J Stat Phys (2013) 153:1022–1038
DOI 10.1007/s10955-013-0879-5

Approach to the Extended States Conjecture

Constanze Liaw

Received: 18 July 2013 / Accepted: 23 October 2013 / Published online: 7 November 2013

Abstract We develop a rather explicit approach concerning the extended states conjecture for the discrete random Schrödinger operator, or more generally for the so-called Anderson-type Hamiltonian. Our work is based on deep mathematical results by Jakšić–Last (Duke Math. J. 133(1):185–204, 2006).

Concretely, we suggest two new directions of research: We provide a formula which may lead the way to a rigorous proof of the conjecture, and an implementation of the proposed approach which yields numerical evidence in favor of the conjecture being true for the discrete random Schrödinger operator in dimension two. Not being based on scaling theory, this method eliminates problems due to boundary conditions, common to previous numerical methods in the field. At the same time, as with any numerical experiment, one cannot exclude finite-size effects with complete certainty.

We numerically track the "bulk distribution" (here: the distribution of where we most likely find an electron) of a wave packet initially located at the origin, after iterative application of the discrete random Schrödinger operator.

Keywords Anderson localization · Discrete random Schrödinger operator · Extended states conjecture

1 Introduction

While Anderson localization is a very well-developed field of physics (see e.g. [3, 7, 14, 15] and the references therein), many questions with striking physical relevance remain open. Among those is a rigorous treatment of the extended states conjecture in the critical dimension 2, which (if true) states that at weak disorder some extended states are preserved.

The author is partially supported by NSF-DMS-1101477, NSF-DMS-1261687.

C. Liaw (✉)
Department of Mathematics, Baylor University, One Bear Place #97328, Waco, TX 76798-7328, USA
e-mail: Constanze_Liaw@baylor.edu

Consider the discrete random Schrödinger operator in dimension d which is given by the self-adjoint operator

$$H_\omega = -\triangle + \sum_{k\in\mathbb{Z}^d} \omega_k \langle \cdot, \delta_k\rangle \delta_k \quad \text{on } l^2(\mathbb{Z}^d).$$

Here $\delta_k \in l^2(\mathbb{Z}^d)$ assumes the value 1 in the $k-$th entry, $k = (k_1, k_2, \ldots, k_d)$, and zero in all other entries. The random variables ω_k are i.i.d. with uniform distribution in $[-c/2, c/2]$, i.e. according to the probability distribution $\mathbb{P} = \prod_k c^{-1}\chi_{[-c/2,c/2]}dx$. (The Laplacian describes a crystal with atoms located at the integer lattice points $\mathbb{Z}^d$. The random part can be interpreted as having the atoms being not perfectly on the lattice points, but randomly displaced.)

This paper pertains to one of the "weaker" definitions of localization which is equivalent to the property that the spectrum of the operator is almost surely purely singular. In dimension $d = 1$ Anderson localization is produced by random disorders of any strength and at all energies (analytic proof, see e.g. [5, 6], or [8]). In dimension $d \geq 2$, localization is proved analytically for disorders c above a certain dimension-dependent threshold c_d. The first result of this type can be found in [9]. Simpler proofs and better constants can be found in [2] as well as [18]. Probably the best results are found in the preprint [16], where in particular a (not sharp) threshold of $c_2 = 22.8$ is proved in dimension 2. Hence, diffusion may only occur for small disorder c. This is precisely the problem of extended states: Does the two dimensional discrete random Schrödinger operator exhibit Anderson localization for small disorder with probability one?

The main contribution of this paper is the introduction of a very explicit approach of how to investigate the extended states conjecture, opening two new directions of research:

- First, we present a simple explicit formula (see Proposition 3.1 below), which is closely related to the extended states conjecture. Together with Corollary 3.2 below, this formula may lead the way to an analytic proof for delocalization at weak disorder for the two dimensional discrete random Schrödinger operator. Another important feature of these results is that they make a numerical experiment feasible: It suffices to track the orbit of just one vector.
- Second, as an application of this formula to the discrete random Schrödinger operator in two dimensions, we present numerical evidence in support of the extended states conjecture (Conjecture B.1 below) via the Numerical Criterion 4.2 below.

Since this conjecture has a long history of contradicting claims and is not in agreement with "conventional wisdom", we would like to briefly discuss its pros and cons. The main advantage of the method at hand is that we do not need to assume the scaling hypothesis which is at the heart of the part of numerical work using multi scale analysis to study the extended states conjecture (see e.g. [14] and the references within). Having said this, it is not immediately obvious whether or not our numerical experiment is subject to other problems that manifest themselves in rapidly increasing localization length. While it eliminates the difficulties of all numerical approaches due to unphysical contributions by the boundary conditions, the numerical part is still subject to finite-size restrictions. Hence, we interpret the numerical evidence in the form of a conjecture, rather than a conclusion. In this experiment, numerical errors due to round off tend to make extended states appear localized are therefore of lesser concern to us.

Yet, further investigations will be necessary and are encouraged. A refinement of the relationship between the proposed method and multi scale analysis or some combination

of the two would be interesting. And a comprehensive study of the re-scaling parameter in Appendices B.2 and B.3 below would be a natural next step to validate the method.

The numerical Sect. 4 through Appendix B are essentially self-contained and can be read almost independently. In Sect. 2, we introduce Anderson-type Hamiltonians, a very general notion of Anderson model including many of those studied in literature. We explain several statements concerning the structure of certain cyclic subspaces as they were obtained in Jakšić–Last [12]. In Sect. 3, we use the results by Jakšić and Last to derive the explicit formula, see Proposition 3.1 and Corollary 3.2. Section 4 is devoted to a description of the numerical experiment. In Sect. 5 we present an investigation of the distribution of energies after repeated application of the random operator of a wave packet initially located at the origin. In Sect. 6 we verify the performance of the method in many examples, e.g. for large disorder, for the free/unperturbed two dimensional discrete Schrödinger operator and for the one dimensional discrete random Schrödinger operator. We briefly remark on computing and memory requirements in Sect. 7. Appendix A contains some technical details concerning the numerical experiment. A summary of numerical results and the conclusions regarding the extended states conjecture, Conjecture B.1, can be found in Appendix B.

2 Preliminaries

2.1 Anderson-Type Hamiltonians, Discrete Schrödinger Operator

While the numerical experiment within pertains to the discrete random Schrödinger operator, we define so-called Anderson-type Hamiltonians. This is a more general class of perturbation problems possessing properties required by the proposed method. The advantage of making such a general definition is that this notion is a generalization of many Anderson models discussed in the literature. Anderson-type Hamiltonians were introduced in [11] and it is worth mentioning that their work explores the connections with so-called rank-one perturbations, although this fact is not the focus of our work.

For $k \in \mathcal{N}$, $\mathcal{N}$ no more than countable, consider the probability space $\Omega_k = (\mathbb{R}, \mathcal{B}, \nu_k)$, where $\mathcal{B}$ is the Borel sigma-algebra on $\mathbb{R}$ and ν_k is a Borel probability measure. In this paper we assume that ν_k is absolutely continuous. Let $\Omega = \prod_{k\in\mathcal{N}} \Omega_k$ be a product space with the probability measure $\mathbb{P}$ on Ω introduced as the product measure of the corresponding measures on Ω_k on the product sigma-algebra $\mathcal{A}$. The elements of Ω are points in $\mathbb{R}^\infty$, $\omega = (\omega_1, \omega_2, \ldots)$ for $\omega_k \in \Omega_k$.

Let $\mathcal{H}$ be a separable Hilbert space and let $\{\varphi_k\}_{k\in\mathcal{N}}$ be unit vectors in $\mathcal{H}$. For each $\omega \in \Omega$ define an Anderson-type Hamiltonian on $\mathcal{H}$ as a self-adjoint operator formally given by

$$H_\omega = H + V_\omega, \quad \text{where } V_\omega = \sum_k \omega_k \langle \cdot, \varphi_k \rangle \varphi_k. \tag{1}$$

Except for degenerate cases, the perturbation V_ω is almost surely a non-compact operator. Hence it is not possible to apply results from classical perturbation theory to study the spectra of H_ω, see e.g. [4] and [13].

In the case of an orthogonal sequence $\{\varphi_k\}$, such operators were studied in [11] and [12]. Probably the most important special case of an Anderson-type Hamiltonian is the discrete random Schrödinger operator on $l^2(\mathbb{Z}^d)$

$$Hf(x) = -\triangle f(x) = -\sum_{|k|=1} (f(x+k) - f(x)), \quad \text{with } \varphi_k(x) = \delta_k(x) = \begin{cases} 1 & x = k \in \mathbb{Z}^d, \\ 0 & \text{else.} \end{cases} \tag{2}$$

2.2 Singular and Absolutely Continuous Parts of Normal Operators

Recall that an operator in a separable Hilbert space is called normal if $T^*T = TT^*$. By the spectral theorem operator T is unitarily equivalent to M_z, multiplication by the independent variable z, in a direct sum of Hilbert spaces

$$\mathcal{H} = \oplus \int \mathcal{H}(z)\, d\mu(z)$$

where μ is a scalar positive measure on $\mathbb{C}$, called a scalar spectral measure of T.

If T is a unitary or self-adjoint operator, its spectral measure μ is supported on the unit circle or on the real line, respectively. Via Radon decomposition, μ can be decomposed into a singular and absolutely continuous parts $\mu = \mu_{\mathrm{s}} + \mu_{\mathrm{ac}}$. The singular component μ_{s} can be further split into singular continuous and pure point parts. For unitary or self-adjoint T, we denote by T_{ac} the restriction of T to its absolutely continuous part, i.e. T_{ac} is unitarily equivalent to $M_t\big|_{\oplus \int \mathcal{H}(t) d\mu_{\mathrm{ac}}(t)}$. Similarly, we denote the singular, singular continuous and the pure point parts of T, by T_{s}, T_{sc} and T_{pp}, respectively.

2.3 Equivalence Classes Generated by Cyclic Subspaces

In [12] obtained many interesting results about the structure of the cyclic subspaces of the Anderson-type Hamiltonian generated by the vectors φ_k and how they can be partitioned into equivalence classes. We explain the statements that are of immediate relevance to us.

By $\mathcal{H}_{\omega,k}$ we denote the cyclic subspace generated by φ_k under the operator H_ω, i.e.

$$\mathcal{H}_{\omega,k} = \operatorname{clos}\operatorname{span}\{H_\omega^l \varphi_k : l \in \mathbb{N} \cup \{0\}\}.$$

Let $\mu_{\omega,k}$ be the spectral measure of H_ω associated with the vector φ_k. Consider the set

$$\mathcal{S} = \{k : \mu_{\omega,k} \text{ is almost surely not equivalent to the Lebesgue measure}\}.$$

Theorem 2.1 (See [12]) *For the Anderson-type Hamiltonian given by* (1) *then we have*:

(i) *For all* k, m *the statement* $\mathcal{H}_{\omega,k} \not\perp \mathcal{H}_{\omega,m}$ *is deterministic, i.e.* $\mathbb{P}(\{\omega : \mathcal{H}_{\omega,k} \not\perp \mathcal{H}_{\omega,m}\})$ *is either* 0 *or* 1. *Further, the relation* $k \sim m$, *defined by the probability being* 1, *is an equivalence relation on* $\mathcal{S}$.

(ii) *The set of parameters* $\mathcal{S}$ *can be partitioned into equivalence classes* $\mathcal{S}_k$ *so that* $k \in \mathcal{S}_k$, $m \in \mathcal{S}_m$ *for* $k \neq m$ *implies* $\mathcal{H}_{\omega,k} \perp \mathcal{H}_{\omega,m}$, *and vice versa.*

As is explained in part (a) of Remark 3.3, in the case of the discrete random Schrödinger operator there is only one equivalence class. That is every k belongs to the same equivalence class. One can obtain an Anderson-type Hamiltonian with two equivalence classes, by taking the orthogonal sum of two discrete random Schrödinger operators. (The latter construction was mentioned to the author by J. Schenker.)

Theorem 2.2 (See [12]) *All vectors of the form* $\psi = \sum a_k \varphi_{k_l}$, $\varphi_{k_l} \in \mathcal{S}_{k_l}$, $a_k \neq 0$, $\sum |a_k|^2 < \infty$ *are cyclic for the singular part* $(H_\omega)_{\mathrm{s}}$ *almost surely.*

Remark 2.3 Further Jakšić and Last noticed that for Anderson-type Hamiltonians with a single equivalence class, the non-cyclicity of some φ_k with positive probability is sufficient to conclude the existence of extended states.

The latter theorem was built upon by Abakumov–Liaw–Poltoratski.

Theorem 2.4 (See [1]) *If some φ_k is cyclic for H_ω almost surely, then so is any non-zero vector.*

3 Distance Formula and Extended States Conjecture

Using Theorems 2.1, 2.2 and Remark 2.3, we deduce an explicit rigorous criterion implying the existence of extended states for Anderson-type Hamiltonians. Consider the Anderson-type Hamiltonian H_ω given by (1), and fix one of the vectors from $\{\varphi_k\}$. Without loss of generality take φ_0. Apply the Gram–Schmidt orthogonalization process (without normalization) to the sequence of vectors $\{\varphi_0, H_\omega\varphi_0, H_\omega^2\varphi_0, \ldots\}$. We denote the resulting sequence of vectors by $\{m_0, m_1, m_2, \ldots\}$. Let $\|\cdot\|_2$ denote the norm in the Hilbert space $\mathcal{H}$. For any unit vector $v \in \mathcal{H}$ consider the distance

$$D_{\omega,c}^n := \operatorname{dist}(v, \operatorname{span}\{H_\omega^l \varphi_0 : l = 0, 1, 2, \ldots, n\})$$

of v to the n dimensional cyclic subspace generated by φ_0 under H_ω.

The orthogonality of the set $\{m_0, m_1, m_2, \ldots\}$ implies that the distances $D_{\omega,c}^n$ can be computed effectively, we have:

Proposition 3.1 $D_{\omega,c}^n = \sqrt{1 - \sum_{k=0}^n \frac{\langle m_k, v\rangle^2}{\|m_k\|_2^2}}$.

The importance of this simple formula for the distance is justified by its impact when combined with Corollary 3.2 below. For the convenience of the reader, we have included the proof of Proposition 3.1 in Appendix A.4 below.

The sequence of distances $\{D_{\omega,c}^n\}_{n\in\mathbb{N}}$ is decreasing with values in $[0, 1]$, and therefore has a limit. If this limit is non-zero, then vector φ_0 is not cyclic for H_ω. In virtue of Remark 2.3, we obtain the following corollary to Theorem 2.1.

Corollary 3.2 *Let ω_k, be i.i.d. random variables with absolutely continuous distribution Ω_k. Assume that $\mathcal{S}$ consists of only one equivalency class. To prove the existence of extended states (i.e. the existence of absolutely continuous spectrum with positive probability), it suffices to find $c > 0$ for which the distance*

$$D_{\omega,c} := \lim_{n\to\infty} D_{\omega,c}^n \neq 0$$

with positive probability.

Remarks 3.3 (a) In the case of the discrete random Schrödinger operator there is only one equivalence class $\mathcal{S}_0$ generated by the $\varphi_0 = \delta_{00}$ vector. Indeed, we have $\mathcal{H}_{\omega,0} \not\perp \mathcal{H}_{\omega,k}$ for all $n \in \mathbb{Z}^2$, because $\mathcal{H}_{\omega,0} \not\perp \delta_k$ for all $n \in \mathbb{Z}^2$. For the numerical experiment below we chose $v = \delta_{11}$.

(b) The converse of Corollary 3.2 is not true. Hence we cannot draw any conclusions, if the distance between a fixed (unit) vector and the subspace generated by the orbit of another vector tends to zero. In particular, we cannot conclude that there must be localization. Even if we show $D_{\omega,c} = 0$ for many or 'all' vectors (instead of just one fixed v). Although it is not

expected, it is possible that the absolutely continuous part has multiplicity one. (c) Theorem 2.4 tells us that we are not restricted to choosing φ_0 to be a vector from $\{\varphi_k\}$. Instead we could take any non-zero vector.

(d) It would be interesting to investigate the following question: Does delocalization in dimension $d_0 \geq 2$ for some (weak) disorder c_0, imply delocalization in some dimension $d > d_0$ for some disorder c?

(e) As it was mentioned above, the corollary in conjunction with Proposition 3.1 provides a very explicit expression, which can potentially be used to rigorously prove the extended states conjecture.

4 Method of Numerical Experiment

Consider the discrete random Schrödinger operator in dimension $d = 2$ (for a definition see (2)). The first step is to **apply the Gram-Schmidt orthogonalization process to numerically compute distance between the unit vector δ_{11} and the subspace obtained by taking the closure of span**$\{\delta_{00}, H_\omega\delta_{00}, H_\omega^2\delta_{00}, \ldots, H_\omega^n\delta_{00}\}$. For fixed c, ω and n, denote this distance by $D^n_{\omega,c}$:

$$D^n_{\omega,c} := \operatorname{dist}(\delta_{11}, \operatorname{span}\{H^k_\omega\delta_{00} : k = 0, 1, 2, \ldots, n\}).$$

To the sequence of vectors $\{\delta_{00}, H_\omega\delta_{00}, H_\omega^2\delta_{00}, \ldots\}$, we apply the Gram–Schmidt orthogonalization process (without normalization). We denote the resulting sequence of vectors by $\{m_0, m_1, m_2, \ldots\}$. Let $\|\cdot\|_2$ denote the Euclidean norm.

Recall that by Proposition 3.1 we have an explicit expression for this distance

$$D^n_{\omega,c} = \sqrt{1 - \sum_{k=0}^{n} \frac{\langle m_k, \delta_{11}\rangle^2}{\|m_k\|_2^2}}\,, \tag{3}$$

and removing the abstraction from Theorem 3.2 we obtain:

Corollary 4.1 *Let ω_k, $k \in \mathbb{Z}^2$, be i.i.d. random variables with uniform (Lebesgue) distribution on $[-c/2, c/2]$, $c > 0$. To prove delocalization (i.e. the existence of absolutely continuous spectrum with positive probability), it suffices to find $c > 0$ for which the distance*

$$D_{\omega,c} := \lim_{n\to\infty} D^n_{\omega,c} \neq 0 \tag{4}$$

with positive probability.

Then we **initially fix c and a realization of the random variable ω**; computer-generated in accordance to the hypotheses of Corollary 4.1. In Appendix A.2 below, we describe the technicalities involved in computing $D^n_{\omega,c}$ for $n \in \{0, 1, 2, \ldots\}$.

Now, for the time being, assume that we have computed $D^n_{\omega,c}$ for $n = 0, \ldots, 4500$ and **find an estimate and a crude lower estimate for $D_{\omega,c}$ via a certain re-scaling argument**:

Figure 1 shows typical trends for the distance $D^n_{\omega,c}$ as a function of n. As the first $n = 200$ points do not contribute to the value of this limit and were generally rather irregular, we have omitted those points in what follows. The graph shows a decreasing function, as expected.

While the results at this point looked promising, they were not yet satisfactory. Most of all they do not provide a reliable estimate for the limit $D_{\omega,c}$. In order to obtain such an

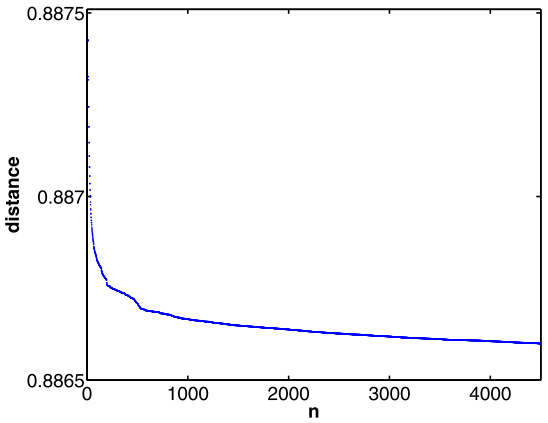

Fig. 1 Typical trend for the distance $D^n_{\omega,c}$ as a function of the number of iterations $n = 1, 2, 3, \ldots, 4500$ and $c = 0.1$. For larger values of c, this graph looks very similar. Notice the fine vertical scale. In order to exclude that the distance tends to zero e.g. logarithmically, we perform a re-scaling of the *horizontal axis* as shown in Fig. 2 below

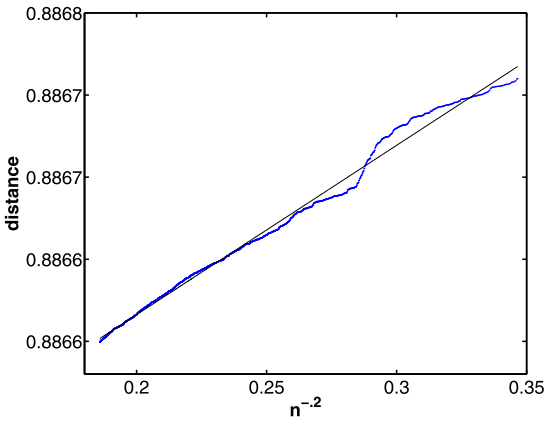

Fig. 2 Re-scaling of the *horizontal axis* in Fig. 1 (for $n = 200, 201, 202, \ldots, 4500$) using the best exponent $a = 0.2$. The y-intercept of the approximating line is the estimate $y_{\omega,c}$ of the value for $D_{\omega,c}$. Again, notice the fine vertical scale. In fact, we have $D^{200}_{\omega,c} \approx 0.88677$ and $D^{4500}_{\omega,c} \approx 0.88660$

estimate for $D_{\omega,c}$, we re-scaled the horizontal axis in Fig. 1 by a negative power n^{-a} (power of the reciprocal, so that the horizontal axis is reverted) and very carefully approximated the resulting graph by a line. This rescaling technique is a well established tool to analyze numerical data and strengthens our conclusions.

The re-scaled graph is shown in Fig. 2. Appendix A.3 contains information about the choice of the re-scaling factor a and explains why, for appropriately small disorders, the graph does not decay to zero, e.g. logarithmically.

The value of $D_{\omega,c}$ is estimated by the y-intercept, $y_{\omega,c}$, of the approximating line.

Since the re-scaled graphs in Fig. 2 were sometimes rather noisy (e.g. a line through the steeper sections of the graph has a lower y-intercept), we include a lower estimate $L_{\omega,c}$ for $y_{\omega,c}$ given by the minimum y-intercept of the lines passing through any two consecutive points.

Summarizing the last few steps, we have

$$D_{\omega,c} \approx y_{\omega,c} \geq L_{\omega,c}\,. \tag{5}$$

Finally, **we repeat the experiment for many values of c and many computer-generated realizations for the random variable ω**: Concerning the different realizations, throughout we took the minimum of $y_{\omega,c}$ and $L_{\omega,c}$ over all the different computer-generated

realizations of ω. Roughly the goal is to show that for some $c > 0$, the limits $D_{\omega,c}$ are bounded away from zero for many realizations ω.

As an alternative way of analyzing the data, for fixed c we take the averages of $D^n_{\omega,c}$, $n = 1, 2, 3, \ldots, 4500$ over the four realizations and apply our re-scaling method to those averages. Let $\tilde{y}_c$ and $\tilde{L}_c$ denote the y-intercepts we obtained from this method.

The numerical data is evaluated against the following criterion:

Numerical Criterion 4.2 *For a fixed value of c, we say that we have delocalization, if the following three properties hold at the same time*:

(1) *The minimum y-intercept of the lines passing through any two consecutive points* $L_{\omega,c} > 0.5$.
(2) *For at least* 75 *% realizations we obtain a re-scaling parameter* $a \in [0.04, 0.85]$. (*Notice that we only need non-zero probability by Corollary* 4.1, *and part* (b) *of Remark* 3.3.)
(3) *The minimum y-intercept of the lines passing through any two consecutive points for the averaged re-scaled distances* $\tilde{L}_c > 0.5$, *and the corresponding re-scaling parameter re-scaling parameter* $\tilde{a} \in [0.04, 0.85]$.

The numerical data obtained and its conclusion is presented in Appendix B.

5 Evolution Under H_ω of the Bulk for Small Values of c

We study the bulk distribution in terms of the distance from the origin of the vector δ_{00} after iterative application of the random Hamiltonian, which describes how a wave packet initially located at the origin progresses. To remind the reader, the "bulk distribution" helps to locate the distance from the origin where we are most likely to find an electron. To be precise, we define

$$E(l, n) = \sqrt{\sum_{|i|+|j|=l} (m_n)^2_{i,j}} \tag{6}$$

for the bulk $E(l, n)$ of the vector m_n at taxicab distance l from the origin. Here $(m_n)_{i,j}$ refers to the (i, j)-entry of the 2-tensor m_n (m_n was defined just before Proposition 3.1). Slightly abusing notation, we normalize m_n and use the same notation for the normalized sequence of vectors.

For example, in order to obtain that the bulk distribution of the vector m_n (defined in Appendix A.2) is at taxicab distance 2 from the origin, we use the elements of m_n which are located on the diamond for which the matrix

$$\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \iddots \\ \ldots & 4 & 3 & 2 & 3 & 4 & \ldots \\ \ldots & 3 & 2 & 1 & 2 & 3 & \ldots \\ \ldots & 2 & 1 & 0 & 1 & 2 & \ldots \\ \ldots & 3 & 2 & 1 & 2 & 3 & \ldots \\ \ldots & 4 & 3 & 2 & 3 & 4 & \ldots \\ \iddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

has entries equal to 2. The bulk $E(2, n)$ of the vector m_n at 'distance' 2 from the origin is equal to the Euclidean norm over the elements in this diamond.

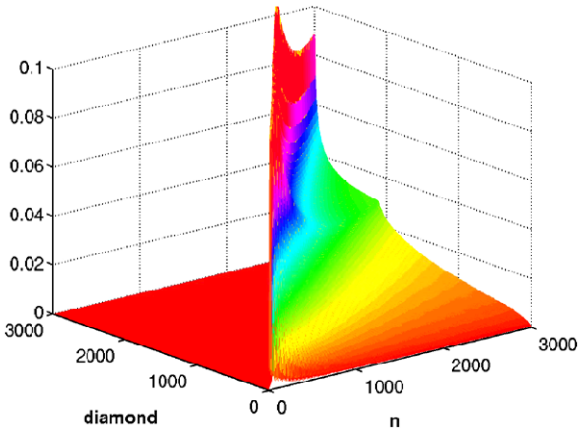

Fig. 3 For $c = 0.05$ we show $E(l, n)$, i.e. the evolution of the bulk distribution of $H^n_{\omega,c}\delta_{00}$ for the diamonds at distance l from the origin. Notice that the bulk travels far out from the origin (the diagonal is the farthest possible)

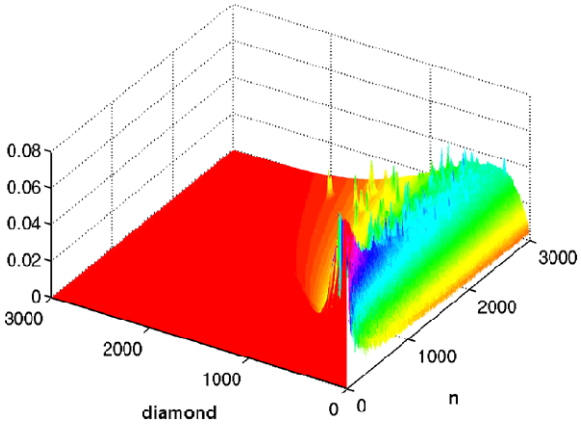

Fig. 4 The analog of Fig. 3 for $c = 0.6$ using different color scale. The bulk remains closer to the origin. Conjecture B.1 predicts extended states

Via small modifications of our programs, we extract the bulk distribution of the orbit of the vector δ_{00} under the random Hamiltonian for the values $c = .1$, $c = 1$ and $c = 5$ of disorder, see Fig. 5. In accordance with Conjecture B.1 below, the bulk for small disorder is far away from the origin whereas it is concentrated close to the origin for large disorder.

Figures 3 and 4 shows the bulk distribution of $H^n_\omega\delta_{00}$ for $n = 2999$ for values of c ranging from $c = .1$ to $c = 1$. Again, the fact that the bulk for small disorder is far away from the origin whereas it shifts much closer to the origin as the disorder increases, supports Conjecture B.1 below. Both figures are the averages obtained from two realizations for each value of c. And again, many repetitions of these experiments for smaller values of n were carried out, and the figures shown represent the behavior obtained in all repetitions.

6 Further Verifications of the Method and the Numerical Experiments

Apart from the usual tests (the program is running stably, checking all subroutines, many verifications for small n), we have also validated the code and versions for other models: the free/unperturbed two dimensional Schrödinger operator and the one dimensional random Schrödinger operator.

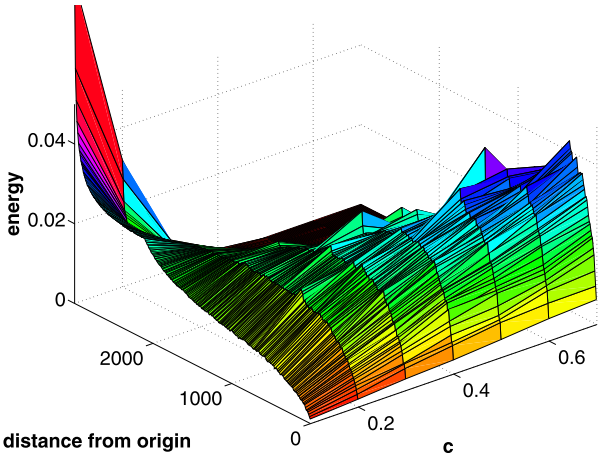

Fig. 5 The *figure* shows $E(l, 2999)$ (the Euclidean norm on the diamonds of m_{2999} given by (6)) as a function of the 'distance' l from the origin for values of disorder $c = 0.1 : 0.1 : 0.7$

6.1 Free Discrete Two Dimensional Schrödinger Operator

When we apply the free discrete Schrödinger operator $H = H_{\mathbf{0}}$ to the vector δ_{00}, it immediately becomes clear that $H\delta_{00}$, as well as all vectors $H^n\delta_{00}$, $n \in \mathbb{N} \cup \{0\}$, are symmetric with respect to the origin. In dimension $d = 2$, it is not hard to see that the distance between δ_{11} and the orbit of δ_{00} under H is at least $\sqrt{3}/2 \approx 0.8660$. Indeed, we have

$$\operatorname{dist}(\delta_{11}, \operatorname{clos}\operatorname{span}\{H^n\delta_{00} : n \in \mathbb{N} \cup \{0\}\}) > \min_x \operatorname{dist}(u_x, \delta_{11}) = \sqrt{3}/2,$$

where

$$u_x = x\delta_{-1-1} + x\delta_{-11} + x\delta_{1-1} + x\delta_{11} = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \iddots \\ \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & x & 0 & x & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & x & 0 & x & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ \iddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

In the experiments for the free discrete two dimensional Schrödinger operator, we obtained a y-intercept of the approximating line approximately equals to 0.8867. The re-scaled graph of distances still had a very convex shape, so the actual distance as $n \to \infty$ would be bigger. In fact, we have extracted from Fig. 6 an upper estimate of 0.8868 by zooming in. Therefore, the distance must lie in the interval [0.8866, 0.8868].

6.2 Verifying Localization for the One Dimensional Random Schrödinger Operator

Consider the discrete random Schrödinger operator in one dimension, see e.g. (1) and (2) with $d = 1$. For this operator, it is well known that localization occurs for random disorders of all strengths (in particular, for small values of c) and at all energies. We have adopted and applied this approach to study the discrete random Schrödinger operator in one dimension. Figure 7 shows a typical re-scaled graph of the distance $D^n_{\omega,c}$ for $n =$

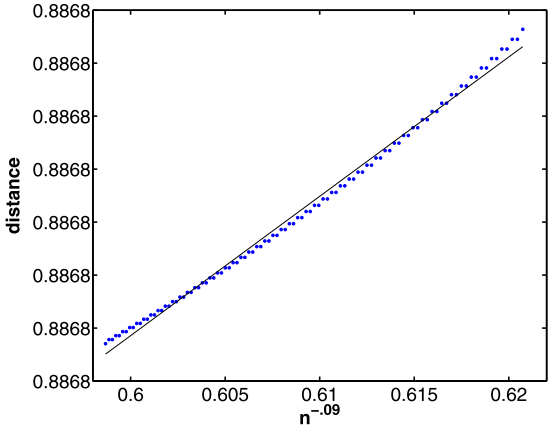

Fig. 6 Convex shape for the distance from δ_{11} to the orbit of δ_{00} under the free Schrödinger operator in two dimensions. The approximating line has y-intercept $y_{\omega,c} \approx 0.8866$. The vertical scale is of order 10^{-5}

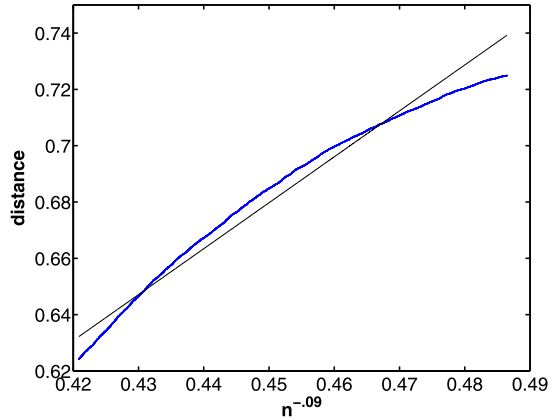

Fig. 7 Discrete random Schrödinger operator in one dimension with disorder $c = 0.05$. For $a = 0.09$, we have $y_{\omega,c} \approx -0.0543$ and a concave graph

$3000, 3001, 3002, \ldots, 15000$ for the disorder $c = 0.05$. With a re-scaling exponent of $a = 0.09$, the graph of $D^n_{\omega,c}$ is still concave. We have $y_{\omega,c} = -0.0543$. Of course, the actual distance will not drop below zero but rather tend to zero. The fact that the numerically computed value for $D^n_{\omega,c}$ tends to zero faster than the exact distance is a manifestation of finite precision. In all our experiments finite precision will always yield a smaller distance than the exact one. This suggests that we will not obtain false "positives"/extended states.

Recall by part (b) of Remark 3.3 that this experiment does not allow us to conclude that there is localization, the result still provides support for the credibility of the method at hand as well as the numerical design.

7 On Computing and Memory Requirements

The numerical implementation of the proposed method uses memory rather efficiently, so that the numerical experiments were mainly limited by the length of the computation. On the machines available to us (we used 2 cores each with 4 GB ram on an 8 core machine), it took 8 1/2 hours to complete one realization for one value of c. Since we need to include many realizations of the random variable and many values of c, it took even several units

a considerable time to finish all the computations. In order to compute D^n_ω described in Sect. 4, our code requires order n^2 (i.e. $\mathcal{O}(n^2)$) memory. Indeed, in order to carry out the Gram–Schmidt orthogonalization process described in Appendix A.2, we must store matrices of size $\mathcal{O}(n) \times \mathcal{O}(n)$. The corresponding code for the d-dimensional discrete random Schrödinger operator will require memory size of order $\mathcal{O}(n^d)$. The random Schrödinger operator on the (dyadic) tree uses memory of $\mathcal{O}(2^n)$. With the resources available to us, memory restrictions would only allow us to compute up to $n \approx 27$ for the tree. In this case, we cannot produce sufficient data to support the fact that the discrete random Schrödinger operator on the tree does indeed exhibit delocalization.

Unfortunately, the method is not well suited for parallelization, as communication after each step n be required. However, it is, of course, possible to run one realization on each core.

Acknowledgements The author would like to thank A. Poltoratski for suggesting the initial mathematical idea of the experiment, which we later noticed was already published in [12]. Further, she is in great debt to G. Berkolaiko for the insightful discussions concerning many aspects of this research as well as for reading and making useful comments on many parts of this paper. The author would like to thank W. Bangerth and G. Kanschat for allowing me to use their computer cluster. Particular thanks also to J. Kuehl for running initial experiments using a code of his, for his comments and edits, and for encouraging her to pursue this idea. Finally, the author would like to thank the reviewers for their valuable comments, especially regarding the re-scaling parameter.

Appendix A: Some Technicalities

A.1 When to Fix the Realization ω

In the experiments described here, we had fixed c and ω at the beginning. For fixed $c = 0.1$, we also computed several cases for which we chose a different realization ω each time the random operator was applied. In other words, for $c = 0.1$ we fix countably many realizations of ω each independently distributed and each in accordance with Corollary 4.1. Let those realizations be denoted by ω, $\tilde{\omega}$, $\hat{\omega}$, etc. Then compute the distance between δ_{11} and the closure of the span of the vectors δ_{00}, $H_\omega\delta_{00}$, $H_{\tilde{\omega}}(H_\omega\delta_{00})$, $H_{\hat{\omega}}(H_{\tilde{\omega}}H_\omega\delta_{00})$, etc. The results obtained from this setup agreed very well with the ones described in Appendix B below.

A.2 Computing the Distance $D^n_{\omega,c}$

For fixed n and ω let us briefly explain the computational approach to obtain $D^n_{\omega,c}$. The main idea is to apply the Gram–Schmidt orthonormalization process in order to recursively compute $D^n_{\omega,c}$. (Unlike in the remainder of the paper, from now on we carry out the normalization in the Gram–Schmidt process.) Recall that the orthogonal sequence $\{m_0, m_1, m_2, \ldots, m_n\}$ was obtained by applying the Gram–Schmidt algorithm to $\{\delta_{00}, H^1_\omega\delta_{00}, H^2_\omega\delta_{00}, \ldots, H^n_\omega\delta_{00}\}$. The following proposition says that in the Gram–Schmidt algorithm all but the two last terms in the sum are zero. While this fact seems to be well known to the physics and the orthogonal polynomials (as the three term recurrence formula, see e.g. [17]) community, the author is grateful to M. Hastings from whom she learned it and its simple proof. This simplification reduces the required memory by the order n (from $\mathcal{O}(n^3)$ to $\mathcal{O}(n^2)$).

Proposition A.1 *The vector $H_\omega m_n$ is orthogonal to m_l for all $l = 0, 1, 2, \ldots, n-2$.*

Proof We use mathematical induction in n. Consider $n = 2$. Assume that we have computed the orthonormal vectors m_0, m_1 and m_2 (via the Gram–Schmidt orthogonalization process). Since the operator is self-adjoint, we have $\langle H_\omega m_2, m_0\rangle = \langle m_2, H_\omega m_0\rangle$. Since $m_1 = H_\omega m_0 - \langle H_\omega m_0, m_0\rangle m_0$ and because m_2 is orthogonal to m_0 and m_2, we obtain

$$\langle H_\omega m_2, m_0\rangle = \langle m_2, m_1 + \langle H_\omega m_0, m_0\rangle m_0\rangle = 0.$$

Assume that the statement of the proposition is true for some $n - 1 \geq 2$. It remains to show that the statement is true for n. Assume that we have computed an orthonormal sequence $m_0, m_1, \ldots, m_n$. For $l \leq n-2$, it suffices to show that $\langle H_\omega m_n, m_l\rangle = 0$. By following the argument for the base case, we obtain

$$\langle H_\omega m_n, m_l\rangle = \langle m_n, H_\omega m_l\rangle = \langle m_n, m_{l+1} + \langle H_\omega m_l, m_l\rangle m_l\rangle.$$

The latter expression equals zero, because we assumed $l \leq n - 2$ and the orthogonality assumption on $m_0, m_1, \ldots, m_n$. □

For the implementation, we normalized the vectors in the sequence $\{m_l\}$. With (3) we obtain the iterative formula

$$(D^n_{\omega,c})^2 = (D^{n-1}_{\omega,c})^2 - ((m_n)_{1,1})^2,$$

where $(m_n)_{1,1}$ equals the $(1, 1)$-entry of m_n.

A.3 Choice of the Re-scaling Parameter

For each fixed c and ω, the re-scaling exponent a is chosen so that the re-scaled graph of the distance function (see Fig. 2) satisfies the least square property; that is, the error when approximating the graph by a line is minimal. With this exponent, we then find the corresponding linear approximation for the re-scaled distance function. The existence of a positive re-scaling factor implies that the graph in Fig. 1 will not decay to zero, e.g. logarithmically. Indeed, if we use a re-scaling factor smaller than the one in the table will result in a 'globally concave' graph for the distances $D^n_{\omega,c}$. In this case, the y-intercept of the line lies below the value expected for $D^\infty_{\omega,c}$.

A.4 Proof of Proposition 3.1

Define the vectors

$$v_N = v - \sum_{k=0}^{N} \frac{\langle m_k, v\rangle}{\|m_k\|_2^2} m_k$$

and observe that $v_N \perp \text{span}\{m_k : k = 0, \ldots, N\}$.

By the Pythagorean Theorem we have

$$1 = \|v\|^2 = \left\| v_N + \sum_{k=0}^{N} \frac{\langle m_k, v\rangle}{\|m_k\|_2^2} m_k \right\|^2 = \|v_N\|^2 + \left\| \sum_{k=0}^{N} \frac{\langle m_k, v\rangle}{\|m_k\|_2^2} m_k \right\|^2$$

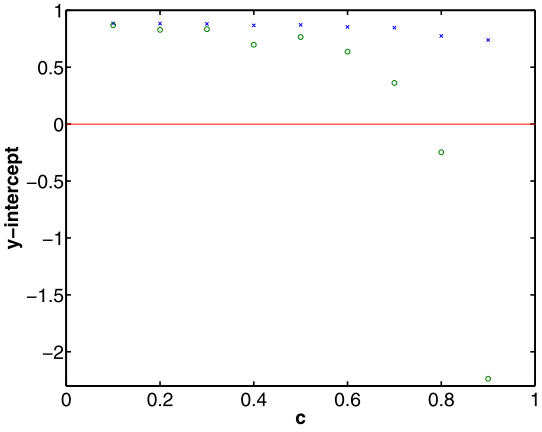

Fig. 8 As a function of c we show the minimum values of $y_{\omega,c}$ (x's; larger values) and $L_{\omega,c}$ (*circles*; smaller values) obtained from four realizations for each c. Notice that $D_{\omega,c} \approx y_{\omega,c} \geq L_{\omega,c} > 0.5$ for $c \leq 0.6$

and solving for $\|v_N\|$ gives

$$\|v_N\| = \sqrt{1 - \left\| \sum_{k=0}^{N} \frac{\langle m_k, v \rangle}{\|m_k\|_2^2} m_k \right\|^2} .$$

Since sequence $\{m_0, m_1, m_2, \ldots\}$ is orthogonal, the distance $D_{\omega,N} = \|v_N\|$ and

$$\left\| \sum_{k=0}^{N} \frac{\langle m_k, v \rangle}{\|m_k\|_2^2} m_k \right\|^2 = \sum_{k=0}^{N} \left\| \frac{\langle m_k, v \rangle}{\|m_k\|_2^2} m_k \right\|^2 = \sum_{k=0}^{N} \frac{\langle m_k, v \rangle^2}{\|m_k\|_2^2} ,$$

which completes the proof. □

Appendix B: Findings and the Extended States Conjecture

Presenting our findings for all three parts of the Numerical Criterion 4.2, we summarize the numerical observations supporting the extended states conjecture, Conjecture B.1 below.

B.1 Observed y-Intercepts

As mentioned in Sect. 4, for a fixed c, we chose many realizations ω. We took the minimum of the quantities for $L_{\omega,c}$ and $y_{\omega,c}$ (the y-intercept of the approximating line and the minimum y-intercept of the lines passing through any two consecutive points, respectively). Figure 8 shows $L_{\omega,c}$ and $y_{\omega,c}$ as a function of c. Being rather cautious, we say that if $y_{\omega,c} \leq 0.5$ then the orbit of δ_{00} may not span the whole space, and we cannot conclude delocalization.

B.2 Observed Re-scaling Factors

We include an extract of the table of best re-scaling exponents, a, which satisfy the least square property for our data. Each entry in the table corresponds to a different realization of the random variable ω for the indicated value of c. As the values for $y_{\omega,c}$ were not sensitively

dependent on the precise value of a, we used a rather coarse mesh $a = 0.05 : 0.05 : 0.85$ and refined using $a = 0.01 : 0.01 : 0.05$, if the best re-scaling exponent was below 0.05.

The N/A indicates that for this particular realization, even the re-scaling parameter $a = 0.02$ yields a concave graph. For such a realization, we do not obtain any information. For values of $c \gtrsim 1.2$ many realizations did not yield a reasonable best fit parameter a. In virtue of part (b) of Remark 3.3, no conclusion can be made for such disorders.

The latter table suggests the following valid question: Why does the re-scaling factor change so much from one realization to another one? Further, with some intuition and heuristic reasoning using ergodic theory one would expect that for a fixed value of c the re-scaling parameter should obey some statistical distribution.

With respect to the first question, i.e. on the probabilistic character of the re-scaling parameter, notice that in order to be able to draw conclusions for a particular realization, it is merely important to find a reasonable re-scaling factor for that realization. Also, notice that for $c \leq 0.6$ there are a total of 27 out of 28 realizations with acceptable re-scaling parameter, and by Corollary 3.2, we only need to show that the distance is bounded away from zero with positive (non-zero) probability. For the one realization that does not have a re-scaling parameter (where the table shows "N/A"), we refer the reader to part (b) of Remark 3.3. It is also useful to observe that the re-scaling parameter has no physical interpretation (unfortunately). In fact, not even the computed distances $D^n_{\omega,c}$ have a physical meaning. Further, by choosing different intervals for n (e.g. $500 \leq n \leq 1000$ or $3500 \leq n \leq 4000$) we found no change in the re-scaling parameter for larger n. Last but not least and very importantly, for fixed c the re-scaling factor depends in a complicated way on the particular "path" of the electron and hence on the particular realization. For example, the direction in which the electron initially "moves" matters: The values of the random variable ω near the origin influence the observed distances $D^n_{\omega,c}$ for small n, which will significantly impact the magnitude of a, but not the sign. Indeed, an initial fast decrease of the distances tends to lead to a larger rescaling factor as we disregard the distances for small n and the curvature of the remaining graph is expected to be smaller. A detailed parameter study of this phenomenon would be interesting, yet require significant computational resources which the author is pursuing.

Considering the second question, we now discuss the behavior of the averaged distances.

B.3 Observed Re-scaling Factors and y-Intercepts for the Averaged Distances

The re-scaling factors, $\tilde{a}$, for the averaged distances over the four realizations, for fixed c, are shown in row two of Table 2 below. Rows three and four, respectively, show the y-intercepts $\tilde{y}_c$ and $\tilde{L}_c$ for the averaged distances (for the re-scaling parameters in row two).

For the disorder strengths $c = 1.2$ and $c = 1.3$, the averaged distances did not yield a re-scaling parameter $\tilde{a} \in [.02, .85]$, but rather $a < .02$. While the values shown in the latter table are limited to four realizations per disorder certainly, the overall trend observed in Tables 1 and 2 shows that the value of a tends to decrease as c increases. Although only further testing can reveal the truth, it is expected that computing many more trials for each disorder would yield a re-scaling factor that decreases smoothly as the disorder increases.

B.4 Conclusion

Combining Fig. 8 with Tables 1 and 2, we observe that for $c \leq 0.6$, all three numerical criteria for delocalization, Numerical Criterion 4.2, are satisfied. Hence the final conclusion of this numerical experiment is the following conjecture.

Table 1 Re-scaling factors a corresponding to four realizations for each disorder strength c

c	.1	.15	.2	.3	.4	.5	.6	.7	.8	.9	1	1.2	1.3
	.2	.04	.2	.25	.1	.15	.1	.1	.05	.02	.05	.1	N/A
	.1	.2	.25	.1	.04	.1	.1	.05	.05	.02	N/A	N/A	.05
	.2	N/A	.2	.2	.15	.1	.2	.1	.1	.04	.03	N/A	N/A
	.05	.1	.05	.15	.15	.1	.05	.1	.05	.03	.05	N/A	N/A

Table 2 Observed re-scaling factors and y-intercepts for the averaged distances

c	.1	.15	.2	.3	.4	.5	.6	.7	.8	.9	1
$\tilde{a}$	.1	.1	.15	.2	.1	.1	.1	.1	.05	.1	.04
$\tilde{y}_c$	.8860	.8851	.8848	.8832	.8763	.8729	.8684	.8611	.8345	.8433	.7925
$\tilde{L}_c$	.8837	.8707	.8799	.8768	.8563	.8343	.8131	.7902	.6656	.6321	.4925

Conjecture B.1 (Delocalization/extended states conjecture) *For weak disorder $c \leq 0.6$, the two dimensional discrete random Schrödinger operator does not exhibit Anderson localization almost surely*; *in the sense that it has non-zero absolutely continuous spectrum almost surely.*

Remark B.2 The presence of absolutely continuous spectrum implies that we do not have what is usually referred to as "strong dynamical localization" (see e.g. [10]).

Also notice that for $c \leq 0.6$ a convincing total number of 27 out of 28 realizations indicated delocalization, and that it sufficed to show that the probability is non-zero (by Corollary 4.1). Further, the author has tested many more realizations for different values of n, all of which yielded comparable results.

It is worth repeating (see part (b) of Remark 3.3) that we cannot conclude localization even if $y_{\omega,c} < 0$, or if only a very small percentage of realizations are usable (i.e. yield re-scaling factors). Therefore, the experiments are inconclusive, and, in particular, do not imply localization for larger values of c.

We do not gain any information about the regimes of physical energy at which delocalization occurs. And the proposed method does not reveal information about the behavior close to the mobility edge. In particular, we cannot draw any conclusions on how precisely the transition from extended to localized states occurs.

Experiments for other geometries are currently conducted.

References

1. Abakumov, E., Liaw, C., Poltoratskiĭ, A.: Cyclic vectors for rank-one perturbations and Anderson-type Hamiltonians. J. Lond. Math. Soc. (2013). doi:10.1112/jlms/jdt028
2. Aizenman, M., Molchanov, S.: Localization at large disorder and at extreme energies: an elementary derivation. Commun. Math. Phys. **157**(2), 245–278 (1993)
3. Anderson, P.W.: Absence of diffusion in certain random lattices. Phys. Rev. **109**, 1492–1505 (1958)
4. Birman, M.S., Solomjak, M.Z.: Spectral Theory of Self-Adjoint Operators in Hilbert Space. Springer, Berlin (1987). Originally published in Russian 1986
5. Carmona, R., Lacroix, J.: Spectral Theory of Random Schrödinger Operators. Birkhäuser, Basel (1990)

6. Cycon, H., Froese, R., Kirsh, W., Simon, B.: Topics in the Theory of Schrödinger Operators. Springer, Berlin (1987)
7. Edwards, J.T., Thouless, D.J.: Numerical studies of localization in disordered systems. J. Phys. C: Solid State Phys. **5** (1972)
8. Figotin, A., Pastur, L.: Spectral Properties of Disordered Systems in the One-Body Approximation. Springer, Berlin (1991)
9. Fröhlich, J., Spencer, T.: Absence of Diffusion in the tight binding model for large disorder of low energy. Commun. Math. Phys. **88**, 151–184 (1983)
10. Hundertmark, D.: A Short Introduction to Anderson Localization, Analysis and Stochastics of Growth Processes and Interface Models pp. 194–218. Oxford Univ. Press, Oxford (2008). MR 2603225 (2011c:82038)
11. Jakšić, V., Last, Y.: Spectral structure of Anderson type Hamiltonians. Invent. Math. **141**(3), 561–577 (2000). MR 1779620 (2001g:47069)
12. Jakšić, V., Last, Y.: Simplicity of singular spectrum in Anderson-type Hamiltonians. Duke Math. J. **133**(1), 185–204 (2006). MR 2219273 (2007g:47062)
13. Kato, T.: Perturbation Theory for Linear Operators. Classics in Mathematics. Springer, Berlin (1995). Reprint of the 1980 edition, MR 1335452 (96a:47025)
14. Lagendijk, A., van Tigglen, B., Wiersma, D.S.: Fifty years of Anderson localization. Phys. Today **82**(8), 24–29 (2009). Featured article
15. Kramer, B., MacKinnon, A.: Localization: theory and experiment. Rep. Prog. Phys. **56**, 1469–1564 (1993). 1–119
16. Schenker, J.: How large is large? Estimating the critical disorder for the Anderson model. Preprint arXiv:1305.6987v1
17. Lieb, H.: Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory. AMS Colloquium Publications. American Mathematical Society, Providence (2005)
18. Lieb, H.: Spectral analysis of rank one perturbations and applications, Mathematical quantum theory. II. Schrödinger operators. In: Trace Ideals and Their Applications, Vancouver, BC, 1993. CRM Proc. Lecture Notes, vol. 8, pp. 109–149. Amer. Math. Soc., Providence (1995). Also available as part of, Simon, B., Mathematical Surveys and Monographs. vol. 120. American Mathematical Society, Providence (2005), viii+150 pp.