a regular tree. This explains the rate $\log_{d-1}^{-1}(n)$ since this approximation typically works well up to distances comparable to $\log_{d-1}(n)$. While this rate is far from the desired $1/n$, our results are strong enough to deduce statements about convergence of the Green function and delocalization of eigenvectors.

Our findings are similar to results recently obtained by Dumitriu and Pal [16] and by Tran, Vu, and Wang [46]. They also study spectral properties of random regular graphs and consider the case where the degree d is not fixed but tends to infinity together with the number of vertices n. The main results also include delocalization of eigenvectors and local convergence of the number of eigenvalues. Again the size of the allowed intervals is comparable to $\log_{d_n-1}^{-1}(n)$. Let us remark that our methods – even though they are tailored for regular graphs with fixed degree – also extend to graphs with growing degree. In Theorem 2.2 we recover results from [16] and [46] and we slightly improve them in a certain sense that is made precise below.

After this prelude, the main purpose of this article is to derive similar results for random Schrödinger operators and to deduce delocalization of eigenvectors. Motivated by recent physical and numerical considerations about eigenvalue statistics [9, 14] we study the Anderson model of random Schrödinger operators on random regular graphs: A random Schrödinger operator consists of the adjacency matrix perturbed by a random potential, that means one adds independent identically distributed entries on the diagonal of the matrix. Again we explore the density of states by comparing with the Anderson model on the infinite regular tree. The strength of our approach lies in the fact that it depends only on local

properties of the graph, therefore it extends to general local operators such as random Schrödinger operators.

The Anderson model on the tree is one of the most studied models of random Schrödinger operators, see [47] for an overview of results and references. It is a natural question in what way spectral properties on the infinite tree extend to the corresponding finite-volume operator on a random regular graph. However, for Schrödinger operators the analysis of spectrum and eigenvectors is more challenging since the corresponding spectral measure on the infinite tree is not purely absolutely continuous but it also contains a pure-point component.

This should influence the behavior of eigenvectors of the finite-volume operator. In spectral regimes that correspond to pure-point spectrum of the infinite-volume operator one expects to find exponential localization while eigenvectors with eigenvalues within the absolutely continuous spectrum are expected to be delocalized. In turn, properties of eigenvectors are conjectured to determine whether the local eigenvalue statistics is Poisson or governed by level repulsion [9,14,47].

Again we are able to make a small step in this direction. In Theorem 2.3 we show that the mean density of states of a random Schrödinger operator on a random regular graph converges to the density of states on the infinite tree. We give bounds on the rate of convergence and show that the mean number of eigenvalues in intervals of size comparable to $\log_{d-1}^{-1}(n)$ converges. Finally, in Theorem 5.3 we apply these results and combine them with recent results about random Schrödinger operators on the infinite tree [3] to prove that typically eigenvectors on a random regular graph with eigenvalues within the absolutely continuous spectrum of the infinite-volume operator are not localized.

The article is organized as follows. In the next section we explain the relevant notation concerning random regular graphs and spectral measures. Then we give the precise statements of the main results about convergence of the density of states.

In Section 3 we provide the main tools: The fact that a random regular graph coincides locally with a tree implies that low moments of the spectral measure on the graph agree with the respective moments on the tree. Therefore also the expectation of polynomials of low degree (typically up to $\log_{d-1}(n)$) is the same. Thus to obtain uniform estimates on the rate of convergence of the density of states one needs to approximate the Heaviside function by polynomials. As was noticed by Chebyshev, Markov, and Stieltjes orthogonal polynomials are well suited for this purpose. In Theorem 3.1 we use this approach to prove a deterministic estimate for the density of states valid for all regular graphs.

In Section 4 we study the distribution of cycles in a random regular graph and we show that it is justified to approximate a random regular graph locally by a tree. We use that to complete the proof of the results from Section 2.

In the final section we apply the developed techniques to prove convergence of the Green function and to deduce results about delocalization of eigenvectors.

2. Notation and results

Consider the set $\mathcal{G}_{n,d}$ of simple regular graphs with n vertices and degree $d \geq 3$. Let $\mathcal{P}_{n,d}$ denote the uniform probability measure on this set and let $\mathcal{E}_{n,d}$ denote the expectation with respect to $\mathcal{P}_{n,d}$. We study this ensemble as n tends to infinity. We say that an event happens asymptotically almost surely if the probability $\mathcal{P}_{n,d}$ of the event tends to one as n tends to infinity.

We assume that the degree d is fixed unless stated otherwise. However, the methods that are employed here are not limited to this case. In particular in Theorem 2.2 we consider the case where $d = d_n$ tends to infinity with n.

2.1. The adjacency matrix. First we study the spectrum of the adjacency matrix A_n of a random regular graph $G_n \in \mathcal{G}_{n,d}$. The adjacency matrix is the self-adjoint operator on the Hilbert space $\ell^2(G_n)$ defined by

$$(A_n \phi)(x) = \sum_{y \in G_n : d(x,y)=1} \phi(y), \quad \phi \in \ell^2(G_n), \; x \in G_n.$$

Here the distance $d(x, y)$ of two vertices $x, y \in G_n$ is the length of the shortest path connecting x and y. The adjacency matrix is the discrete Laplace operator on G_n with the diagonal terms removed.

Let $(\lambda_j)_{j=1}^n$ denote the eigenvalues of A_n and let $(\varphi_j)_{j=1}^n$ be the corresponding $\ell^2(G_n)$-normalized eigenvectors. If an eigenvalue has multiplicity higher than one we repeat the value according to its multiplicity and we choose the eigenvectors as an orthonormal basis of the eigenspace. Since G_n has n vertices, A_n is a symmetric n by n matrix and we obtain n eigenvalues and n eigenvectors. To study the distribution of the eigenvalues we introduce the following spectral measures.

For a vertex $x \in G_n$ we write $\delta_x \in \ell^2(G_n)$ for its characteristic function: $\delta_x(x) = 1$ and $\delta_x(y) = 0$ for $y \neq x$. Also, for a set $I \subset \mathbb{R}$ let χ_I denote its characteristic function: $\chi_I(t) = 1$ for $t \in I$ and $\chi_I(t) = 0$ for $t \notin I$. We write $|I|$ for the length of the set. The local spectral measure $\mu_{n,x}$ with respect to a vertex

$x \in G_n$ is given by

$$\mu_{n,x}(I) = (\delta_x, \chi_I(A_n)\delta_x)_{\ell^2(G_n)} = \sum_{\lambda_j \in I} |\varphi_j(x)|^2. \tag{2.1}$$

With $\mathcal{N}_I(G_n)$ we denote the counting measure that counts the number of eigenvalues λ_j in the set I,

$$\mathcal{N}_I(G_n) = \sum_{j=1}^{n} \chi_I(\lambda_j) = \text{Tr}[\chi_I(A_n)]$$

and we remark the identity

$$\mathcal{N}_I(G_n) = \sum_{x \in G_n} \mu_{n,x}(I). \tag{2.2}$$

Our main tool in the analysis of the spectral distribution of A_n is the fact that a typical regular graph locally resembles a tree in the sense that it contains large regions without cycles. (A cycle refers to an unoriented, unlabelled closed path without repetitions of vertices and edges other than the starting and ending vertex.) So we also consider the infinite regular tree $\mathcal{T}_d$ of degree d. On the tree the adjacency matrix $A_{\mathcal{T}_d}$ is again defined by

$$\left(A_{\mathcal{T}_d}\phi\right)(x) = \sum_{y \in \mathcal{T}_d : d(x,y)=1} \phi(y), \quad \phi \in \ell^2(\mathcal{T}_d), \ x \in \mathcal{T}_d.$$

Since the degree d is finite $A_{\mathcal{T}_d}$ is a bounded and self-adjoint operator with domain $\ell(\mathcal{T}_d)$. In analogy with the local spectral measure (2.1) we define the local density of states measure

$$\sigma_0(I) = (\delta_x, \chi_I(A_{\mathcal{T}_d})\delta_x)_{l^2(\mathcal{T}_d)} \tag{2.3}$$

which is independent of $x \in \mathcal{T}_d$. On the tree this can be calculated explicitly and is given by the Kesten–McKay measure:

$$\sigma_0(d\lambda) = \frac{d}{2\pi} \frac{\sqrt{4(d-1)-\lambda^2}}{d^2-\lambda^2} \chi_{(-2\sqrt{d-1}, 2\sqrt{d-1})}(\lambda)d\lambda. \tag{2.4}$$

The fact that a random regular graph is locally identical to a tree has already been used by McKay to determine the limiting distribution of eigenvalues of the adjacency matrix. Assume that for each fixed $k \geq 3$ the number of cycles of length k in G_n is of order $o(n)$ as n tends to infinity. Then it is shown in [33] that the measure $\mathcal{N}_{(\cdot)}(G_n)/n$ converges in distribution to σ_0. This local convergence was generalized in [39] to the much richer class of self-adjoint graphs that includes

trees and, in particular, regular trees. Since we aim at comparing with results on regular trees we restrict ourselves here to local convergence of random regular graphs to regular trees.

We refine the result of McKay by giving an estimate on the rate of convergence. Note that the measure σ_0 is supported in the interval $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ and that its density is bounded by γ_d, where

$$
\gamma_d = \begin{cases} \dfrac{d}{4\pi} \dfrac{1}{\sqrt{d^2 - 4(d-1)}} & \text{if } d \leq 6, \\[3mm] \dfrac{\sqrt{d-1}}{d\pi} & \text{if } d \geq 7. \end{cases}
$$

Theorem 2.1. *The local density of states of the adjacency matrix of a random regular graph satisfies the estimate*

$$
\sup_{t \in \mathbb{R}} \left[\frac{1}{n} \sum_{x \in G_n} |\mu_{n,x}((-\infty, t]) - \sigma_0((-\infty, t])| \right] \leq C \gamma_d \sqrt{d-1} \log_{d-1}^{-1}(n)
$$

asymptotically almost surely for any constant $C > 8\pi$.

In particular, the estimate

$$
\left| \frac{1}{n} \mathcal{N}_I(G_n) - \sigma_0(I) \right| \leq \delta |I| \tag{2.5}
$$

holds asymptotically almost surely, for all $\delta > 0$ and all intervals $I \subset \mathbb{R}$ satisfying

$$
|I| \geq \frac{2C \gamma_d \sqrt{d-1}}{\delta} \log_{d-1}^{-1}(n).
$$

Remark. As mentioned in the introduction this result can be seen as a generalization of recent results found in [16, 46], where the spectral distribution of the adjacency matrix is analyzed on random regular graphs with degree $d = d_n$ that tends to infinity as $n \to \infty$. In fact from [16, Lemma 10] one can also derive a result for fixed degree d, namely that (2.5) holds if the size of the interval I is larger than $\log(\log(n)) \log^{-1}(n)$.

To emphasize that our approach can be applied in various situations let us now briefly consider the case where the degree $d = d_n$ depends on the number of vertices and tends to infinity as $n \to \infty$, as studied in [16, 46]. To confine the spectrum to a finite interval one considers the rescaled operator

$$
\tilde{A}_n = \frac{1}{2\sqrt{d_n - 1}} A_n
$$

on $\ell^2(G_n)$ and the correspondingly rescaled operator $\widetilde{A}_{\mathfrak{T}_{d_n}}$ on $\ell^2(\mathfrak{T}_{d_n})$. Then the local density of states measure

$$\tilde{\sigma}_0(I) = (\delta_x, \chi_I(\widetilde{A}_{\mathfrak{T}_{d_n}})\delta_x)_{\ell^2(\mathfrak{T}_{d_n})}$$

is given by

$$\tilde{\sigma}_0(d\lambda) = \frac{2}{\pi} \frac{d_n(d_n - 1)}{d_n^2 - 4(d_n - 1)\lambda^2} \sqrt{1 - \lambda^2} \chi_{(-1,1)}(\lambda) d\lambda. \tag{2.6}$$

As n tends to infinity this measure converges to the semicircle measure

$$\sigma_{\mathrm{sc}}(d\lambda) = \frac{2}{\pi} \sqrt{1 - \lambda^2} \chi_{(-1,1)}(\lambda) d\lambda. \tag{2.7}$$

In the same way as above we define the local spectral measure

$$\tilde{\mu}_{n,x}(I) = (\delta_x, \chi_I(\widetilde{A}_n)\delta_x)_{\ell^2(G_n)}$$

and the counting measure

$$\widetilde{\mathcal{N}}_I(G_n) = \mathrm{Tr}[\chi_I(\widetilde{A}_n)] = \sum_{x \in G_n} \tilde{\mu}_{n,x}(I) \tag{2.8}$$

and we obtain the following local semicircle law.

Theorem 2.2. *Let $d_n \to \infty$ as $n \to \infty$ with $d_n \le (n/\ln(n))^{1/3}$. Then the local density of states of the operator $\widetilde{A}_n$ satisfies the estimate*

$$\sup_{t \in \mathbb{R}} \left[\frac{1}{n} \sum_{x \in G_n} |\tilde{\mu}_{n,x}((-\infty, t]) - \sigma_{\mathrm{sc}}((-\infty, t])| \right] \le C \left(\frac{\ln(d_n - 1)}{\ln(n)} + \frac{1}{d_n} \right)$$

asymptotically almost surely for any constant $C > 8$.

In particular, the estimate

$$\left| \frac{1}{n} \widetilde{\mathcal{N}}_I(G_n) - \sigma_{\mathrm{sc}}(I) \right| \le \delta |I| \tag{2.9}$$

holds asymptotically almost surely, for all $\delta > 0$ and all intervals $I \subset \mathbb{R}$ satisfying

$$|I| \ge \frac{2C}{\delta} \left(\frac{\ln(d_n - 1)}{\ln(n)} + \frac{1}{d_n} \right).$$

Remark. This theorem is similar to recent results from [16,46]. In [16, Theorem 2] it is shown that (2.9) holds with probability at least $1 - o(1/n)$, if $d_n = \ln^\gamma(n)$ and the size of the considered interval I is comparable to $\ln^{-\beta\gamma}(n)$ for $0 < \gamma < 1$ and comparable to $\ln^{-\beta}(n)$ for $\gamma \ge 1$ both with $\beta < 1$. Theorem 2.2 above improves this in the sense that one can consider slightly smaller intervals I.

In [46, Theorem 1.6] a similar estimate is shown to hold with probability at least $1 - O(\exp(-cn\sqrt{d_n}\ln(d_n)))$ if the size of the considered interval I is comparable to $\ln^{1/5}(d_n)\,d_n^{-1/10}$. As far as the size of the interval is concerned Theorem 2.2 above gives better bounds if d_n is less than $\ln^{10}(n)$. For d_n larger than that the result in [46] is stronger.

2.2. Random Schrödinger operators.

Now we study the distribution of eigenvalues of a random Schrödinger operator on a random regular graph $G_n \in \mathcal{G}_{n,d}$ with fixed degree $d \geq 3$. We define the operator

$$H_n(V) = A_n + V$$

on $\ell^2(G_n)$. The operator V denotes a random potential, a multiplication operator,

$$(V\phi)(x) = \omega_x \phi(x), \quad \phi \in \ell^2(G_n),\ x \in G_n,$$

where $(\omega_x)_{x \in G_n}$ stands for a collection of independent identically distributed real random variables with density ρ. We assume that ρ has bounded support such that $\operatorname{supp}\rho \subset (-\rho_0, \rho_0)$ for some $0 < \rho_0 < \infty$ and we write $\|\rho\|_\infty = \sup_{t \in \mathbb{R}} \rho(t)$.

For random Schrödinger operators we apply the same notation as above for eigenvalues and eigenvectors. Here the eigenvalues $(\lambda_j)_{j=1}^n$ and eigenvectors $(\varphi_j)_{j=1}^n$ are random objects even for a fixed graph G_n since they depend on the potential V. So for $x \in G_n$ we consider the random spectral measure

$$\mu_{n,x}(I; V) = (\delta_x, \chi_I(H_n(V))\,\delta_x)_{\ell^2(G_n)} = \sum_{\lambda_j \in I} |\varphi_j(x)|^2 \qquad (2.10)$$

that corresponds to (2.1). As in (2.2), with $\mathcal{N}_I(G_n; V)$ we denote the random variable that counts the number of eigenvalues λ_j in the set I,

$$\mathcal{N}_I(G_n; V) = \sum_{j=1}^n \chi_I(\lambda_j) = \operatorname{Tr}[\chi_I(H_n(V))] = \sum_{x \in G_n} \mu_{n,x}(I; V). \qquad (2.11)$$

In the same way we define the corresponding objects on the tree $\mathcal{T}_d$: First the operator

$$H_{\mathcal{T}_d}(V) = A_{\mathcal{T}_d} + V$$

on $\ell^2(\mathcal{T}_d)$. Here $V = (\omega_x)_{x \in \mathcal{T}_d}$ denotes again a collection of independent identically distributed real random variables with density ρ. We write $\mathbb{P}$ and $\mathbb{E}$ for the probability and expectation with respect to the distribution of $(\omega_x)_{x \in \mathcal{T}_d}$. We recall that the density of the random potential has finite support. Hence, $H_{\mathcal{T}_d}$ is a bounded self-adjoint operator with domain $\ell^2(\mathcal{T}_d)$.

For $x \in \mathcal{T}_d$ we define the local density of states measure

$$\mu_{\mathcal{T}_d,x}(I;V) = (\delta_x, \chi_I(H_{\mathcal{T}_d}(V))\delta_x)_{\ell^2(\mathcal{T}_d)}.$$

This measure depends on the potential V and thus on the vertex $x \in \mathcal{T}_d$. To obtain an invariant measure we take the expectation with respect to the random potential and set

$$\sigma_\rho(I) = \mathbb{E}[\mu_{\mathcal{T}_d,x}(I;V)]. \tag{2.12}$$

By translation invariance of the operator $H_{\mathcal{T}_d}(V)$ the measure σ_ρ is independent of $x \in \mathcal{T}_d$ and thus depends only on the density ρ. We refer to Appendix A.2, where we state selected properties of $\mu_{\mathcal{T}_d,x}$ and σ_ρ.

To identify the random potential on the graph G_n with the random potential on the tree $\mathcal{T}_d$ we consider the tree as the universal cover of the graph. To construct the universal cover one starts with an arbitrary vertex $o \in G_n$ and the set of non-backtracking walks in G_n that start at o. This is the set of finite sequences $(x_j)_{j=1}^k$ such that $x_1 = o$, x_j is adjacent to x_{j+1} in G_n for $j = 1, \ldots, k-1$, and $x_{j-1} \neq x_{j+1}$ for $j = 2, \ldots, k-1$. Two such walks are said to be adjacent if their lengths differ by one and if they agree except for the last vertex of the longer walk. The set of non-backtracking walks in G_n starting at o with this notion of adjacency is isomorphic to the tree $\mathcal{T}_d$ and forms the universal cover of the graph G_n.

The graph G_n induces an equivalence relation on its universal cover and thus on the tree: Two vertices of $\mathcal{T}_d$ are equivalent if the corresponding non-backtracking walks in G_n have the same endpoints. Hence, the graph G_n can be recovered from the universal cover as the set of equivalence classes. This induces a map

$$\iota \colon \mathcal{T}_d \longrightarrow G_n$$

which is onto and preserves the local geometry in the sense explained in the following.

For a vertex $x \in G_n$ and $k \in \mathbb{N}$ let

$$B_k(x) = \{y \in G_n \colon d(x,y) \leq k\}$$

denote the k-neighborhood of x, including all edges of G_n that are incident with at least one vertex y with $d(x,y) < k$. If $B_k(x)$ is acyclic then $B_k(x)$ is a finite tree of depth k. To compare the graph G_n locally to the tree we define

$$R(x) = \max\{k \in \mathbb{N} \colon B_k(x) \text{ is acyclic}\}. \tag{2.13}$$

(Since G_n does not contain double edges we have $R(x) \geq 1$ for all $x \in G_n$.)

Let $\hat{x} \in \mathcal{T}_d$ be an arbitrary vertex from the preimage of x under the map ι and let

$$\iota_{\hat{x}} \colon B_{R(x)}(\hat{x}) \subset \mathcal{T}_d \longrightarrow B_{R(x)}(x) \subset G_n \tag{2.14}$$

be the restriction of ι to the neighborhood

$$B_{R(x)}(\hat{x}) = \{\hat{y} \in \mathcal{T}_d \colon d(\hat{x}, \hat{y}) \le R(x)\}.$$

By (2.13) the map $\iota_{\hat{x}}$ is an isomorphism from $B_{R(x)}(\hat{x})$ to $B_{R(x)}(x)$.

Given a realization of the random potential $V = (\omega_{\hat{x}})_{\hat{x} \in \mathcal{T}_d}$ on the tree, this map generates a realization of the potential on the graph: For $x \in G_n$ we choose $\hat{x} \in \mathcal{T}_d$ as above and for $y \in B_{R(x)}(x)$ we set $\omega_y = \omega_{\hat{y}}$, where $\hat{y} \in B_{R(x)}(\hat{x})$ is the unique preimage of y under $\iota_{\hat{x}}$. For $y \notin B_{R(x)}(x)$ we set $\omega_y = \omega_{\hat{y}}$, where $\hat{y} \in \mathcal{T}_d$ is an arbitrary vertex from the preimage of y under ι. This procedure yields a collection of independent and identically distributed random variables $(\omega_y)_{y \in G_n}$ with density ρ. In fact, this realization of the random potential depends on $x \in G_n$ and on the choice of preimages. However, by independence of the random variables $(\omega_y)_{y \in \mathcal{T}_d}$ this dependence disappears after taking expectations. Thus we denote the resulting random potential on G_n again by V.

With this construction the local geometry and the random potential in $B_{R(x)}(x) \subset G_n$ and $B_{R(x)}(\hat{x}) \subset \mathcal{T}_d$ coincide. By induction in $k \in \mathbb{N}$, it is easy to see that $(\delta_x, H_n(V)^k \delta_x)_{\ell^2(G_n)}$ and $(\delta_{\hat{x}}, H_{\mathcal{T}_d}(V)^k \delta_{\hat{x}})_{\ell^2(\mathcal{T}_d)}$ only depend on the local geometry and on the random potential in $B_m(x)$ and $B_m(\hat{x})$ respectively, where $m = k/2$ for k even and $m = (k+1)/2$ for k odd. In particular, it follows that

$$(\delta_x, H_n(V)^k \delta_x)_{\ell^2(G_n)} = (\delta_{\hat{x}}, H_{\mathcal{T}_d}(V)^k \delta_{\hat{x}})_{\ell^2(\mathcal{T}_d)} \tag{2.15}$$

for all $k = 0, 1, \ldots, 2R(x)$. This identity is a key ingredient in the proof of the following result.

Theorem 2.3. *Assume that the density ρ of the random potential V satisfies $\|\rho\|_\infty < \infty$ and $\mathrm{supp}(\rho) \subset (-\rho_0, \rho_0)$ with $0 < \rho_0 < \infty$. Then the local density of states of the operator $H_n(V)$ on a random regular graph satisfies the estimate*

$$\sup_{t \in \mathbb{R}} \left[\frac{1}{n} \sum_{x \in G_n} \mathbb{E} |\mu_{n,x}((-\infty, t]; V) - \mu_{\mathcal{T}_d, \hat{x}}((-\infty, t]; V)| \right]$$

$$\le C \|\rho\|_\infty (2\sqrt{d-1} + \rho_0) \log_{d-1}^{-1}(n)$$

asymptotically almost surely for any constant $C > 4\pi$.

In particular, the estimate

$$\left|\frac{1}{n}\mathbb{E}\left[\mathcal{N}_I(G_n; V)\right] - \sigma_\rho(I)\right| \le \delta|I|$$

holds asymptotically almost surely, for all $\delta > 0$ and all intervals $I \subset \mathbb{R}$ satisfying

$$|I| \ge \frac{2C\|\rho\|_\infty(2\sqrt{d-1} + \rho_0)}{\delta}\log_{d-1}^{-1}(n)\,.$$

In Section 5 we apply the developed methods to deduce estimates for the Green function. We establish convergence of the imaginary part of the Green function on a random regular graph to the respective quantity on the infinite regular tree, see Corollary 5.1. Based on these bounds we prove statements about delocalization of eigenvectors. In particular, in Theorem 5.3 we show that eigenvectors of the operator $H_n(V)$ with eigenvalues within the absolutely continuous spectrum of the infinite-volume operator $H_{\mathcal{T}_d}(V)$ are not localized. Since these results require more notation and since they ask for some discussion we state these results in Section 5.

3. A deterministic estimate for the integrated density of states

In this section we fix a graph $G_n \in \mathcal{G}_{n,d}$ and a vertex $x \in G_n$. We derive an estimate for the difference between the spectral measure on G_n and the respective measure on the tree $\mathcal{T}_d$ that depends on the local geometry of the graph.

Recall the definition of $R(x)$ from (2.13) and set $R(x)^* = R(x)$ for $R(x)$ odd and $R(x)^* = R(x) - 1$ for $R(x)$ even. As in Theorem 2.3 we write $\hat{x} \in \mathcal{T}_d$ for a vertex from the preimage of $x \in G_n$ under the universal cover.

Theorem 3.1. *For all $G_n \in \mathcal{G}_{n,d}$ and all $x \in G_n$ the local density of states of A_n satisfies*

$$\sup_{t \in \mathbb{R}}[|\mu_{n,x}((-\infty, t)) - \sigma_0((-\infty, t))|] \le 4\pi\gamma_d\sqrt{d-1}\frac{1}{R(x)^*}\,.$$

Moreover, under the conditions of Theorem 2.3, the local density of states of $H_n(V)$ satisfies

$$\sup_{t \in \mathbb{R}}[\mathbb{E}|\mu_{n,x}((-\infty, t); V) - \mu_{\mathcal{T}_d, \hat{x}}((-\infty, t); V)|]$$

$$\le 2\pi\|\rho\|_\infty(2\sqrt{d-1} + \rho_0)\frac{1}{R(x)^*}$$

for all $G_n \in \mathcal{G}_{n,d}$ and all $x \in G_n$.

Remark. For the adjacency matrix A_n there is a variant of this result. In [41] it is shown that there is a constant $C > 0$ such that, for all $m \in \mathbb{N}$,

$$\sup_{t \in \mathbb{R}}[|\mu_{n,x}\,((-\infty, t]) - \sigma_0\,((-\infty, t])\,|] \leq C\Big(\frac{1}{m} + m^6\Big(\sum_{k=1}^{2m-2} W_k(x, G_n)^2\Big)^{1/2}\Big),$$

where $W_k(x, G_n)$ is related to the number of closed non-backtracking walks of length k in G_n that start at x. For $m < R(x)$ the second summand is zero. More generally, $W_k(x, G_n)$ can be estimated in terms of the number of cycles in G_n and the resulting bounds are similar to Theorem 2.1.

In the remainder of this section we prove Theorem 3.1. The proof is based on a general estimate for measures on the real line that we give in the next subsection. In Subsection 3.2 we show how the theorem can be deduced from Proposition 3.2.

3.1. A general estimate. The following result is related to the classical moment problem and the Chebyshev–Markov–Stieltjes inequality; it is based on an approximation of the Heaviside function by orthogonal polynomials. We refer to the books [5, 31, 43] for background information regarding this approach.

Proposition 3.2. *Let σ be a measure on the real line with bounded density w and with support in the finite interval $(-w_0, w_0)$. Let $N \in \mathbb{N}$ and assume that μ is another measure on the real line such that, for all $k = 0, \ldots, 2N$,*

$$\int_{\mathbb{R}} \lambda^k \sigma(d\lambda) = \int_{\mathbb{R}} \lambda^k \mu(d\lambda)\,. \tag{3.1}$$

Then the estimate

$$\sup_{t \in \mathbb{R}} |\sigma((-\infty, t]) - \mu((-\infty, t])| \leq \frac{2\pi}{N^*}\|w\|_\infty w_0$$

holds with $N^* = N$ for N odd and with $N^* = N - 1$ for N even.

Proof. First we note that for $t \leq -w_0$ we have

$$|\sigma((-\infty, t]) - \mu((-\infty, t])| \leq |\mu((-\infty, -\omega_0])|$$
$$= |\sigma((-\infty, -w_0]) - \mu((-\infty, -w_0])|$$

and for $t \geq w_0$ (by (3.1) with $k = 0$),

$$|\sigma((-\infty, t]) - \mu((-\infty, t])| = \sigma((-\infty, \omega_0]) - \mu((-\infty, t])$$
$$\leq |\sigma((-\infty, w_0]) - \mu((-\infty, w_0])|.$$

Hence, we can assume $t \in (-w_0, w_0)$.

For $n \in \mathbb{N}_0$, let P_n denote the orthonormal polynomial of degree n with respect to the measure σ. We claim that

$$|\sigma((-\infty, t]) - \mu((-\infty, t])| \leq \frac{1}{\sum_{n=0}^{N} P_n(t)^2}. \tag{3.2}$$

Combining this bound with Lemma 3.3 below proves the proposition.

To establish (3.2) let us first assume that t is a zero of the polynomial P_N. We remark that these zeros are real and simple [5] and we denote them by $\lambda_1 < \lambda_2 < \cdots < \lambda_N$. Then this assumption means that $t = \lambda_j$ for an index $j \in \{1, \ldots, N\}$. We construct a polynomial R_{2N-2} of degree $2N - 2$ that satisfies

$$R_{2N-2}(\lambda_1) = \cdots = R_{2N-2}(\lambda_j) = 1,$$

$$R_{2N-2}(\lambda_{j+1}) = \cdots = R_{2N-2}(\lambda_N) = 0,$$

and

$$R'_{2N-2}(\lambda_i) = 0$$

for all $i \neq j$. These $2N - 1$ assumptions determine the polynomial R_{2N-2} uniquely. A short argument using Rolle's theorem shows that

$$R_{2N-2}(\lambda) \geq \chi_{(-\infty, \lambda_j]}(\lambda)$$

for all $\lambda \in \mathbb{R}$ [5, p. 65]. In the same way we construct a polynomial Q_{2N-2} of degree $2N - 2$ by changing only the condition at λ_j to $Q_{2N-2}(\lambda_j) = 0$. Then we get $Q_{2N-2}(\lambda) \leq \chi_{(-\infty, \lambda_j]}(\lambda)$ for all $\lambda \in \mathbb{R}$. Hence, we can estimate

$$\sigma((-\infty, \lambda_j]) = \int_{\mathbb{R}} \chi_{(-\infty, \lambda_j]} \sigma(d\lambda) \leq \int_{\mathbb{R}} R_{2N-2}(\lambda) \sigma(d\lambda),$$

$$\mu((-\infty, \lambda_j]) = \int_{\mathbb{R}} \chi_{(-\infty, \lambda_j]} \mu(d\lambda) \geq \int_{\mathbb{R}} Q_{2N-2}(\lambda) \mu(d\lambda) \tag{3.3}$$

and

$$\sigma((-\infty, \lambda_j]) - \mu((-\infty, \lambda_j]) \leq \int_{\mathbb{R}} R_{2N-2}(\lambda) \sigma(d\lambda) - \int_{\mathbb{R}} Q_{2N-2}(\lambda) \mu(d\lambda). \tag{3.4}$$

To bound the right-hand side we invoke the Gaussian quadrature formula from Lemma A.1 in Appendix A.1: With $M = N - 1$ and $s = 0$ we find

$$\int_{\mathbb{R}} R_{2N-2}(\lambda)\sigma(d\lambda) = \sum_{k=1}^{N} \frac{R_{2N-2}(\lambda_k)}{\sum_{n=0}^{N-1} P_n(\lambda_k)^2} = \sum_{k=1}^{j} \frac{1}{\sum_{n=0}^{N-1} P_n(\lambda_k)^2}.$$

By assumption, the first $2N$ moments of the measures σ and μ agree, so we also get

$$\int_{\mathbb{R}} Q_{2N-2}(\lambda)\mu(d\lambda) = \int_{\mathbb{R}} Q_{2N-2}(\lambda)\sigma(d\lambda)$$

$$= \sum_{k=1}^{N} \frac{Q_{2N-2}(\lambda_k)}{\sum_{n=0}^{N-1} P_n(\lambda_k)^2}$$

$$= \sum_{k=1}^{j-1} \frac{1}{\sum_{n=0}^{N-1} P_n(\lambda_k)^2}.$$

Combining these identities with the estimate (3.4) yields the upper bound

$$\sigma((-\infty, \lambda_j]) - \mu((-\infty, \lambda_j]) \leq \frac{1}{\sum_{n=0}^{N-1} P_n(\lambda_j)^2} = \frac{1}{\sum_{n=0}^{N} P_n(\lambda_j)^2},$$

where we used the assumption $P_N(\lambda_j) = 0$ in the last step. The lower bound is proved in the same way by exchanging the roles of σ and μ in (3.3) and (3.4). This proves (3.2) if t is a zero of P_N.

It remains to prove (3.2) if t is not a zero of P_N. For $s \in \mathbb{R}$ we define a polynomial of degree $N + 1$ by

$$\widehat{P}_{N+1}(\lambda) = P_{N+1}(\lambda) + s P_N(\lambda).$$

Since $P_N(t) \neq 0$ we can choose s in such a way that

$$\widehat{P}_{N+1}(t) = 0.$$

Thus we can argue in the same way as before, with N replaced by $N + 1$.

This establishes (3.2) and completes the proof. $\square$

The proof of Proposition 3.2 is based on the following estimate of the so-called Christoffel numbers $(\sum_{n=0}^{N} P_n(t)^2)^{-1}$.

Lemma 3.3. *Assume that σ is a measure on the real line with bounded density w and with support in the finite interval $(-w_0, w_0)$. Let P_n, $n \in \mathbb{N}_0$, denote the orthonormal polynomial of degree n with respect to σ.*

Then for all $N \in \mathbb{N}$ and all $t \in (-w_0, w_0)$ the estimate

$$\frac{1}{\sum_{n=0}^{N} P_n(t)^2} \leq \frac{2\pi w_0 \|w\|_\infty}{N^*}$$

holds with $N^* = N$ for N odd and with $N^* = N - 1$ for N even.

Proof. We fix $N \in \mathbb{N}$ and $t \in (-w_0, w_0)$. Below we construct a polynomial $S_{2N-2}^{(t)}$ of degree less or equal than $2N - 2$ with the properties

$$S_{2N-2}^{(t)}(\lambda) \geq 0 \tag{3.5}$$

for all $\lambda \in \mathbb{R}$,

$$S_{2N-2}^{(t)}(t) = 1, \tag{3.6}$$

and

$$\int_{-w_0}^{w_0} S_{2N-2}^{(t)}(\lambda) d\lambda \leq \frac{2\pi w_0}{N^*}. \tag{3.7}$$

To estimate $\sum_{n=0}^{N} P_n(t)^2$ in terms of this polynomial let us first assume that t is a zero of P_N. Let $\lambda_1 < \lambda_2 < \cdots < \lambda_N$ denote the zeros of P_N and assume that $t = \lambda_j$ for an index $j \in \{1, \ldots, N\}$. Then the assumption $P_N(\lambda_j) = 0$ together with (3.5) and (3.6) implies

$$\frac{1}{\sum_{n=0}^{N} P_n(t)^2} = \frac{1}{\sum_{n=0}^{N-1} P_n(\lambda_j)^2} = \frac{S_{2N-2}^{(t)}(\lambda_j)}{\sum_{n=0}^{N-1} P_n(\lambda_j)^2} \leq \sum_{k=1}^{N} \frac{S_{2N-2}^{(t)}(\lambda_k)}{\sum_{n=0}^{N-1} P_n(\lambda_k)^2}.$$

We combine this estimate with the quadrature formula from Lemma A.1 and get

$$\frac{1}{\sum_{n=0}^{N} P_n(t)^2} \leq \int_{\mathbb{R}} S_{2N-2}^{(t)}(\lambda) \sigma(d\lambda). \tag{3.8}$$

If $t \in (-w_0, w_0)$ is not a zero of P_N we define, for $s \in \mathbb{R}$, a polynomial

$$\widehat{P}_{N+1}(\lambda) = P_{N+1}(\lambda) + s P_N(\lambda).$$

Since $P_N(t) \neq 0$ we can choose s such that $\widehat{P}_{N+1}(t) = 0$. Now we can argue similarly as above: Again let $\lambda_1 < \cdots < \lambda_{N+1}$ denote the zeros of $\widehat{P}_{N+1}$ so that $t = \lambda_j$ for an index $j \in \{1, \ldots, N+1\}$. Now we apply Lemma A.1 directly to

$$\frac{1}{\sum_{n=0}^{N} P_n(t)^2} = \frac{S_{2N-2}^{(t)}(\lambda_j)}{\sum_{n=0}^{N} P_n(\lambda_j)^2} \leq \sum_{k=1}^{N+1} \frac{S_{2N-2}^{(t)}(\lambda_k)}{\sum_{n=0}^{N} P_n(\lambda_k)^2}$$

and we obtain (3.8). We have shown that the estimate (3.8) is valid for all $t \in (-w_0, w_0)$. Now we combine this estimate with (3.7) and arrive at

$$\frac{1}{\sum_{n=0}^{N} P_n(t)^2} \le \int_{\mathbb{R}} S_{2N-2}^{(t)}(\lambda)\sigma(d\lambda)$$

$$\le \|w\|_\infty \int_{-w_0}^{w_0} S_{2N-2}^{(t)}(\lambda)d\lambda$$

$$\le \frac{2\pi w_0 \|w\|_\infty}{N^*}$$

which is the claimed estimate.

It remains to construct the polynomial $S_{2N-2}^{(t)}$. Let T_m denote the Chebyshev polynomial of degree $m \in \mathbb{N}_0$ that is defined by the equation

$$T_m(\cos \theta) = \cos(m\theta).$$

For $N \in \mathbb{N}$, let $n \in \mathbb{N}$ denote the largest integer satisfying $4n \le 2N - 2$. Then, for $x \in \mathbb{R}$, we set

$$F_{2N-2}(x) = \frac{1}{2n+1} + \frac{2}{(2n+1)^2} \sum_{m=1}^{2n} (2n - m + 1)(-1)^m T_{2m}(x).$$

This defines a polynomial of degree $4n \le 2N - 2$.

Let us note the following relation to the Fejér kernel. For $x \in (-1, 1)$ write $x = \cos \theta$ with $\theta \in (0, \pi)$ and calculate

$$F_{2N-2}(x) = F_{2N-2}(\cos \theta)$$

$$= \frac{1}{2n+1} + \frac{2}{(2n+1)^2} \sum_{m=1}^{2n} (2n - m + 1)(-1)^m \cos(2m\theta)$$

$$= \frac{1}{2} \frac{1}{(2n+1)^2} \left(\frac{\sin^2((2n+1)(\theta - \frac{\pi}{2}))}{\sin^2(\theta - \frac{\pi}{2})} + \frac{\sin^2((2n+1)(\theta + \frac{\pi}{2}))}{\sin^2(\theta + \frac{\pi}{2})} \right).$$

This identity shows that $F_{2N-2}(0) = F_{2N-2}(\cos(\pi/2)) = 1$, that $F_{2N-2}(x) \ge 0$ for $x \in (-1, 1)$, and that

$$\int_{-1}^{1} F_{2N-2}(x)dx \le \frac{\pi}{2n+1} \le \frac{\pi}{N^*}.$$

To see that $F_{2N-2}(x)$ is non-negative for all $x \in \mathbb{R}$ note that it vanishes together with its derivative at the points $x_k = \cos(\pi/2 + k\pi/(2n+1))$, $k = 1, 2, \ldots, n$. Together with the condition $F_{2N-2}(0) = 1$ these are $2n + 1$ conditions. Since F_{2N-2} is by definition an even polynomial of exact degree $4n$ these conditions determine the polynomial uniquely and show that it is non-negative.

Thus, for $t \in (-w_0, w_0)$ and $\lambda \in \mathbb{R}$, we set

$$S_{2N-2}^{(t)}(\lambda) = F_{2N-2}\left(\frac{\lambda - t}{2w_0}\right)$$

and the properties (3.5), (3.6), and (3.7) follow directly from the properties of F_{2N-2}. This completes the proof. □

3.2. Proof of Theorem 3.1. Let us first consider the adjacency matrix A_n. For this operator the claim follows directly from Proposition 3.2. We only have to show that the measures $\mu_{n,x}$ and σ_0 satisfy the conditions of the proposition with $N = R(x)$: Recall that the measure σ_0 is supported in the finite interval $(-2\sqrt{d-1}, 2\sqrt{d-1})$ and that its density is bounded by γ_d. Hence it remains to establish that the k-th moments of σ_0 and $\mu_{n,x}$ agree for $k = 0, \ldots, 2R(x)$.

For $x \in G_n$ we pick a vertex $\hat{x} \in \mathcal{T}_d$ from the preimage of x under the universal cover and consider the map $\iota_{\hat{x}}$ from (2.14). It maps the neighborhood $B_{R(x)}(\hat{x}) \subset \mathcal{T}_d$ isomorphically to $B_{R(x)}(x) \subset G_n$ and as in (2.15) we find $(\delta_x, A_n^k \delta_x)_{\ell^2(G_n)} = (\delta_{\hat{x}}, A_{\mathcal{T}_d}^k \delta_{\hat{x}})_{\ell^2(\mathcal{T}_d)}$ for $k = 0, 1, \ldots, 2R(x)$. By definition of the measures $\mu_{n,x}$ and σ_0, see (2.1) and (2.3), this implies

$$\int_{\mathbb{R}} \lambda^k \mu_{n,x}(d\lambda) = (\delta_x, A_n^k \delta_x)_{\ell^2(G_n)} = (\delta_{\hat{x}}, A_{\mathcal{T}_d}^k \delta_{\hat{x}})_{\ell^2(\mathcal{T}_d)} = \int_{\mathbb{R}} \lambda^k \sigma_0(d\lambda)$$

for $k = 0, 1, \ldots, 2R(x)$. Hence, these measures satisfy the conditions of Proposition 3.2 and the proof of the first statement is complete.

To prove the second claim we have to argue a bit more carefully. For the random operator $H_n(V)$ the spectral measure $\mu_{\mathcal{T}_d,x}(I; V)$ is not necessarily absolutely continuous, hence we cannot apply Proposition 3.2 directly. (We note that σ_ρ has bounded density by the Wegner estimate (A.2), so we can apply the proposition to the measures σ_ρ and $\mathbb{E}[\mu_{n,x}]$. This immediately yields a similar estimate for the difference of $\mathbb{E}[\mu_{n,x}]$ and σ_ρ, however, we want to prove a stronger statement.)

From identity (2.15) we see that

$$\int_{\mathbb{R}} \lambda^k \mu_{n,x}(d\lambda; V) = (\delta_x, H_n(V)^k \delta_x)_{\ell^2(G_n)}$$

$$= (\delta_{\hat{x}}, H_{\mathbb{T}_d}(V)^k \delta_{\hat{x}})_{\ell^2(\mathbb{T}_d)}$$

$$= \int_{\mathbb{R}} \lambda^k \mu_{\mathbb{T}_d, \hat{x}}(d\lambda; V)$$

is valid for all $k = 0, \ldots, 2R(x)$. Hence, we can apply estimate (3.2) to the measures $\mu_{n,x}$ and $\mu_{\mathbb{T}_d, \hat{x}}$ and obtain

$$|\mu_{n,x}((-\infty, t]; V) - \mu_{\mathbb{T}_d, \hat{x}}((-\infty, t]; V)| \leq \frac{1}{\sum_{n=0}^{R(x)} P_n(t)^2}, \qquad (3.9)$$

where P_n denotes the orthonormal polynomial of degree n with respect to the random measure $\mu_{\mathbb{T}_d, \hat{x}}$.

From the general property (A.1) we learn the the support of this measure is almost surely contained in the interval $(-2\sqrt{d-1} - \rho_0, 2\sqrt{d-1} + \rho_0)$. Hence, in the same way as in the beginning of the proof of Proposition 3.2 we can reduce the problem to $t \in (-2\sqrt{d-1} - \rho_0, 2\sqrt{d-1} + \rho_0)$. For such t estimate (3.8) is valid and combining it with (3.9) we find that the bound

$$\left|\mu_{n,x}((-\infty, t]; V) - \mu_{\mathbb{T}_d, \hat{x}}((-\infty, t]; V)\right| \leq \int_{\mathbb{R}} S_{2R(x)-2}^{(t)}(\lambda) \mu_{\mathbb{T}_d, \hat{x}}(d\lambda; V)$$

holds almost surely. Hence, recalling definition (2.12), the Wegner estimate (A.2), and the bound (3.7), we obtain

$$\mathbb{E}|\mu_{n,x}((-\infty, t]; V) - \mu_{\mathbb{T}_d, \hat{x}}((-\infty, t]; V)| \leq \int_{\mathbb{R}} S_{2R(x)-2}^{(t)}(\lambda) \sigma_\rho(d\lambda)$$

$$\leq \frac{2\pi \|\rho\|_\infty (2\sqrt{d-1} + \rho_0)}{R^*(x)}.$$

This proves the second claim and completes the proof of Theorem 3.1.

4. Asymptotically almost sure bounds on the rate of convergence

In this section we combine the deterministic estimates of Theorem 3.1 with bounds on the number of cycles in random regular graphs to prove the results from Section 2.

4.1. Acyclic regions in random regular graphs. Here we collect information about cycles in random regular graphs based on results from [34,35]. We establish the fact that typically at most sites the graph looks locally like a tree. This is made precise in Lemma 4.2 below.

For a set A we write $|A|$ for the number of elements, in particular for a subset $F \subset G_n$ of a graph, $|F|$ denotes the number of vertices. For a graph $G_n \in \mathcal{G}_{n,d}$ and $k \in \mathbb{N}$ let $\mathcal{C}(k)$ denote the set of cycles of length k in G_n.

Lemma 4.1. *For $3 \le k \le nd/4 - 2d^2$ one has*

$$\mathcal{E}_{n,d}[|\mathcal{C}(k)|] \le \frac{(d-1)^k}{2k}\Big(1 + \frac{8}{n}\Big(d + \frac{k}{2d}\Big)\Big)^k.$$

Proof. Let G_n be an arbitrary graph from $\mathcal{G}_{n,d}$ and let $e(G_n)$ denote the set of edges of G_n. We also introduce K_n, the complete graph of n vertices and the set $e(K_n)$ of edges of K_n. We consider a cycle $\mathfrak{c} \subset K_n$ of length k and its set of edges $e(\mathfrak{c})$. To estimate the expectation of $|\mathcal{C}(k)|$ we use the following relation to the number of cycles $\mathfrak{c}$ of length k in K_n:

$$\begin{aligned}
\mathcal{E}_{n,d}[|\mathcal{C}(k)|] &= \sum_{\mathfrak{c} \subset K_n} \mathcal{P}_{n,d}[e(\mathfrak{c}) \subset e(G_n)] \\
&\le |\{\mathfrak{c} \subset K_n\}| \max_{\mathfrak{c} \subset K_n} \mathcal{P}_{n,d}[e(\mathfrak{c}) \subset e(G_n)].
\end{aligned} \tag{4.1}$$

From [35, Theorem 3] (see also [34, Theorem 2.10]) it follows that for $k \le nd/4 - 2d^2$ one has

$$\begin{aligned}
\mathcal{P}_{n,d}[e(\mathfrak{c}) \subset e(G_n)] &\le \frac{d^k(d-1)^k}{2^k}\Big(\frac{2}{nd - 4d^2 - 2k}\Big)^k \\
&= \frac{(d-1)^k}{n^k}\Big(\frac{nd}{nd - 4d^2 - 2k}\Big)^k.
\end{aligned}$$

Moreover, a counting argument shows that

$$|\{\mathfrak{c} \subset K_n\}| = \frac{n!}{2k(n-k)!} \le \frac{n^k}{2k}.$$

We combine these estimates with (4.1) and get

$$\mathcal{E}_{n,d}[|\mathcal{C}(k)|] \le \frac{(d-1)^k}{2k}\Big(\frac{nd}{nd - 4d^2 - 2k}\Big)^k.$$

A simple estimate using the fact that $k \le nd/4 - 2d^2$ yields the claim. $\qquad\square$

Remark. A similar argument leads to a lower bound on $\mathcal{E}_{n,d}\left[|\mathcal{C}(k)|\right]$ with the same leading term $(d-1)^k/2k$. In this sense the estimate in Lemma 4.1 is sharp.

The following result is a simplified version of [16, Lemma 4]. It quantifies how well one can approximate a graph $G_n \in \mathcal{G}_{n,d}$ by a tree.

Lemma 4.2. *For $k \in \mathbb{N}$ let $F_n(k) \subset G_n$ denote the set of vertices $x \in G_n$ such that $B_k(x)$ is acyclic. Then for all $\epsilon > 0$ and $k \leq n/4d - 2d^2$ we have*

$$\mathcal{P}_{n,d}\left[1 - \frac{|F_n(k)|}{n} > \epsilon\right] \leq \frac{1}{2n\epsilon} \frac{(d-1)^{2k+1/2}}{\sqrt{d-1}-1}\left(1 + \frac{8}{n}\left(d + \frac{k}{d}\right)\right)^{2k}.$$

Proof. We apply the Markov inequality, namely that for all $\epsilon > 0$,

$$\mathcal{P}_{n,d}\left[1 - \frac{|F_n(k)|}{n} > \epsilon\right] = \mathcal{P}_{n,d}\left[|G_n \setminus F_n(k)| > n\epsilon\right]$$

$$\leq \frac{1}{n\epsilon}\mathcal{E}_{n,d}\left[|G_n \setminus F_n(k)|\right]. \tag{4.2}$$

Hence, we have to estimate $\mathcal{E}_{n,d}\left[|G_n \setminus F_n(k)|\right]$.

Consider $\mathfrak{c} \in \mathcal{C}(m)$, i.e. a cycle $\mathfrak{c} \subset G_n$ of length $m \in \mathbb{N}$. For $k \in \mathbb{N}$ set

$$N_{\mathfrak{c}}(k) = \{x \in G_n : \mathfrak{c} \subset B_k(x)\}.$$

For $k < m/2$ the set $N_{\mathfrak{c}}(k)$ is empty. For $k \geq m/2$ it is included in the set of vertices that are at distance less or equal than $k - m/2$ from $\mathfrak{c}$. Hence, we have $|N_{\mathfrak{c}}(k)| = 0$ for $k < m/2$ and

$$|N_{\mathfrak{c}}(k)| \leq m(d-1)^{k-m/2}$$

for $k \geq m/2$. For each vertex $x \in G_n \setminus F_n(k)$ the neighborhood $B_k(x)$ contains at least one cycle so that

$$G_n \setminus F_n(k) \subset \bigcup_{\mathfrak{c}} N_{\mathfrak{c}}(k).$$

Hence,

$$|G_n \setminus F_n(k)| \leq \sum_{m \geq 3} \sum_{\mathfrak{c} \in \mathcal{C}(m)} |N_{\mathfrak{c}}(k)| \leq \sum_{m=3}^{2k} m(d-1)^{k-m/2}|\mathcal{C}(m)|.$$

We combine this estimate with Lemma 4.1 and conclude that

$$\mathcal{E}_{n,d}[|G_n \setminus F_n(k)|] \leq \frac{1}{2} \sum_{m=3}^{2k} (d-1)^{k+m/2}\Big(1 + \frac{8}{n}\Big(d + \frac{m}{2d}\Big)\Big)^m$$

$$\leq \frac{1}{2} \frac{(d-1)^{2k+1/2}}{\sqrt{d-1}-1}\Big(1 + \frac{8}{n}\Big(d + \frac{k}{d}\Big)\Big)^{2k}.$$

Inserting this into (4.2) finishes the proof. □

4.2. Proof of the main results. With Lemma 4.2 at hand we can deduce the results of Section 2 from Theorem 3.1. Here we give the proofs of Theorem 2.3 and Theorem 2.2. The proof of Theorem 2.1 is similar.

Proof of Theorem 2.3. Recall the definition of $R(x)$ from the beginning of Section 3 and the definition of $F_n(k)$ from Lemma 4.2. We can always estimate $R(x) \geq 1$ and for $x \in F_n(k)$ we have $R(x) \geq k$. Thus Theorem 3.1 implies that, for all $k \in \mathbb{N}$,

$$\sup_{t \in \mathbb{R}} \Big[\frac{1}{n} \sum_{x \in G_n} \mathbb{E}|\mu_{n,x}((-\infty, t]; V) - \mu_{\mathcal{T}_d, \hat{x}}((-\infty, t]; V)|\Big]$$

$$\leq 2\pi \|\rho\|_\infty (2\sqrt{d-1} + \rho_0) \frac{1}{n}\Big(\frac{1}{k-1}|F_n(k)| + |G_n \setminus F_n(k)|\Big)$$

$$\leq 2\pi \|\rho\|_\infty (2\sqrt{d-1} + \rho_0)\Big(\frac{1}{k-1} + \frac{1}{n}|G_n \setminus F_n(k)|\Big). \qquad (4.3)$$

We apply Lemma 4.2 to estimate the second term. For $\epsilon > 0$ let $\Omega(\epsilon, k)$ denote the event $\{|G_n \setminus F_n(k)| \leq n\epsilon/(k-1)\}$ and note that by Lemma 4.2

$$\mathcal{P}_{n,d}[\Omega(\epsilon, k)] \geq 1 - \frac{k-1}{2n\epsilon} \frac{(d-1)^{2k+1/2}}{\sqrt{d-1}-1}\Big(1 + \frac{8}{n}\Big(d + \frac{k}{d}\Big)\Big)^{2k}. \qquad (4.4)$$

Now we choose $k = \kappa \log_{d-1}(n) + 1$ with $\kappa < 1/2$. Then we see that for fixed $\epsilon > 0$ the event $\Omega(\epsilon, \kappa \log_{d-1}(n) + 1)$ holds asymptotically almost surely. With this choice of k estimate (4.3) shows that the bound

$$\sup_{t \in \mathbb{R}} \Big[\frac{1}{n} \sum_{x \in G_n} \mathbb{E}|\mu_{n,x}((-\infty, t]; V) - \mu_{\mathcal{T}_d, \hat{x}}((-\infty, t]; V)|\Big]$$

$$\leq \frac{(1+\epsilon)2\pi}{k-1} \|\rho\|_\infty (2\sqrt{d-1} + \rho_0)$$

$$\leq \frac{(1+\epsilon)2\pi}{\kappa \log_{d-1}(n)} \|\rho\|_\infty (2\sqrt{d-1} + \rho_0)$$

holds on the event $\Omega(\epsilon, \kappa \log_{d-1}(n)+1)$, so the bound holds asymptotically almost surely. This proves the first part of the theorem. In view of (2.11) and (2.12) the second statement is a direct consequence of the first. $\qquad\square$

Proof of Theorem 2.2. Note that the measure $\tilde{\sigma}_0$, given in (2.6), is supported in $(-1, 1)$ and that its density is bounded by

$$\tilde{\gamma}_d = \begin{cases} \dfrac{d\sqrt{d-1}}{2\pi}\dfrac{1}{\sqrt{d^2-4(d-1)}} & \text{if } d \leq 6, \\[3mm] \dfrac{2}{\pi}\dfrac{d(d-1)}{d^2} & \text{if } d \geq 7. \end{cases}$$

Thus we can argue as in Section 3.2 to obtain the estimate

$$\sup_{t\in\mathbb{R}}[|\tilde{\mu}_{n,x}((-\infty, t]) - \tilde{\sigma}_0((-\infty, t])|] \leq \frac{2\pi\tilde{\gamma}_d}{R(x)^*}. \tag{4.5}$$

Hence, on the event

$$\Omega(\epsilon, k) = \left\{|G_n \setminus F_n(k)| \leq \frac{n\epsilon}{(k-1)}\right\}$$

we find that

$$\sup_{t\in\mathbb{R}}\left[\frac{1}{n}\sum_{x\in G_n}|\tilde{\mu}_{n,x}((-\infty, t]) - \tilde{\sigma}_0((-\infty, t])|\right]$$

$$\leq 2\pi\tilde{\gamma}_{d_n}\left(\frac{1}{k-1} + \frac{1}{n}|G_n \setminus F_n(k)|\right)$$

$$\leq (1+\epsilon)\tilde{\gamma}_{d_n}\frac{2\pi}{k-1}.$$

The probability of $\Omega(\epsilon, k)$ is bounded by (4.4). So we can again choose $k = \kappa \ln(n)/\ln(d_n - 1) + 1$ with $\kappa < 1/2$ and we note that the assumption $d_n \leq (n/\ln(n))^{1/3}$ ensures that the condition of Lemma 4.2 is satisfied. We deduce that the event $\Omega(\epsilon, \kappa \ln(n)/\ln(d_n-1)+1)$ holds, for all $\epsilon > 0$, asymptotically almost surely. Thus the estimate

$$\sup_{t\in\mathbb{R}}\left[\frac{1}{n}\sum_{x\in G_n}|\tilde{\mu}_{n,x}((-\infty, t]) - \tilde{\sigma}_0((-\infty, t])|\right] \leq \frac{4(1+2\epsilon)}{\kappa}\frac{\ln(d_n-1)}{\ln(n)}$$

holds, for all $\epsilon > 0$, asymptotically almost surely. Here we used the fact that $\tilde{\gamma}_{d_n}$ tends to $2/\pi$ as d_n tends to infinity. It remains to note that

$$\sup_{t\in\mathbb{R}}[|\tilde{\sigma}_0((-\infty, t]) - \sigma_{\text{sc}}((-\infty, t])|] \leq (1+\epsilon)\frac{2}{\pi d_n} \tag{4.6}$$

for d_n large enough. Thus applying the triangle inequality yields the first claim. By (2.8) the second claim follows from the first. $\qquad\square$

5. Estimates for the Green function and delocalization of eigenvectors

Here we apply the results of the previous sections to compare Green functions on the finite graph G_n with the respective Green functions on the infinite tree. Then we deduce delocalization of eigenvectors for the adjacency matrix and for random Schrödinger operators. In particular, we show that eigenvectors of the operator $H_n(V)$ with eigenvalues in the regime of absolutely continuous spectrum of the infinite-volume operator $H_{\mathcal{T}_d}(V)$ are not uniformly localized as n tends to infinity.

5.1. Convergence of the Green function. Let us first introduce some notation. For $z \in \mathbb{C}_+$ and $x \in G_n$ we consider diagonal elements of the Green function, the Stieltjes transform of the local spectral measures:

$$\Gamma_n(x, z) = (\delta_x, (A_n - z)^{-1}\delta_x)_{\ell^2(G_n)}$$

$$= \int_{\mathbb{R}} (\lambda - z)^{-1} \mu_{n,x}(d\lambda),$$

$$\tilde{\Gamma}_n(x, z) = (\delta_x, (\tilde{A}_n - z)^{-1}\delta_x)_{\ell^2(G_n)}$$

$$= \int_{\mathbb{R}} (\lambda - z)^{-1} \tilde{\mu}_{n,x}(d\lambda),$$

$$\Gamma_n(x, z; V) = (\delta_x, (H_n(V) - z)^{-1}\delta_x)_{\ell^2(G_n)}$$

$$= \int_{\mathbb{R}} (\lambda - z)^{-1} \mu_{n,x}(d\lambda; V). \tag{5.1}$$

In the same way, we define the corresponding Green function on the infinite tree $\mathcal{T}_d$. One can either use resolvent expansions and the geometric structure of the tree or the explicit representations of the measures σ_0 and σ_{sc} (see (2.4) and (2.7)) to derive that, for any $\hat{x} \in \mathcal{T}_d$,

$$\Gamma_{\mathcal{T}_d}(\hat{x}, z) = \Gamma_{\mathcal{T}_d}(z)$$

$$= \int_{\mathbb{R}} (\lambda - z)^{-1} \sigma_0(d\lambda) \tag{5.2}$$

$$= \frac{-z(d - 2) - d\sqrt{z^2 - 4(d - 1)}}{2(z^2 - d^2)}$$

and

$$\lim_{n \to \infty} \tilde{\Gamma}_{\mathcal{T}_{d_n}}(\hat{x}, z) = \Gamma_{\mathrm{sc}}(z)$$

with

$$\Gamma_{\mathrm{sc}}(z) = \int_{\mathbb{R}} (\lambda - z)^{-1} \sigma_{\mathrm{sc}}(d\lambda) = -\frac{1}{2}(z - \sqrt{z^2 - 4}).$$

Here we specify the root of a complex number in $\mathbb{C}_+$ as the one with positive imaginary part. As in Theorem 2.3, for given $x \in G_n$, we choose $\hat{x} \in \mathcal{T}_d$ from the preimage of x under the universal cover.

Corollary 5.1. *Let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a sequence of positive numbers such that $\varepsilon_n = o(1)$ as $n \to \infty$. Then the following estimates hold asymptotically almost surely.*

(1) *Assume that $(z_n)_{n \in \mathbb{N}}$ is a sequence of complex numbers with*

$$\Im z_n \geq C(\varepsilon_n \log_{d-1}(n))^{-1}$$

for a constant

$$C > 16\pi \gamma_d \sqrt{d-1}.$$

Then we have

$$\frac{1}{n} \sum_{x \in G_n} |\Im \Gamma_n(x, z_n) - \Im \Gamma_{\mathcal{T}_d}(z_n)| \leq \varepsilon_n.$$

(2) *Let $d_n \to \infty$ as $n \to \infty$ with $d_n \leq (n/\ln(n))^{1/3}$ and assume that $(z_n)_{n \in \mathbb{N}}$ is a sequence of complex numbers with*

$$\Im z_n \geq C \varepsilon_n^{-1}(\log_{d_n-1}^{-1}(n) + d_n^{-1})$$

for a constant

$$C > 16.$$

Then we have

$$\frac{1}{n} \sum_{x \in G_n} |\Im \widetilde{\Gamma}_n(x, z_n) - \Im \Gamma_{\mathrm{sc}}(z_n)| \leq \varepsilon_n .$$

(3) *Let the random potential V satisfy the conditions of Theorem 2.3 and assume that $(z_n)_{n \in \mathbb{N}}$ is a sequence of complex numbers with*

$$\Im z_n \geq C(\varepsilon_n \log_{d-1}(n))^{-1}$$

for a constant

$$C > 8\pi \|\rho\|_\infty (2\sqrt{d-1} + \rho_0).$$

Then we have

$$\frac{1}{n} \sum_{x \in G_n} \mathbb{E}|\Im \Gamma_n(x, z_n; V) - \Im \Gamma_{\mathcal{T}_d}(\hat{x}, z_n; V)| \leq \varepsilon_n.$$

Proof. Let us show how the first statement can be deduced from Theorem 2.1. The other two statements follow in the same way from Theorem 2.2 and Theorem 2.3 respectively.

For $z \in \mathbb{C}_+$ write

$$z = E + i\eta,$$

with $E \in \mathbb{R}$ and $\eta > 0$, and

$$g_{E,\eta}(\lambda) = \Im(\lambda - z)^{-1} = \frac{\eta}{(E - \lambda)^2 + \eta^2}. \tag{5.3}$$

We note that

$$\int_{\mathbb{R}} \left| \frac{\partial g_{E,\eta}}{\partial \lambda}(\lambda) \right| d\lambda = \frac{2}{\eta}. \tag{5.4}$$

By (5.1) we have

$$\Im\Gamma_n(x, z_n) = \int_{\mathbb{R}} g_{E_n,\eta_n}(\lambda)\mu_{n,x}(d\lambda) = -\int_{\mathbb{R}} \mu_{n,x}\left((-\infty, \lambda]\right) dg_{E_n,\eta_n}(\lambda)$$

and by (5.2)

$$\Im\Gamma_{\mathfrak{T}_d}(z_n) = -\int_{\mathbb{R}} \sigma_0\left((-\infty, \lambda]\right) dg_{E_n,\eta_n}(\lambda).$$

Combining these identities with Theorem 2.1 and with (5.4) yields the claim. □

Remark. The third statement of the corollary gives a partial answer to a question raised in [25], whether

$$\mathbb{E}|\Gamma_n(0, E + i\eta_n; V) - \Gamma_{\mathfrak{T}_d}(0, E + i\eta_n; V)| \longrightarrow 0$$

for a sequence of positive numbers η_n that is of order $o(1)$ as $n \to \infty$. At least for the imaginary part of the Green function, the third statement of Corollary 5.1 can be interpreted in this way, if the vertex 0 is chosen at random with uniform probability from G_n.

5.2. Delocalization of eigenvectors. In this subsection we analyze how the behavior of eigenvectors of the finite-volume operator $H_n(V)$ is related to spectral properties of the infinite-volume operator $H_{\mathfrak{T}_d}(V)$. In the regime of absolutely continuous spectrum of the infinite-volume operator the corresponding generalized eigenfunctions are delocalized: They are not square-summable and in particular not localized to a bounded set (see for example [23] for the behavior of eigenfunctions of the adjacency matrix on an infinite regular tree and [4, 30] for random Schrödinger operators).

In finite volume the spectrum is of course always pure point and eigenvectors are square-summable. In the following we show that one can nevertheless find remnants of delocalization for eigenvectors of the finite volume operator.

Let us first consider the rescaled adjacency matrix $\widetilde{A}_n$ on a random regular graph with degree d_n tending to infinity as $n \to \infty$. In this case the spectral measure in the limit of infinite volume $n \to \infty$ is given by the semicircle measure σ_{sc} defined in (2.7). This measure is purely absolutely continuous with bounded density. Thus eigenvectors of $\widetilde{A}_n$ are expected to be delocalized for large n. Proposition 5.2 shows that this is justified. The result is similar to [16, Theorem 3] and the remark after Theorem 2.2 applies. We include the statement and its proof because it serves as an illustration for the more involved result about eigenvectors of random Schrödinger operators that is given in Theorem 5.3 below.

The proof relies on the fact that the Stieltjes transform of an absolutely continuous measure has a uniformly bounded imaginary part. In this way absolute continuous spectrum is related to boundedness of the imaginary part of the Green function: Since the semicircle distribution is absolutely continuous with bounded density we find, for all $E \in \mathbb{R}$ and $\eta > 0$,

$$\Im \Gamma_{\mathrm{sc}}(E + i\eta) = \int_{\mathbb{R}} g_{E,\eta}(\lambda) \sigma_{\mathrm{sc}}(d\lambda) \leq \frac{2}{\pi} \int_{\mathbb{R}} g_{E,\eta}(\lambda) d\lambda = 2 . \qquad (5.5)$$

This uniform estimate is a crucial ingredient in the proof of the following proposition.

Proposition 5.2. *Let d_n satisfy the conditions of Theorem 2.2 and let Λ_n be a deterministic subset from n vertices with $|\Lambda_n| \leq \ln(n)$.*

Let $G_n \in \mathcal{G}_{n,d_n}$ be a random regular graph of degree d_n. Then for any $\ell^2(G_n)$-normalized eigenvector ϕ of the operator $\widetilde{A}_n$ the estimate

$$\sum_{x \in \Lambda_n} |\phi(x)|^2 \leq C \left(\frac{\ln(d_n - 1)}{\ln(n)} + \frac{1}{d_n} \right) |\Lambda_n|$$

holds asymptotically almost surely with a uniform constant $C > 0$.

The same estimate holds asymptotically almost surely if one first chooses $G_n \in \mathcal{G}_{n,d}$ at random and then the subset $\Lambda_n \subset G_n$ at random with uniform probability.

Proof of Proposition 5.2. Let $(\varphi_j)_{j=1}^n$ and $(\lambda_j)_{j=1}^n$ denote the eigenvectors and eigenvalues of the operator $\widetilde{A}_n$. Then for any $m \in \{1, \ldots, n\}$ and $\eta > 0$ we estimate

$$|\varphi_m(x)|^2 \leq \eta \sum_{j=1}^n \frac{\eta}{(\lambda_m - \lambda_j)^2 + \eta^2} |\varphi_j(x)|^2 = \eta \Im \widetilde{\Gamma}_n(x, \lambda_m + i\eta) , \qquad (5.6)$$

where in the last step we used (5.1). To derive an upper bound on $\Im\widetilde{\Gamma}_n(x, \lambda_m + i\eta)$ we compare with the Stieltjes transform of the semicircle distribution:

$$
\begin{aligned}
&|\Im\widetilde{\Gamma}_n(x, \lambda_m + i\eta) - \Im\Gamma_{\mathrm{sc}}(\lambda_m + i\eta)| \\
&= \left| \int_{\mathbb{R}} g_{\lambda_m,\eta}(\lambda) d\tilde{\mu}_{n,x}(\lambda) - \int_{\mathbb{R}} g_{\lambda_m,\eta}(\lambda) d\sigma_{\mathrm{sc}}(\lambda) \right| \\
&= \left| \int_{\mathbb{R}} (\tilde{\mu}_{n,x}((-\infty, \lambda]) - \sigma_{\mathrm{sc}}((-\infty, \lambda])) dg_{\lambda_m,\eta}(\lambda) \right| \\
&\leq \int_{\mathbb{R}} |\tilde{\mu}_{n,x}((-\infty, \lambda]) - \sigma_{\mathrm{sc}}((-\infty, \lambda])| \, |g'_{\lambda_m,\eta}(\lambda)| d\lambda.
\end{aligned}
$$

Inserting estimates (4.5), (4.6), and (5.4) yields

$$
|\Im\widetilde{\Gamma}_n(x, \lambda_m + i\eta) - \Im\Gamma_{\mathrm{sc}}(\lambda_m + i\eta)| \leq C\left(\frac{1}{R(x)^*} + \frac{1}{d_n} \right) \int_{\mathbb{R}} |g'_{\lambda_m,\eta}(\lambda)| d\lambda
$$

$$
\leq \frac{2C}{\eta}\left(\frac{1}{R(x)^*} + \frac{1}{d_n} \right)
$$

for all $x \in G_n$ and $\eta > 0$ with a constant $C > 0$ independent of x and n. We combine this bound with (5.5) and obtain

$$
\Im\widetilde{\Gamma}_n(x, \lambda_m + i\eta) \leq 2 + \frac{2C}{\eta}\left(\frac{1}{R(x)^*} + \frac{1}{d_n} \right). \tag{5.7}
$$

Recall the definition of $F_n(k) \subset G_n$ from Lemma 4.2 and the fact that $R(x) \geq k$ for $x \in F_n(k)$. Let us assume that $\Lambda_n \subset F_n(k)$. Under this assumption we combine (5.7) with (5.6) and take the limit $\eta \downarrow 0$. This yields

$$
\sum_{x \in \Lambda_n} |\varphi_m(x)|^2 \leq 2C\left(\frac{1}{k-1} + \frac{1}{d_n} \right) |\Lambda_n| \tag{5.8}
$$

for any eigenfunction φ_m, $m \in \{1, \ldots, n\}$, on the event $\{\Lambda_n \subset F_n(k)\}$.

Now we estimate the probability of this event with the help of Lemma 4.2. We remark that, for $|\Lambda_n| < |F_n(k)|$,

$$
\mathcal{P}_{n,d}[\Lambda_n \subset F_n(k)] = \binom{n - |\Lambda_n|}{|F_n(k)| - |\Lambda_n|} \binom{n}{|F_n(k)|}^{-1}
$$

$$
= \frac{(n - |\Lambda_n|)! \, |F_n(k)|!}{(|F_n(k)| - |\Lambda_n|)! \, n!}.
$$

For a parameter $0 < \tau_n < 1 - |\Lambda_n|/n$ we introduce the event

$$\Omega(\tau_n, k) = \{|F_n(k)| > n(1 - \tau_n)\}$$

and estimate

$$\mathcal{P}_{n,d}\left[\Lambda_n \subset F_n(k)\right] \geq \mathcal{P}_{n,d}\left[\Lambda_n \subset F_n(k) \mid \Omega(\tau_n, k)\right] \mathcal{P}_{n,d}\left[\Omega(\tau_n, k)\right].$$

The first factor is bounded below by

$$\frac{(n - |\Lambda_n|)! \, (n(1 - \tau_n))!}{(n(1 - \tau_n) - |\Lambda_n|)! \, n!} \geq \left(\frac{n(1 - \tau_n) - |\Lambda_n|}{n}\right)^{|\Lambda_n|} = \left(1 - \tau_n - \frac{|\Lambda_n|}{n}\right)^{|\Lambda_n|}$$

and from Lemma 4.2 we obtain that the second factor is bounded below by

$$1 - \frac{1}{2n\tau_n}(d-1)^{2k} \frac{\sqrt{d-1}}{\sqrt{d-1}-1}\left(1 + \frac{8}{n}\left(d + \frac{k}{d}\right)\right)^{2k}.$$

Now we choose τ_n comparable to $1/\sqrt{n}$ and $k = \kappa \log_{d_n-1}(n) + 1$ with $\kappa < 1/4$. Then both lower bounds tend to 1 as $n \to \infty$ and we get

$$\mathcal{P}_{n,d}[\Lambda_n \subset F_n\left(\kappa \log_{d_n-1}(n) + 1\right)] = 1 - o(1)$$

as $n \to \infty$. Inserting this choice of k in the bound (5.8) proves the claim. □

Remark. The same methods yield a similar statement for the adjacency matrix A_n on a random regular graph with fixed degree. However, for this operator better results have recently been derived in [11]: Brooks and Lindenstrauss also investigate delocalization on regular graphs by comparing with the regular tree. They use the explicit representation of spherical generalized eigenfunctions on the tree and estimate the norms of certain propagation operators (also build from orthogonal polynomials). From these estimates they deduce information about eigenvectors on regular graphs. This direct comparison of eigenvectors leads to delocalization bounds that decay not logarithmically but with a small power of n.

Results about the behavior of eigenvectors on regular graphs related to quantum ergodicity have also been obtained in [6], where delocalization is tested by averaging an observable. Both results rely on explicit formulas for eigenfunctions on trees that are not available for random Schrödinger operators.

Let us now study eigenvectors of random Schrödinger operators on random regular graphs with fixed degree d. In this case the analysis is more complicated since the spectral measure of the corresponding infinite volume operator is not purely absolutely continuous: The spectrum of a random Schrödinger operator on

an infinite tree can consist of different components, including absolutely continuous spectrum but also pure-point spectrum. We refer to [47] for an overview of spectral properties of the operator $H_{\mathfrak{T}_d}(V)$, see also Appendix A.2, where we state selected results.

Existence of pure-point spectrum and exponential localization of the corresponding eigenfunctions of $H_{\mathfrak{T}_d}(V)$ was proved in [1]. Therefore one cannot expect that all eigenvectors of the finite volume operator $H_n(V)$ on a random regular graph are delocalized. The existence of absolutely continuous spectrum of $H_{\mathfrak{T}_d}(V)$ was also established, first in [30] and later in [2,26] and the regime where absolutely continuous spectrum can be found was recently extended in [4]. A relevant criterion for absolutely continuous spectrum is positivity of the imaginary part of the Green function. Thus one defines

$$\sigma_{\mathrm{ac}}(H_{\mathfrak{T}_d}) = \{\lambda \in \mathbb{R} : \mathbb{P}[\lim_{\eta \downarrow 0} \Im\Gamma_{\mathfrak{T}_d}(x, \lambda + i\eta; V) > 0] > 0\}.$$

This set is independent of $x \in \mathfrak{T}_d$, deterministic and forms the support of the absolutely continuous component of the spectrum (almost surely with respect to the random potential), see [3,4]. In Theorem 5.3 we show that eigenvectors of the finite volume operator $H_n(V)$ with eigenvalues in $\sigma_{\mathrm{ac}}(H_{\mathfrak{T}_d})$ are typically delocalized for large n.

An important ingredient in the proof of delocalization for the adjacency matrix alone is the fact that the limiting spectral measure is absolutely continuous with uniformly boundedness of the density. This allows for estimate (5.5). For random Schrödinger operators the density of the limiting spectral measure $\mu_{\mathfrak{T}_d,x}$ is given by $\lim_{\eta \downarrow 0} \Im\Gamma_{\mathfrak{T}_d}(x, \lambda + i\eta; V)$. This limit exists almost everywhere and it is finite if λ lies within the absolutely continuous spectrum [4,45]. However, even for compact intervals $I \subset \sigma_{\mathrm{ac}}(H_{\mathfrak{T}_d})$ it is not clear whether $\sup_{\lambda \in I, \eta > 0} \Im\Gamma_{\mathfrak{T}_d}(x, \lambda + i\eta; V)$ has finite expectation. Therefore we can not treat single eigenvectors but we have to select a suitable combination as follows.

For a given realization of the random potential V on the tree $\mathfrak{T}_d$, a graph $G_n \in \mathcal{G}_{n,d}$, and a vertex $x_0 \in G_n$ we identify the potential on the graph with the potential on the tree as described in Section 2.2. Let $(\lambda_j)_{j=1}^n$ and $(\varphi_j)_{j=1}^n$ denote the eigenvalues and corresponding $\ell^2(G_n)$-normalized eigenvectors of the operator $H_n(V)$ and let $I \subset \mathbb{R}$ be bounded and measurable. For $j = 1, 2, \ldots, n$ we define non-negative coefficients

$$c_j(x_0, I) = |\varphi_j(x_0)|^2 \text{ if } \lambda_j \in I \quad \text{and} \quad c_j(x_0, I) = 0 \text{ if } \lambda_j \notin I \qquad (5.9)$$

and we note that $\sum_{j=1}^n c_j(x_0, I) \leq 1$.

In Lemma 5.4 below we prove the following estimate under the assumption $I \subset \sigma_{\mathrm{ac}}(H_{\mathcal{T}_d})$: For all $\hat{x} \in \mathcal{T}_d$ and $\eta > 0$ we have

$$\mathbb{E}\Big[\sum_{j=1}^{n} c_j(x_0, I) \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta; V) \Big] \leq C|I| \qquad (5.10)$$

with a constant $C > 0$ depending only on the degree d and on the density of the random potential. In the proof of Theorem 5.3 we use this estimate in the same way as we used (5.5) in the proof of Proposition 5.2.

We remark that an eigenvector φ_j of $H_n(V)$ that is localized close to the vertex x_0 leads to a coefficient c_j of order 1. The following result shows that this can not happen for large n for eigenvectors with eigenvalues within the absolutely continuous spectrum of $H_{\mathcal{T}_d}$. Recall that we write $B_r(x_0) = \{y \in G_n : d(x_0, y) \leq r\}$ and that $|B_r(x_0)|$ denotes the number of vertices in this neighborhood.

Theorem 5.3. *Assume that the density of the random potential is bounded and has bounded support. Let $I \subset \sigma_{\mathrm{ac}}(H_{\mathcal{T}_d})$ be bounded and measurable. For a random regular graph $G_n \in \mathcal{G}_{n,d}$ choose a vertex $x_0 \in G_n$ at random with uniform probability. Let $(\varphi_j)_{j=1}^{n}$ denote the $\ell^2(G_n)$-normalized eigenvectors of $H_n(V)$ and let $c_j(x_0, I)$ be as defined in (5.9).*

Then for any sequence $(\varepsilon_n)_{n \in \mathbb{N}}$ of positive numbers with $\varepsilon_n = o(1)$ as $n \to \infty$ the estimate

$$\sum_{x \in B_r(x_0)} \sum_{j=1}^{n} c_j(x_0, I)|\varphi_j(x)|^2 \leq \frac{|B_r(x_0)|}{\varepsilon_n \sqrt{\log_{d-1}(n)}}$$

holds asymptotically almost surely (with respect to the random potential V and the choice of graph $G_n \in \mathcal{G}_{n,d}$ and vertex $x_0 \in G_n$) for all $r \leq \ln(\ln(n))$.

Proof. First we fix a graph G_n and a vertex $x_0 \in G_n$ and we choose $x \in B_r(x_0)$. For a given realization of the potential V on the tree $\mathcal{T}_d$ we consider the corresponding potential on the graph G_n as explained in Section 2.2.

As in (5.6), we have for all $j \in \{1, \ldots, n\}$ and all $\eta > 0$

$$|\varphi_j(x)|^2 \leq \eta \Im \Gamma(x, \lambda_j + i\eta; V)$$

and therefore

$$\mathbb{E}\Big[\sum_{j=1}^{n} c_j |\varphi_j(x)|^2 \Big] \leq \eta \, \mathbb{E}\Big[\sum_{j=1}^{n} c_j \Im \Gamma_n(x, \lambda_j + i\eta; V) \Big] \qquad (5.11)$$

for all $\eta > 0$. Here and in the remainder of the proof we write $c_j = c_j(x_0, I)$ for short.

To derive an upper bound on $\Im\Gamma_n(x, \lambda_j + i\eta; V)$ we compare with the Green function on the tree $\mathfrak{T}_d$. Consider the map $\iota_{\hat{x}_0}$ given in (2.14) and let $\hat{x} \in \mathfrak{T}_d$ be the preimage of $x \in G_n$ under this map. Recall the definition of $g_{E,\eta}$ from (5.3). For $\eta > 0$ and any eigenvalue $\lambda_j \in I$ we have

$$
\begin{aligned}
&\left| \Im\Gamma_n(x, \lambda_j + i\eta; V) - \Im\Gamma_{\mathfrak{T}_d}(\hat{x}, \lambda_j + i\eta; V) \right| \\
&= \left| \int_{\mathbb{R}} g_{\lambda_j,\eta}(\lambda)\mu_{n,x}(d\lambda; V) - \int_{\mathbb{R}} g_{\lambda_j,\eta}(\lambda)\mu_{\mathfrak{T}_d,\hat{x}}(d\lambda; V) \right| \\
&\leq \int_{\mathbb{R}} \left| \mu_{n,x}((-\infty, \lambda]; V) - \mu_{\mathfrak{T}_d,\hat{x}}((-\infty, \lambda]; V) \right| \sup_{\xi \in I} |g'_{\xi,\eta}(\lambda)| d\lambda.
\end{aligned}
$$

Note that a coefficient c_j is non-zero only if the corresponding eigenvalue λ_j lies in I and that the sum of the coefficients c_j is bounded by one. Hence, we find

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{j=1}^{n} c_j |\Im\Gamma_n(x, \lambda_j + i\eta; V) - \Im\Gamma_{\mathfrak{T}_d}(\hat{x}, \lambda_j + i\eta; V)| \right] \\
&\leq \int_{\mathbb{R}} \mathbb{E}[|\mu_{n,x}((-\infty, \lambda]; V) - \mu_{\mathfrak{T}_d,\hat{x}}((-\infty, \lambda]; V)|] \sup_{\xi \in I} |g'_{\xi,\eta}(\lambda)| d\lambda
\end{aligned}
$$

and Theorem 3.1 yields the upper bound

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{j=1}^{n} c_j |\Im\Gamma_n(x, \lambda_j + i\eta; V) - \Im\Gamma_{\mathfrak{T}_d}(\hat{x}, \lambda_j + i\eta; V)| \right] \\
&\leq 2\pi \|\rho\|_\infty (2\sqrt{d-1} + \rho_0) \frac{1}{R(x)^*} \int_{\mathbb{R}} (\sup_{\xi \in I} |g'_{\xi,\eta}(\lambda)|) d\lambda \\
&\leq C\left(1 + \frac{|I|}{\eta} \right) \frac{1}{\eta R^*(x)},
\end{aligned}
\tag{5.12}
$$

where we used the fact that the integral of the supremum is bounded by a constant times $\eta^{-1}(1 + |I|\eta^{-1})$. Here ρ denotes the density of the random potential and we use that $\|\rho\|_\infty < \infty$ and that $\operatorname{supp}\rho = [-\rho_0, \rho_0]$ with $\rho_0 < \infty$. To shorten notation C denotes various positive constants that may depend only on d and ρ.

Combining (5.11) and (5.12) with (5.10) yields, for $\eta > 0$ small enough,

$$\mathbb{E}\Big[\sum_{j=1}^{n} c_j |\varphi_j(x)|^2\Big] \le C\Big(|I|\eta + \frac{1}{R(x)^*} + \frac{|I|}{\eta R(x)^*}\Big) \le C|I|\Big(\eta + \frac{1}{\eta R(x)^*}\Big).$$

Applying the Markov inequality we deduce that for any $0 < \delta_n < 1$

$$\mathbb{P}\Big[\sum_{x \in B_r(x_0)} \sum_{j=1}^{n} c_j |\varphi_j(x)|^2 \le \frac{C|I|}{\delta_n} \sum_{x \in B_r(x_0)} \Big(\eta + \frac{1}{\eta R(x)^*}\Big)\Big] \ge 1 - \delta_n. \quad (5.13)$$

It remains to estimate $R(x)^*$. Recall the definition of $F_n(k) \subset G_n$ from Lemma 4.2 and assume that $x_0 \in F_n(k)$. Since x_0 is chosen from G_n at random with uniform probability we can estimate the probability of this event with the help of Lemma 4.2. Indeed, we have, for any $0 < \tau_n < 1$,

$$\mathcal{P}_{n,d}[x_0 \in F_n(k)]$$
$$\ge \mathcal{P}_{n,d}[x_0 \in F_n(k) \mid |F_n(k)| \ge n(1 - \tau_n)]\mathcal{P}_{n,d}[|F_n(k)| \ge n(1 - \tau_n)]$$
$$\ge (1 - \tau_n)\mathcal{P}_{n,d}[|F_n(k)| \ge n(1 - \tau_n)].$$

By Lemma 4.2 the latter probability is bounded below by

$$1 - \frac{1}{2n\tau_n}(d-1)^{2k}\frac{\sqrt{d-1}}{\sqrt{d-1}-1}\Big(1 + \frac{8}{n}\Big(d + \frac{k}{d}\Big)\Big)^{2k}.$$

Let us now choose $k = \kappa \log_{d-1}(n) + 1$ with $\kappa < 1/4$ and τ_n comparable to $1/\sqrt{n}$. Then the lower bound tends to 1 as $n \to \infty$ and we get

$$\mathcal{P}_{n,d}[x_0 \in F_n(\kappa \log_{d-1}(n) + 1)] = 1 - o(1) \quad (5.14)$$

as $n \to \infty$.

By assumption, $r \le \ln(\ln(n)) < \kappa \log_{d-1}(n) + 1$ for n large enough. Hence, for $x \in B_r(x_0)$ we have

$$B_{\kappa \log_{d-1}(n)+1-r}(x) \subset B_{\kappa \log_{d-1}(n)+1}(x_0).$$

Thus $x_0 \in F_n(\kappa \log_{d-1}(n) + 1)$ implies that these neighborhoods are acyclic. In particular we find $x \in F_n(\kappa \log_{d-1}(n) - r + 1)$ so that $R(x)^* \ge \kappa \log_{d-1}(n) - r$.

We choose $\eta = (\kappa \log_{d-1}(n) - r)^{-1/2}$ and arrive at

$$\sum_{x \in B_r(x_0)} \left(\eta + \frac{1}{\eta R(x)^*} \right) \leq \frac{2|B_r(x_0)|}{\sqrt{\kappa \log_{d-1}(n) - r}} \leq \frac{4|B_r(x_0)|}{\sqrt{\kappa \log_{d-1}(n)}}$$

for n large enough.

With this choice of parameters (5.13) reads as

$$\mathbb{P}\Big[\sum_{x \in B_r(x_0)} \sum_{j=1}^{n} c_j |\varphi_j(x)|^2 \leq \frac{C|I|}{\delta_n} \frac{|B_r(x_0)|}{\sqrt{\kappa \log_{d-1}(n)}} \Big] \geq 1 - \delta_n$$

and this estimate is valid for n large enough on the event

$$\{x_0 \in F_n(\kappa \log_{d-1}(n) + 1)\}.$$

Finally, we choose δ_n comparable to $C|I|\varepsilon_n/\sqrt{\kappa}$ such that

$$\mathbb{P}\Big[\sum_{x \in B_r(x_0)} \sum_{j=1}^{n} c_j |\varphi_j(x)|^2 \leq \frac{|B_r(x_0)|}{\varepsilon_n \sqrt{\log_{d-1}(n)}} \Big] \geq 1 - \frac{C|I|}{\sqrt{\kappa}} \varepsilon_n \qquad (5.15)$$

for n large enough on the event

$$\{x_0 \in F_n(\kappa \log_{d-1}(n) + 1)\}.$$

Relations (5.14) and (5.15) show that the claimed estimate holds asymptotically almost surely and the proof is complete.   □

The proof of Theorem 5.3 relies on the following estimate for the Green function on the tree. For this estimate it is essential that $I \subset \sigma_{\mathrm{ac}}(H_{\mathcal{T}_d})$.

Lemma 5.4. *Under the conditions of Theorem* 5.3 *the bound* (5.10) *holds for all* $\hat{x} \in \mathcal{T}_d$ *and* $\eta > 0$ *with a constant* $C > 0$ *depending only on the degree d and on the density of the random potential.*

Proof. The proof is based on estimate (A.3), the fact that within the absolutely continuous spectrum the imaginary part of the Green function has finite inverse moments. This was recently proved in [3]. To apply this result we rely on recursion properties of the Green function on trees and on results from rank-one perturbation theory.

We fix a graph G_n and a vertex $x_0 \in G_n$. Recall the definition of the coefficients $c_j(x_0, I)$ from (5.9). Again we write $c_j = c_j(x_0, I)$ for short. By (2.10) we have, for any $\hat{x} \in \mathcal{T}_d$,

$$\sum_{j=1}^{n} c_j \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta; V) = \sum_{\lambda_j \in I} |\varphi_j(x_0)|^2 \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta; V)$$
$$= \int_I \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda + i\eta; V) \mu_{n,x_0}(d\lambda; V).$$

We use the Stieltjes inversion formula, see for example [45, Theorem 3.21], to write the spectral measure μ_{n,x_0} in terms of the Green function and we obtain

$$\sum_{j=1}^{n} c_j \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta; V)$$
$$= \lim_{\epsilon \downarrow 0} \frac{1}{\pi} \int_I \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda + i\eta; V) \Im \Gamma_n(x_0, \lambda + i\epsilon; V) \, d\lambda. \tag{5.16}$$

To estimate the expectation of the right-hand side, we use the following results from rank-one perturbation theory [1, 15, 40]. This allows to analyze the dependence on the single random variable ω_{x_0}.

First we rewrite $\Im \Gamma_n(x_0, \lambda + i\epsilon; V)$: For a realization of the random potential $V = (\omega_x)_{x \in G_n}$ we denote by $\widehat{V}$ the same collection of random variables with ω_{x_0} replaced by zero. Then

$$H_n(V) = H_n(\widehat{V}) + \omega_{x_0} \delta_{x_0},$$

where $\delta_{x_0}(x_0) = 1$ and $\delta_{x_0}(x) = 0$ for $x \neq x_0$. The resolvent identity yields, for $z \in \mathbb{C}_+$,

$$\frac{1}{H_n(V) - z} - \frac{1}{H_n(\widehat{V}) - z} = \frac{1}{H_n(V) - z} \omega_{x_0} \delta_{x_0} \frac{1}{H_n(\widehat{V}) - z}$$

and thus

$$\Gamma_n(x_0, z; V) = \frac{\Gamma_n(x_0, z; \widehat{V})}{1 + \omega_{x_0} \Gamma_n(x_0, z; \widehat{V})} = \frac{1}{\omega_{x_0} - \Xi_{x_0}(z)}$$

with

$$\Xi_{x_0}(z) = -\Gamma_n(x_0, z; \widehat{V})^{-1}.$$

It follows that

$$\Im\Gamma_n(x_0, \lambda + i\epsilon; V) = \frac{\Im\Xi_{x_0}(\lambda + i\epsilon)}{(\omega_{x_0} - \Re\Xi_{x_0}(\lambda + i\epsilon))^2 + \Im\Xi_{x_0}(\lambda + i\epsilon)^2}. \qquad (5.17)$$

We emphasize that Ξ_{x_0} is independent of ω_{x_0}. We also note that the limit

$$\Xi_{x_0}(\lambda) = \lim_{\epsilon \downarrow 0} \Xi_{x_0}(\lambda + i\epsilon)$$

exists almost everywhere with respect to Lebesgue measure, see for example [45, Theorem 3.23].

Next, we estimate $\Im\Gamma_{\mathcal{T}_d}(\hat{x}, \lambda + i\eta; V)$: If we remove the vertex $\hat{x}$ from the tree $\mathcal{T}_d$, it is decomposed into d disjoint infinite rooted trees that are rooted at the the nearest neighbors of $\hat{x}$. (We remark that these trees are no longer regular, since the degree at the root equals $d - 1$.) Let

$$N(\hat{x}) = \{y \in \mathcal{T}_d : d(\hat{x}, y) = 1\}$$

denote the set of nearest neighbors of $\hat{x}$. For $y \in N(\hat{x})$, let T_y denote the rooted tree with root at y. In the same way as on the regular tree we define the bounded self-adjoint operators $H_{T_y}(V)$ with domain $\ell^2(T_y)$. For $u \in T_y$ and $z \in \mathbb{C}_+$ we write

$$\Gamma_{T_y}(u, z; V) = (\delta_u, (H_{T_y}(V) - z)^{-1}\delta_u)_{\ell^2(T_y)}$$

for the respective Green function. The trees T_y with $y \in N(\hat{x})$ are not connected to each other, hence the random variables $\Gamma_{T_y}(y, z; V)$, $y \in N(\hat{x})$, are independent.

Employing the resolvent equation one can derive the recursion formula, see for example [4],

$$\Gamma_{\mathcal{T}_d}(\hat{x}, z; V) = \frac{1}{\omega_{\hat{x}} - z - \sum_{y \in N(\hat{x})} \Gamma_{T_y}(y, z; V)}.$$

Taking the imaginary part, we see that for all $z \in \mathbb{C}_+$

$$\Im\Gamma_{\mathcal{T}_d}(\hat{x}, z; V) = |\Gamma_{\mathcal{T}_d}(\hat{x}, z; V)|^2 \Big(\sum_{y \in N(\hat{x})} \Im\Gamma_{T_y}(y, z; V) + \Im z \Big).$$

Applying the recursion formula one also gets that

$$|\Gamma_{\mathcal{T}_d}(\hat{x}, z; V)|^2 \leq \frac{1}{(\Im(\omega_{\hat{x}} - z - \sum_{y \in N(\hat{x})} \Gamma_{T_y}(y, z; V)))^2}$$

$$= \frac{1}{(\sum_{y \in N(\hat{x})} \Im\Gamma_{T_y}(y, z; V) + \Im z)^2}$$

and we obtain

$$\Im\Gamma_{\mathcal{T}_d}(\hat{x}, z; V) \leq \frac{1}{\sum_{y \in N(\hat{x})} \Im\Gamma_{T_y}(y, z; V)} . \tag{5.18}$$

Let again $\hat{x}_0 \in \mathcal{T}_d$ denote the vertex corresponding to $x_0 \in G_n$ under the universal cover and let us for the moment assume that $\hat{x} \neq \hat{x}_0$. Then there is one unique vertex $y^* \in N(\hat{x})$ such that the corresponding rooted tree T_{y^*} contains $\hat{x}_0$. Let

$$N^+(\hat{x}) = \{y \in N(\hat{x}) : \hat{x}_0 \notin T_y\}$$

denote the subset of all other nearest neighbors of $\hat{x}$. In particular, we find that $N^+(\hat{x})$ contains $d-1$ vertices. For $y \in N^+(\hat{x})$ the tree T_y does not contain $\hat{x}_0$, thus the collection of random variables

$$(\Gamma_{T_y}(y, \lambda + i\eta; V))_{y \in N^+(\hat{x})}$$

is independent of $\omega_{\hat{x}_0}$, the value of the random potential V at $\hat{x}_0$. If $\hat{x} = \hat{x}_0$, then all random variables $\Gamma_{T_y}(y, \lambda + i\eta, V)$ with $y \in N(\hat{x})$ are independent of $\omega_{\hat{x}_0}$ and we can continue the proof in the same way with $N^+(\hat{x})$ replaced by $N(\hat{x})$.

We note that, for $z \in \mathbb{C}_+$, the imaginary part of the Green function is positive. Hence, in view of (5.18) we have

$$\Im\Gamma_{\mathcal{T}_d}(\hat{x}, z; V) \leq \frac{1}{\sum_{y \in N^+(\hat{x})} \Im\Gamma_{T_y}(y, z; V)} . \tag{5.19}$$

Combining (5.16), (5.17), and (5.19) we arrive at the bound

$$\sum_{j=1}^{n} c_j \Im\Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta; V)$$

$$\leq \lim_{\epsilon \downarrow 0} \frac{1}{\pi} \int_I S_1 S_2 d\lambda,$$

where

$$S_1 = \frac{1}{\sum_{y \in N^+(\hat{x})} \Im \Gamma_{T_y}(y, \lambda + i\eta; V)}$$

and

$$S_2 = \frac{\Im \Xi_{x_0}(\lambda + i\epsilon)}{(\omega_{x_0} - \Re \Xi_{x_0}(\lambda + i\epsilon))^2 + \Im \Xi_{x_0}(\lambda + i\epsilon)^2}$$

Recall that $\omega_{\hat{x}_0} = \omega_{x_0}$ and that the random variables $\Xi_{x_0}(\lambda + i\epsilon)$ and $\Gamma_{T_y}(y, \lambda + i\eta; V)$ with $y \in N^+(\hat{x})$ are independent of $\omega_{\hat{x}_0}$. Hence, we condition on the random potential $V = (\omega_y)_{y \in \mathcal{T}_d}$ at all other vertices and take the expectation with respect to $\omega_{\hat{x}_0}$ only. By dominated convergence, we find

$$\mathbb{E}\Big[\sum_{j=1}^{n} c_j \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta; V) \,\Big|\, (\omega_y)_{y \neq \hat{x}_0} \Big]$$
$$\leq \lim_{\epsilon \downarrow 0} \frac{1}{\pi} \int_{\mathbb{R}} \int_I S_1 S_3 d\lambda \rho(v) dv,$$

where

$$S_3 = \frac{\Im \Xi_{x_0}(\lambda + i\epsilon)}{(v - \Re \Xi_{x_0}(\lambda + i\epsilon))^2 + \Im \Xi_{x_0}(\lambda + i\epsilon)^2}.$$

Now we estimate $\rho(v) \leq \|\rho\|_\infty$ and apply Fubini's theorem to perform the integration in v. Since

$$\int_{\mathbb{R}} \frac{\Im \Xi_{x_0}(\lambda + i\epsilon)}{(v - \Re \Xi_{x_0}(\lambda + i\epsilon))^2 + \Im \Xi_{x_0}(\lambda + i\epsilon)^2} dv = \pi$$

we obtain

$$\mathbb{E}\Big[\sum_{j=1}^{n} c_j \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta; V) \,\Big|\, (\omega_y)_{y \neq \hat{x}_0} \Big]$$
$$\leq \|\rho\|_\infty \int_I \frac{1}{\sum_{y \in N^+(\hat{x})} \Im \Gamma_{T_y}(y, \lambda + i\eta; V)} \, d\lambda.$$

Taking now the expectation with respect to $(\omega_y)_{y \neq \hat{x}_0}$ yields

$$\mathbb{E}\Big[\sum_{j=1}^{n} c_j \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta, V) \Big]$$
$$\leq \|\rho\|_\infty \int_I \mathbb{E}\Big[\frac{1}{\sum_{y \in N^+(\hat{x})} \Im \Gamma_{T_y}(y, \lambda + i\eta, V)} \Big] d\lambda.$$

Finally, we use that the set $N^+(x)$ contains $d - 1$ elements. The inequality of arithmetic and geometric means tells us that

$$\sum_{y \in N^+(\hat{x})} \Im \Gamma_{T_y}(y, \lambda + i\eta; V) \geq (d-1) \prod_{y \in N^+(\hat{x})} \Im \Gamma_{T_y}(y, \lambda + i\eta; V)^{1/(d-1)}$$

and using the fact that the random variables $\Im \Gamma_{T_y}(y, \lambda + i\eta; V)$, $y \in N^+(\hat{x})$, are independent we conclude

$$\mathbb{E}\Big[\sum_{j=1}^{n} c_j \Im \Gamma_{\mathcal{T}_d}(\hat{x}, \lambda_j + i\eta; V) \Big]$$

$$\leq \frac{\|\rho\|_\infty}{d-1} \int_I \prod_{y \in N^+(\hat{x})} \mathbb{E}[\Im \Gamma_{T_y}(y, \lambda + i\eta; V)^{-1/(d-1)}] d\lambda.$$

Hence, applying (A.3) yields the claim of the lemma.   □

Appendix A. Auxiliary results

In the appendix we collect some known results that are used in the previous sections. We indicate where proofs can be found in the literature.

A.1. Orthonormal polynomials and Gaussian quadrature.

In Section 3 we repeatedly used the following quadrature formula for polynomials. A proof and further references can be found for example in [5, Chapter 1.4.1].

Lemma A.1. *Let σ be a measure on the real line with finite moments and let $(P_n)_{n \in \mathbb{N}_0}$ denote the orthonormal polynomials with respect to σ. For arbitrary $M \in \mathbb{N}$ and $s \in \mathbb{R}$ set $\hat{P}_{M+1} = P_{M+1} + s P_M$ and let $\lambda_1 < \lambda_2 < \cdots < \lambda_{M+1}$ denote the zeros of $\hat{P}_{M+1}$.*

Then the identity

$$\int_{\mathbb{R}} R(\lambda) \sigma(d\lambda) = \sum_{k=1}^{M+1} \frac{R(\lambda_k)}{\sum_{n=0}^{M} P_n(\lambda_k)}$$

holds for any polynomial R of degree less or equal than $2M$.

A.2. Spectrum and Green function on the infinite tree. Here we mention some results about the spectrum of the random Schrödinger operator $H_{\mathcal{T}_d}(V)$ defined in Section 2.2 and about the Green function $\Gamma_{\mathcal{T}_d}(x, z; V)$ defined in Section 5.1. We refer to the books [12, 29, 38, 45] for more information and further references.

Recall that the random potential V is defined as a multiplication operator

$$(V\phi)(x) = \omega_x \phi(x), \quad \phi \in \ell^2(\mathcal{T}_d), \ x \in \mathcal{T}_d,$$

where $(\omega_x)_{x \in \mathcal{T}_d}$ are independent identically distributed real random variables with density ρ. Hence one can refer to the theory of ergodic operators to determine the spectrum of $H_{\mathcal{T}_d}$. In [28, 32] it is shown that the spectrum corresponds almost surely to the set-sum of the spectrum of the adjacency matrix and the support of ρ. On the tree the spectrum of $A_{\mathcal{T}_d}$ is given by $(-2\sqrt{d-1}, 2\sqrt{d-1})$. So under the assumption $\text{supp}(\rho) = [-\rho_s, \rho_s]$ with $\rho_s < \infty$ the spectrum of $H_{\mathcal{T}_d}(V)$ is almost surely given by the deterministic set $[-2\sqrt{d-1} - \rho_s, 2\sqrt{d-1} + \rho_s]$. In particular, the spectral measure $\mu_{\mathcal{T}_d, x}$ satisfies, for all $x \in \mathcal{T}_d$,

$$\text{supp}(\mu_{\mathcal{T}_d, x}) = [-2\sqrt{d-1} - \rho_s, 2\sqrt{d-1} + \rho_s] \tag{A.1}$$

almost surely and this implies

$$\text{supp}(\sigma_\rho) = [-2\sqrt{d-1} - \rho_s, 2\sqrt{d-1} + \rho_s].$$

It was noticed by Wegner [48] that regularity of the distribution of ω_x implies regularity of the density of states measure: Under the assumption $\|\rho\|_\infty < \infty$ using spectral averaging one can show that

$$\left\| \frac{d\sigma_\rho}{d\lambda} \right\|_\infty \leq \|\rho\|_\infty. \tag{A.2}$$

Along with the spectrum also the spectral components, the pure-point spectrum, the singular continuous spectrum, and the absolutely continuous spectrum form almost surely deterministic sets. It is the subject of extensive research to determine the location of these spectral components and we refer to [47] for an overview of results and further references.

One useful criterion for absolutely continuous spectrum is that the imaginary part of the green function does not vanish. So one considers the set

$$\sigma_{\text{ac}}(H_{\mathcal{T}_d}) = \left\{ \lambda \in \mathbb{R} \colon \mathbb{P}\left[\lim_{\eta \downarrow 0} \Im \Gamma_{\mathcal{T}_d}(x, \lambda + i\eta; V) > 0 \right] > 0 \right\}$$

that also forms a deterministic set that does not depend on $x \in \mathcal{T}_d$. For almost every realization of the randomness $\sigma_{\mathrm{ac}}(H_{\mathcal{T}_d})$ is the support of the absolutely continuous component of the spectrum [3, 4]. Within this set the imaginary part of the Green function has finite inverse moments. This fact was recently established in [3, Theorem 2.4]: Let $I \subset \sigma_{\mathrm{ac}}(H_{\mathcal{T}_d})$ be a bounded and measurable set. Consider an infinite rooted tree T and let 0 denote the vertex at the root. Then there is a $\delta > 0$ such that the estimate

$$\operatorname*{ess\,sup}_{\lambda \in I, \eta > 0} \mathbb{E}[(\Im \Gamma_T(0, \lambda + i\eta; V))^{-3-\delta}] < \infty \qquad (A.3)$$

holds.

References

[1] M. Aizenman, Localization at weak disorder: some elementary bounds. *Rev. Math. Phys.* **6** (1994), no. 5A, 1163–1182. Special issue dedicated to E. H. Lieb. MR 1301371 Zbl 0843.47039

[2] M. Aizenman, R. Sims, and S. Warzel, Stability of the absolutely continuous spectrum of random *Probab. Theory Related Fields* **136** (2006), no. 3, 363–394. MR 2257129 Zbl 1169.82007

[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** (2012), no. 9, article id. 095205. MR 2905787 Zbl 1278.81090

[4] M. Aizenman and S. Warzel, Resonant delocalization for random Schrödinger operators on tree graphs. *J. Eur. Math. Soc.* (*JEMS*) **15** (2013), no. 4, 1167–1222. MR 3055759 Zbl 1267.47064

[5] N. I. Akhiezer, *The classical moment problem and some related questions in analysis.* Translated by N. Kemmer. Hafner, New York, 1965. MR 0184042 Zbl 0135.33803

[6] N. Anantharaman and E. Le Masson, Quantum ergodicity on large regular graphs. *Duke Math. J.* **164** (2015), no. 4, 723–765. MR 3322309 Zbl 06434640

[7] Greg W. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices.* Cambridge Studies in Advanced Mathematics, 118. Cambridge University Press, Cambridge, 2010. MR 2760897 Zbl 1184.15023

[8] Z. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices.* 2nd ed. Springer Series in Statistics. Springer, New York, 2010. MR 2567175 Zbl 1301.60002

[9] G. Biroli, A. C. Ribeiro-Teixeira, and M. Tarzia, Difference between level statistics, ergodicity and localization transitions on the Bethe lattice. Preprint 2012. arXiv:1211.7334 [cond-mat.dis-nn]

[10] B. Bollobás, *Random graphs.* 2nd ed. Cambridge Studies in Advanced Mathematics, 73. Cambridge University Press, Cambridge, 2001. MR 1864966 Zbl 0979.05003

[11] S. Brooks and E. Lindenstrauss, Non-localization of eigenfunctions on large regular graphs. *Israel J. Math.* **193** (2013), no. 1, 1–14. MR 3038543 Zbl 1317.05110

[12] R. Carmona and J. Lacroix, *Spectral theory of random Schrödinger operators.* Probability and its Applications. Birkhäuser Boston, Boston, MA, 1990. MR 1102675 Zbl 0717.60074

[13] P. Deift, Universality for mathematical and physical systems. In M. Sanz-Solé, J. Soria, J. L. Varona, and J. Verdera (eds.), *International Congress of Mathematicians.* Vol. I. Plenary lectures and ceremonies. Proceedings of the congress held in Madrid, August 22–30, 2006. European Mathematical Society, Zürich, 2007, 125–152. MR 2334189 MR 2334180 (collection) Zbl 1149.82012 Zbl 1111.00009 (collection)

[14] A. De Luca and A. Scardicchio, Ergodicity breaking in a model showing many-body localization. *Europhysics Letters* **101** (2013), no. 3, article id. 37003.

[15] F. Delyon, Y. Lévy, and B. Souillard, Anderson localization for multidimensional systems at large disorder or large energy. *Comm. Math. Phys.* **100** (1985), no. 4, 463–470. MR 0806247 Zbl 0576.60053

[16] I. Dumitriu and S. Pal, Sparse regular random graphs: Spectral density and eigenvectors. *Ann. Probab.* **40** (2012), no. 5, 2197–2235. MR 3025715 Zbl 1255.05173

[17] Y. Elon, Eigenvectors of the discrete Laplacian on regular graphs – a statistical approach. J. Phys. A **41** (2008), no. 43, *J. Phys. A* **41** (2008), no. 43, article id. 435203. MR 2453165 Zbl 1157.39314

[18] L. Erdős, A. Knowles, H. T. Yau, and J. Yin, Spectral statistics of Erdős–Rényi graphs II: Eigenvalue spacing and the extreme eigenvalues. *Comm. Math. Phys.* **314** (2012), no. 3, 587–640. MR 2964770 Zbl 1251.05162

[19] L. Erdős, A. Knowles, H. T. Yau, and J. Yin, Spectral statistics of Erdős–Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** (2013), no. 3B, 2279–2375. MR 3098073 Zbl 1272.05111

[20] L. Erdős, S. Péché, J. A. Ramírez, B. Schlein, and H.-T. Yau, Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63** (2010), no. 7, 895–925. MR 2662426 Zbl 1216.15025

[21] L. Erdős, J. Ramírez, B. Schlein, T. Tao, V. Vu, and H.-T. Yau, Bulk universality for Wigner Hermitian matrices with subexponential decay. *Math. Res. Lett.* **17** (2010), no. 4, 667–674. MR 2661171 Zbl 1277.15027

[22] L. Erdős, H.-T. Yau, and J. Yin, Bulk universality for generalized Wigner matrices. *Probab. Theory Related Fields* **154** (2012), no. 1-2, 341–407. MR 2981427 Zbl 1277.15026

[23] A. Figà-Talamanca and C. Nebbia, *Harmonic analysis and representation theory for groups acting on homogeneous trees.* London Mathematical Society Lecture Note Series, 162. Cambridge University Press, Cambridge, 1991. MR 1152801 Zbl 1154.22301

[24] J. Friedman, A proof of Alon's second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.* **195** (2008), no. 910. MR 2437174 Zbl 1177.05070

[25] R. Froese, Spectral concentration estimates for trees. *Oberwolfach Reports* **8** (2011), no. 4, 2850–2851.

[26] R. Froese, D. Hasler, and W. Spitzer, Absolutely continuous spectrum for the Anderson model on a tree: a geometric proof of Klein's theorem. *Comm. Math. Phys.* **269** (2007), no. 1, 239–257. MR 2274470 Zbl 1117.82024

[27] D. Jakobson, S. D. Miller, I. Rivin, and Z. Rudnick, Eigenvalue spacings for regular graphs. In D. A. Hejhal, J. Friedman, M. C. Gutzwiller and A. M. Odlyzko (eds.), *Emerging applications of number theory.* Papers from the IMA Summer Program held at the University of Minnesota, Minneapolis, MN, July 15–26, 1996, 317–327. MR 1691538 MR 1691526 (collection) Zbl 1071.81522 Zbl 0919.00047 (collection)

[28] W. Kirsch and F. Martinelli, On the ergodic properties of the spectrum of general random operators. *J. Reine Angew. Math.* **334** (1982), 141–156. MR 0667454 Zbl 0476.60058

[29] W. Kirsch, Random Schrödinger operators. A course. In H. Holden and A. Jensen (eds.), *Schrödinger operators.* Proceedings of the Nordic Summer School in Mathematics held in Sønderborg, August 1–12, 1988. Lecture Notes in Physics, 345. Springer-Verlag, Berlin, 1989, 264–370. MR 1037323 MR 1037314 (collection) Zbl 0712.35070 Zbl 0686.00012 (collection)

[30] A. Klein, Extended states in the Anderson model on the Bethe lattice. *Adv. Math.* **133** (1998), no. 1, 163–184. MR 1492789 Zbl 0899.60088

[31] M. G. Kreĭn and A. A. Nudelman, *The Markov moment problem and extremal problems.* Ideas and problems of P. L. Čebyšev and A. A. Markov and their further development. Translated from the Russian by D. Louvish. Translations of Mathematical Monographs, 50. American Mathematical Society, Providence, R.I., 1977. MR 0458081 Zbl 0361.42014

[32] H. Kunz and B. Souillard, Sur le spectre des opérateurs aux différences finies aléatoires. *Comm. Math. Phys.* **78** (1980/81), no. 2, 201–246. MR 0597748 Zbl 0449.60048

[33] B. D. McKay, The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.* **40** (1981), 203–216. MR 0629617 Zbl 0468.05039

[34] B. D. McKay, Subgraphs of random graphs with specified degrees. In F. Hoffman, K. B. Reid, K.Brooks, R. C. Mullin, and R. G. Stanton (eds.), *Proceedings of the Twelfth Southeastern Conference on Combinatorics, Graph Theory and Computing.* Vol. II. (Louisiana State University, Baton Rouge, March 2-5, 1981.) Congressus Numerantium, 33. Winnipeg, Utilitas Mathematica, Winnipeg, 213–223. MR 0681916 Zbl 0492.05038 Zbl 0483.00001 (collection)

[35] B. D. McKay, N. C. Wormald, and B. Wysocka, Short cycles in random regular graphs. *Electron. J. Combin.* **11** (2004), no. 1, Research Paper 66. MR 2097332 Zbl 1063.05122

[36] M. L. Mehta, *Random matrices.* 3rd ed. Pure and Applied Mathematics (Amsterdam), 142. Elsevier/Academic Press, Amsterdam, 2004. MR 2129906 Zbl 1107.15019

[37] I. Oren and U. Smilansky, Trace formulas and spectral statistics for discrete Laplacians on regular graphs (II). *J. Phys. A* **43** (2010), no. 22, article id. 225205. MR 2647344 Zbl 1190.81052

[38] L. Pastur and A. Figotin, *Spectra of random and almost-periodic operators.* Grundlehren der Mathematischen Wissenschaften, 297. Springer, Berlin etc., 1992. MR 1223779 Zbl 0752.47002

[39] J. Salez, *Quelques conséquences de la convergence locale faible pour les graph aléatoires.* Ph.D. Thesis. Université Pierre et Marie Curie, Paris, 2011. http://opac.inria.fr/record=b1133316

[40] B. Simon and T. Wolff, Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians. *Comm. Pure Appl. Math.* **39** (1986), no. 1, 75–90. MR 0820340 Zbl 0609.47001

[41] S. Sodin, Random matrices, nonbacktracking walks, and orthogonal polynomials. *J. Math. Phys.* **48** (2007), no. 12, article id. 123503. MR 2377835 Zbl 1153.81436

[42] S. Sodin, The Tracy–Widom law for some sparse random matrices. *J. Stat. Phys.* **136** (2009), no. 5, 834–841. MR 2545551 Zbl 1177.82066

[43] G. Szegő, *Orthogonal polynomials.* 4th ed. American Mathematical Society, Colloquium Publications, XXIII. American Mathematical Society, Providence, R.I., 1975. MR 0372517 Zbl 0305.42011

[44] T. Tao and V. Vu, Random matrices: universality of local eigenvalue statistics. *Acta Math.* **206** (2011), no. 1, 127–204. MR 2784665 Zbl 1217.15043

[45] G. Teschl, *Mathematical methods in quantum mechanics.* With applications to Schrödinger operators. Graduate Studies in Mathematics, 99. American Mathematical Society, Providence, RI, 2009. MR 2499016 Zbl 1166.81004

[46] L. V. Tran, V. H. Vu, and K. Wang, Sparse random graphs: Eigenvalues and eigen-vectors. *Random Structures Algorithms* **42** (2013), no. 1, 110–134. MR 2999215 Zbl 1257.05089

[47] S. Warzel, Surprises in the phase diagram of the Anderson model on the Bethe lat-tice. In A. Jensen (ed.), XVIIth *International Congress on Mathematical Physics.* Proceedings of the congress (ICMP12) held in Aalborg, August 6–11, 2012. World Scientific, Hackensack, N.J., 2014, 239–253. MR 3204475 MR 3204717 (collection) Zbl 1305.82033 Zbl 1279.00048 (collection)

[48] F. Wegner, Bounds on the density of states in disordered systems. *Z. Phys. B* **44** (1981), no. 1-2, 9–15. MR 0639135

[49] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.* (2) **62** (1955), 548–564. MR 0077805 Zbl 0067.08403

Leander Geisinger, Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544, USA

e-mail: leander.geisinger@gmail.com