Delocalization for the 3D discrete random Schrödinger operator at weak disorder

2014 J. Phys. A: Math. Theor. 47 305202

(http://iopscience.iop.org/1751-8121/47/30/305202)

IOP Publishing Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **47** (2014) 305202 (14pp) doi:10.1088/1751-8113/47/30/305202

Delocalization for the 3D discrete random Schrödinger operator at weak disorder

Constanze Liaw, Westin King and Robert C Kirby

Department of Mathematics, Baylor University, One Bear Place #97328, Waco, TX 76798-7328, USA

E-mail: Constanze-Liaw@baylor.edu, Westin-King@alumni.baylor.edu and Robert-Kirby@baylor.edu

Received 23 January 2014, revised 22 May 2014
Accepted for publication 6 June 2014
Published 10 July 2014

Abstract

We apply a recently developed approach (Liaw 2013 *J. Stat. Phys.* **153** 1022–38) to study the existence of extended states for the three-dimensional discrete random Schrödinger operator at small disorder. The conclusion of delocalization at small disorder agrees with other numerical and experimental observations (see e.g. (Lagendijk *et al* 2009 *Phys. Today* **82** 24–29)). Further the work furnishes a verification of the numerical approach and its implementation. 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. Our work can be thought of as a new and quite different use of Lanczos' algorithm; *a posteriori* tests show that the orthogonality loss is very small. 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: 3D Anderson model, extended states, Lanczos algorithm
PACS numbers: 05.45.Mt, 05.60.−k, 71.23.An, 72.15.Rn
Mathematics Subject Classification: 47A16, 47B80, 81Q10

(Some figures may appear in colour only in the online journal)

1. Introduction

Consider the discrete three-dimensional (3D) Schrödinger operator, given by:

$$-\triangle f(x) = -\sum_{|i|=1}(f(x+i) - f(x)),$$

when i is of the form $(i_1, i_2, i_3)^T$, $i_k \in \mathbb{Z}$, and consider an element $\delta_i(x)$ of $l^2(\mathbb{Z}^3)$ given by

$$\delta_i(x) = \begin{cases} 1 & x = i \in \mathbb{Z}^3, \\ 0 & \text{else.} \end{cases}$$

Let the random variables ω_i be i.i.d. with uniform distribution in $[-c/2, c/2]$, i.e. according to the probability distribution $\mathbb{P} = c^{-1} \Pi_i \chi_{[-c/2,c/2]} dx$.

The 3D random discrete Schrödinger operator, formally given by

$$H_\omega = -\triangle + \sum_{i \in \mathbb{Z}^3} \omega_i < \cdot, \delta_i > \delta_i \quad \text{on} \quad l^2(\mathbb{Z}^3), \tag{1}$$

is the main object of study.

This operator has been studied extensively, see e.g. [10, 19] and the references therein. The first part of the operator $-\triangle$ describes the movement of an electron inside a crystal with atoms located at all integer lattice points $\mathbb{Z}^3$. The perturbation $\sum_{i \in \mathbb{Z}^3} \omega_i < \cdot, \delta_i > \delta_i$ can be interpreted as having the atoms randomly displaced around the lattice points. It is important to notice that the perturbation is almost surely non-compact, so that classical perturbation theory (e.g. Weyl–von Neuman theorem, which states the invariance of the essential spectrum under compact perturbations) cannot be applied almost surely. It is known that the absolutely continuous spectrum is deterministic, i.e. it occurs with probability one or zero, see e.g. [15]. Localization in the sense of exponentially decaying eigenfunctions was proved analytically for disorders c *above* some threshold C_0 (see e.g. [3, 6, 19]). Currently, the smallest threshold in three dimensions is $C_0 = 100.6$ (see table 1 in [18]).

Diffusion is expected but not proved for small disorder $c > 0$. We numerically determine a regime of disorders for which the 3D discrete random Schrödinger operator does not exhibit localization. Our calculations are based on the Lanczos algorithm [12] for determining orthogonal bases for Krylov spaces [21]. Although we are not the first to use the Lanczos algorithm (see e.g. [11, 20] and the references therein), our application of it is quite different. In particular, our method is not based on scaling theory (for further discussion see [14]). In [17], the Lanczos algorithm is employed to compute a set of eigenvalues and eigenvectors. A few so-called interior eigenvectors and eigenvalues are calculated effectively by pre-conditioning methods. However, we test for localization without computing eigenvalues or eigenvectors, but only compute the distance between δ_{111} and the orbit of δ_{000}. The orbit is the span of $\{H_\omega^k \delta_{000} : k \in \mathbb{N} \cup \{0\}\}$, which is exactly a Krylov subspace. At each step of the Lanczos iteration, we use the orthogonality of the generated vectors to update the distance of interest. In this way, we maintain the low memory cost of a three-term recurrence, bypassing the need to store any eigenvectors at all. In addition to this, we have performed some *a posteriori* tests of the Lanczos algorithm on smaller cases to measure the degree to which orthogonality may be lost.

Besides computational advantages, our approach also offers a different mathematical perspective. By utilizing eigenvectors, it is (tacitly) assumed that all spectral points are in fact eigenvalues, while our approach merely generates an orbit without attempting to rule out other kinds of spectral points.

While the contributions of this paper are numeric, the method (see [14]) provides an explicit analytic expression, which may yield a proof of the following numerically supported main result.

Main result 1. *For disorder $c \lesssim 2.0$, numerical experiments indicate that the 3D discrete random Schrödinger operator does not exhibit Anderson localization with positive probability, in the sense that it has non-zero absolutely continuous spectrum with probability 1.*

As was pointed out by one of the reviewers, the provided data sets in fact (non-rigorously) suggest that a phase transition might happen near $c \approx 2$.

Analogous numerical experiments [14] in dimension 2 suggest extended states for $c \leqslant 0.6$.

We take a brief moment to allude to some aspects in the field, without attempting to be exhaustive by any means. Initiated by [4], the topic of Anderson localization has grown into a rich field that is researched by many communities. Even the meaning of localization can be interpreted in several different ways. The interesting work by Aizenman and Warzel concerning the Anderson model on the Bethe lattice (see [22] and the references therein) suggests that there might be more than one transition point, depending on how the transition is defined—spectrally, dynamically or statistically. (Due to memory requirements, numerical experiments are not feasible for the geometry of the Bethe lattice, see section 7.)

The Thouless criterion uses the ratio g of the so-called Heisenberg time over the so-called Thouless time as a measure of conductance. Localization is said to occur when $g < 1$ (see e.g. [11] for further explanation). In scaling theory the dimensionless ratio g is used as parameter in terms of which the scaling function gives the expected conductance depending on the system size. The scaling function is determined for many materials. At the moment it is not immediately clear how to directly compare scaling theory with the current experiment.

In the review article [22] (also see [8, 13]), several other definitions of localization can be found. Dynamical localization considers the supremum over all times of the part of the energy that we expect to find at distance R from the initial state δ_0. For example, strong exponential dynamical localization means exponential decay of this energy in terms of R. By the RAGE theorem, our numerical results indicate that we do not have strong dynamical localization (i.e. of all orders). See [8, section 1.2] for more details. Alternatively, transport is measured by an index in terms of the speed at which energy diffuses as time increases. For example, ballistic transport is a special case.

The problem is also studied via the local eigenvalue statistics under finite volume restrictions. Here delocalization is conjectured to correspond to GOE statistics. Localization lengths are investigated in terms of the band width W of a Hermitian and symmetric random matrix with uniformly distributed i.i.d. entries. The strength of the disorder is fixed. For example, in dimension $d = 1$, numerical, theoretical and symmetry arguments [7] indicate a localization length of order W^2 through a change in the local distribution of eigenvalues from Poisson to GOE statistics. Much research has been conducted in this field, see e.g. [5] and the many references therein. In two dimensions, the localization length is expected to be exponentially large in W (see [2]), while in higher dimensions the localization length is supposed to be macroscopic, and independent of W. A direct comparison with the work at hand is not available. In our work in the two-dimensional setting [14], the situation is more subtle. In particular, more concrete expressions for the localization lengths would provide useful information for the two-dimensional experiment.

The key analytical tool to our method is stated in proposition 2 below. Section 3 is devoted to a description of the numerical experiment. The numerical testing criterion we applied is given by numerical criterion 4 below. Our numerical findings and the conclusions can be found in section 4. In subsections 4.1 and 4.2, we study the averaged data and find further numerical validation of our method. In section 5 we verify the performance of the method in many examples. In subsection 5.3, we present the distribution of energies after repeated application of the random operator of a wave packet initially located at the origin. We briefly remark on computing and memory requirements in section 6 and on possible further projects in section 7.

2. Preliminaries

2.1. 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

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

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 $M_t\big|_{\oplus \int \mathcal{H}(z)\mathrm{d}\mu_{\mathrm{ac}}(z)}$. Similarly, define the singular, singular continuous and the pure point parts of T, denoted by T_{s}, T_{sc} and T_{pp}, respectively.

2.2. Key tool

As mentioned above delocalization is deterministic. Therefore demonstrating that it does not occur with probability zero is sufficient to determine delocalization.

This following result makes our numerical experiment possible as it suffices to check the evolution of only one vector through repeated operations by the Anderson Hamiltonian and 3D random Schrödinger operator.

Fix the vectors $\delta_{000} \in l^2(\mathbb{Z}^3)$ and $\delta_{111} \in l^2(\mathbb{Z}^3)$, i.e. 3-tensors with zero entries, except for the $(0, 0, 0)-$position and the $(1, 1, 1)-$position, respectively, which equal 1.

Notice that

$$D^n_{\omega,c} := \mathrm{dist}\big(\delta_{111}, \mathrm{span}\big\{H^k_\omega \delta_{000} : k = 0, 1, 2, \ldots, n\big\}\big) \tag{2}$$

describes the distance between the unit vector δ_{111} and the subspace obtained taking the closure of the span of the vectors $\delta_{000}, H_\omega\delta_{000}, H^2_\omega\delta_{000}, \ldots, H^n_\omega\delta_{000}$.

In numerical linear algebra, this space is called a Krylov subspace, and the Lanczos algorithm [12] provides a classical approach for finding an orthonormal basis. Our distance calculation (2) relies on the orthogonality of these vectors, iteratively updating the distance with each new Krylov vector.

Jakšić–Last [9] showed that certain vectors are cyclic for the singular part $(H_\omega)_{\mathrm{s}}$ almost surely. In [14] it was showed that the vector δ_{000} is cyclic for the singular part $(H_\omega)_{\mathrm{s}}$ almost surely. So in order to show that the existence of absolutely continuous spectrum, it suffices to show that δ_{000} is not cyclic for H_ω. In other words, $(H_\omega)_{\mathrm{ac}}$ is non-trivial. More concretely

Proposition 2. *Consider the discrete random Schrödinger operator given by equation (1). Let ω_i, $i \in \mathbb{Z}^3$, be i.i.d. random variables with uniform (Lebesgue) distribution on $[-c, c]$, $c > 0$. To prove delocalization (i.e. the existence of absolutely continuous spectrum with positive probability), it suffices to find $c > 0$ for which the distance*

$$D_{\omega,c} := \lim_{n\to\infty} D^n_{\omega,c} > 0 \tag{3}$$

with non-zero probability. (Notice that the limit exists by the monotone convergence theorem.)

Although it is not of immediate use to us here, we mention [1] that apart from the zero vector any vector is cyclic almost surely.

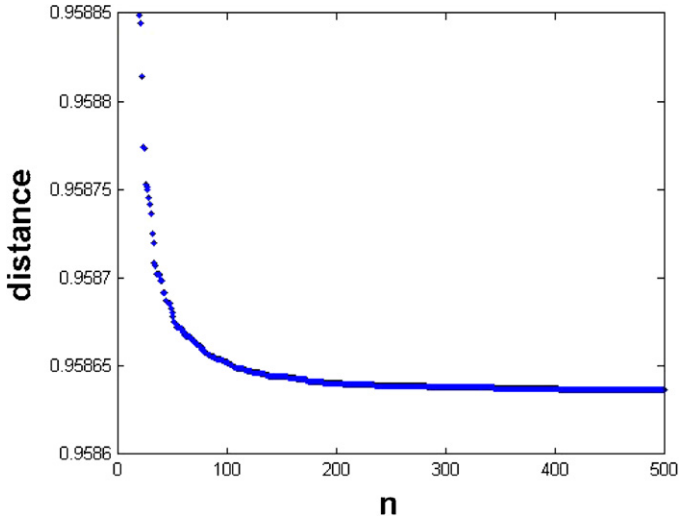

Figure 1. Typical trend for $D^n_{\omega,c}$ for $n = 500$ iterations and $c = 0.3$.

Remark 3. The converse of proposition 2 is not true. And 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 (3) for many or 'all' vectors (instead of just δ_{111}), it could be possible that the absolutely continuous part has multiplicity one and that δ_{000} is cyclic, that is, $l^2(\mathbb{Z}^3) = \text{clos span}\{H^k_\omega \delta_{000} : k \in \mathbb{N} \cup \{0\}\}$.

3. Method of numerical experiment

Consider the discrete random Schrödinger operator given by (1) with random variable ω distributed according to the hypotheses of proposition 2.

By proposition 2, we obtain delocalization if we can find $c > 0$ for which (3) happens with non-zero probability. Let us now explain precisely how we verify delocalization numerically, leading up to the numerical criterion 4 below.

In the numerical experiment, we initially fix c and fix one computer-generated realization of the random variable ω (with distribution in accordance to the hypotheses of proposition 2). We then calculate the distances $D^n_{\omega,c}$ for $n \in \{0, 1, 2, \ldots\}$.

Assuming that we know $D^n_{\omega,c}$ for $n = 0, \ldots, 500$, let us find a lower estimate for the limit

$$D_{\omega,c} = \lim_{n\to\infty} D^n_{\omega,c}.$$

Figure 1 displays a typical trend for the distance $D^n_{\omega,c}$ as n increases. Notice that the graph is decreasing, as is expected. Although it certainly appears that the limit does not go to 0, the graph could have logarithmic decay, approaching zero very slowly. To attain an estimate for $D_{\omega,c}$, which excludes the case of such slow decay, we re-scaled the graph by n^{-a}, $0.1 \leqslant a \leqslant 2$, so that the x-axis is inverted and the y-intercept, $y_{\omega,c}$, of a line of best fit will estimate $D_{\omega,c}$. The choice of the parameter a aims at identifying the correct power law in the decay of $D^n_{\omega,c}$,i.e., as $D_{\omega,c} + Bn^{-1/a}$. Because the majority of the change in distance occurs in the first $n = 119$ points, they were omitted when re-scaling in order to obtain a better approximation as $n \to \infty$.

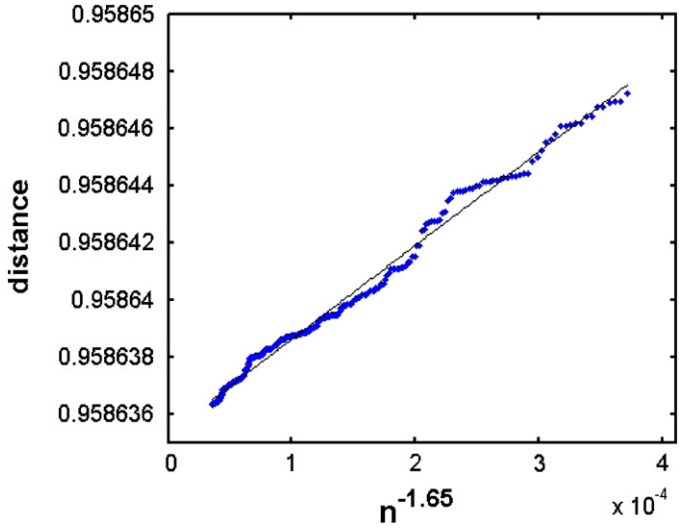

Figure 2. Figure 1 re-scaled using $a = 1.65$. Notice the fine y-scale and proximity to the y-axis.

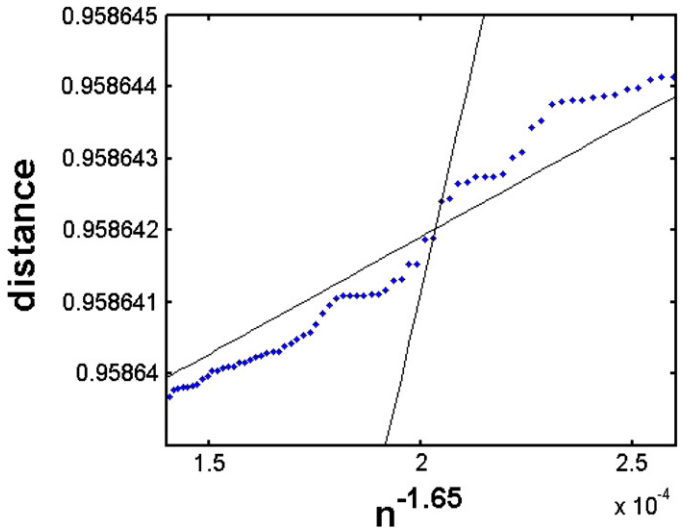

Figure 3. A closer look at a window in figure 2. The steeper line is used to determine $L_{\omega,c}$, and the line of best fit provides $y_{\omega,c}$. For this realization we have $L_{\omega,c} = 0.958\,5894$ and $y_{\omega,c} = 0.958\,6354$.

This does not change the sign of the rescaling factor. Figure 2 shows the re-scaled graph for $n = 120, 121, 122, \ldots, 500$.

Subsection 3.2 describes the choice of a and why, for small values of c, $D_{\omega,c}$ does not decay to 0.

Since an approximating line is only an estimate, for further confidence in our results, we also calculated the minimum y-intercept of all lines through two consecutive points and call it $L_{\omega,c}$ (see the steep line in figure 3). This is essentially the 'worst case,' and ought to

Table 1. For $n = 200$ we present the data of 15 realizations for each c between 0.1 and 5, and four realizations for each $c \geqslant 10$ and $c = 0$.

c	0	0.1	0.2	0.3	0.4	0.5	0.6
$\mathbb{P}$	1	1	1	1	1	1	0.93
$y_{\omega,c}$	0.958 69	0.9587	0.9586	0.9584	0.9582	0.9580	0.9573
$L_{\omega,c}$	0.958 63	0.9586	0.9585	0.9582	0.9579	0.9549	0.9454
c	0.7	0.8	0.9	1	1.5	2	2.5
$\mathbb{P}$	0.93	1	0.93	0.8	1	0.87	1
$y_{\omega,c}$	0.9573	0.9576	0.9559	0.9554	0.9540	0.9404	0.9447
$L_{\omega,c}$	0.9488	0.9569	0.9440	0.9455	0.9490	0.8748	0.8978
c	3	3.5	4	4.5	5	10	15
$\mathbb{P}$	0.8	0.67	0.73	0.83	0.6	1	0.75
$y_{\omega,c}$	0.9298	0.9105	0.8857	0.8843	0.8020	0.6458	0.0935
$L_{\omega,c}$	0.7765	0.6765	0.3895	0.3396	−0.3895	0.1098	−4.9095
c	20	25	30	35	40		
$\mathbb{P}$	0.25	0.75	0.25	0.75	0.25		
$y_{\omega,c}$	−2.1507	−1.9412	−3.3271	−3.3301	−9.7171		
$L_{\omega,c}$	−19.1480	−16.9035	−20.5279	−52.8516	−52.4555		

underestimate $D_{\omega,c}$, yielding the relationship

$$L_{\omega,c} \leqslant y_{\omega,c} \approx D_{\omega,c}.$$

We repeated this process for many values of c and multiple, different, computer-generated instances of the random variable ω. We took the minimum of $y_{\omega,c}$ and $L_{\omega,c}$ across all instances of ω, with the intent to demonstrate that $D_{\omega,c}$ is above 0 for many different ω.

In order to give confidence to our calculations to account for random error occurring in the computer, we introduce the following restrictions even though proposition 2 only requires that $L_{\omega,c} > 0$.

Numerical Criterion 4. *For a fixed value of c, we say that we have delocalization, if for at least 90% realizations we obtain* $L_{\omega,c} > .9 > 0$ *and* $y_{\omega,c} - L_{\omega,c}$ *is of order* 10^{-3}. *(Notice that we only need non-zero probability by proposition 2, and remark 3.)*

3.1. Notes on $L_{\omega,c}$

In the computation of $L_{\omega,c}$, we divide numerically the difference of approximately equal numbers $D^n_{\omega,c} - D^{n-1}_{\omega,c}$ by a small number $n^{-a} - (n-1)^{-a}$. This can have the effect of magnifying roundoff error in floating point computations. We explain how this potential problem is controlled here.

Most importantly, an *a posteriori* check is given by our restriction in the numerical criterion 4, where we assume that $y_{\omega,c} - L_{\omega,c}$ is of order 10^{-3}. Whence in the data used to support the main result 1 for $c \leqslant 2$, the error introduced by this potential problem is small. It might well be the case that this error causes some of the larger differences between $y_{\omega,c} - L_{\omega,c}$ for values of $c > 2$, but we do not draw conclusions from those data points. In general, the values of $L_{\omega,c}$ should be thought of as rather crude lower estimates of the actual y-intercept.

Finally, please note that we only examine the values of $L_{\omega,c}$. In particular, we do not substitute them into a subsequent calculation.

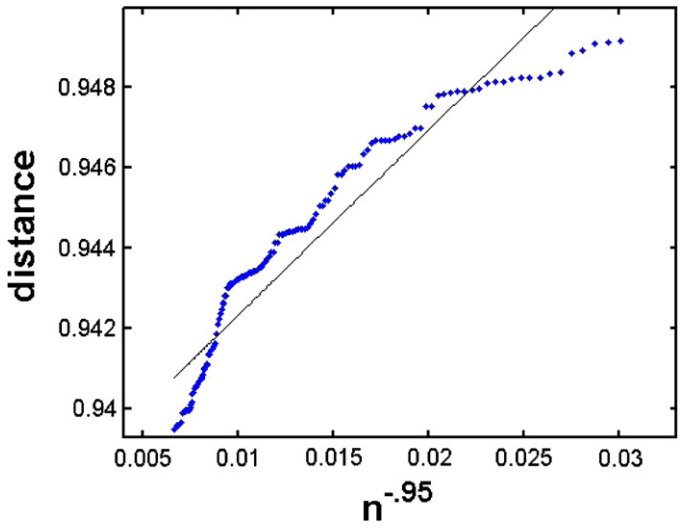

Figure 4. A concave graph that yields no usable data ($c = 6$). The concave shape of the data implies that $y_{\omega,c}$ is not necessarily a lower bound for $D_{\omega,c}$.

3.2. 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 figure 2) satisfies the least square property; that is, the error with respect to square-norm 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.

To find optimal a, we used the mesh $a = 0.05 : 0.05 : 2$. In table 1 we give the percentage of usable trials (those for which an optimal $a \geqslant 0.1$ was found) for many values of c. Trials are not usable if the re-scaling parameter $a = 0.05$ yields a concave graph. If this happens, we do not obtain any information (according to remark 3). See figure 4 below. Note that a small value ($\leqslant 0.05$) of a is 'bad', since the graph rescaled with $a = 0.05$ will be concave, and thus it is not expected for a line of best fit to underestimate the limit of the distance.

A positive re-scaling factor implies that the graph in figure 1 will not decay to zero. Indeed, using a re-scaling factor smaller than the optimal one will result in a convex graph for the distances $D^n_{\omega,c}$. And the y-intercept of the line lies below the value expected for $D^\infty_{\omega,c}$.

4. Conclusions

As mentioned in section 3, for a fixed c we chose several realizations ω. For every value of c, we took the minimum of the resulting quantities for $y_{\omega,c}$ and $L_{\omega,c}$ (the y-intercept of the approximating line and the minimum y-intercept of the lines passing through any two consecutive points, respectively).

We present our observations for the Numerical Criterion 4 for $n = 500$. For fixed disorder, we will comment in subsection 4.1 on the re-scaling parameters of averages over the distances $D^n_{\omega,c}$, $n = 0, 1, 2, \ldots, 200$ and $n = 0, 1, 2, \ldots, 500$.

The data in table 1 documents the data obtained for n=200 by taking 15 realizations for each c between 0.1 and 5, and four realizations for each $c \geqslant 10$ and $c = 0$. By $\mathbb{P}$ we denote the probability of finding a re-scaling factor $a \in [0.1, 2]$. For $c \leqslant 2$ a total of 173 out of 180 trials yielded a good re-scaling factor (while by proposition 2 and remark 3 extended states follow

Table 2. For $n = 500$ we present the data of 13 trials for $c \leqslant 1$ and four trials for all other values.

c	0	0.1	0.2	0.3	0.4	0.5	0.6
$\mathbb{P}$	1	1	1	1	1	1	1
$y_{\omega,c}$	0.958 69	0.958 66	0.958 56	0.958 38	0.958 13	0.9578	0.9575
$L_{\omega,c}$	0.958 68	0.958 65	0.958 54	0.958 33	0.958 04	0.9577	0.9573
c	0.7	0.8	0.9	1	1.5	2	2.5
$\mathbb{P}$	1	1	1	1	1	1	1
$y_{\omega,c}$	0.9570	0.9565	0.9559	0.9552	0.9520	0.9518	0.9523
$L_{\omega,c}$	0.9568	0.9563	0.9557	0.9549	0.9505	0.9489	0.9362
c	3	3.5	4	4.5	5	10	15
$\mathbb{P}$	1	1	1	1	1	0.5	0.75
$y_{\omega,c}$	0.9451	0.9556	0.9405	0.9271	0.9149	0.2600	−0.5149
$L_{\omega,c}$	0.9403	0.9318	0.9184	0.8829	0.7460	−5.0995	−13.8931
c	20	25	30	35	40		
$\mathbb{P}$	0.5	0.5	0.5	0.75	0.75		
$y_{\omega,c}$	−2.8571	−1.5957	−2.5188	−2.2407	−3.1759		
$L_{\omega,c}$	−22.8975	−33.8175	−24.2294	−27.5353	−35.9414		

from showing non-zero probability). The data for $c > 2$ is included to show the reader what goes wrong. In accordance with remark 3, the data for $c > 2$ does not have any implications.

While for some $c \geqslant 2.5$, we have $\mathbb{P} \geqslant 0.9$ the difference between $y_{\omega,c}$ and $L_{\omega,c}$ is relatively large, which means that the line from taking the least square approximation is likely not a good approximation for the distances.

We also repeated the experiment for $n = 500$ and table 2 documents the findings. In these trials, the first 119 entries were removed instead of the first 44, as in the $n = 200$ case. This larger crop makes the data more stable by giving better estimates for $y_{\omega,c}$ and $L_{\omega,c}$ and by more consistently finding a usable rescaling factor a. We ran 13 trials for $c \leqslant 1$ and 4 trials for all other values.

A good rescaling factor a was found for all 143 of the trials for $c \leqslant 1$ and all $c \leqslant 2.0$ satisfy Criterion 4, an improvement from the $n = 200$ case. Hence the final conclusion of this numerical experiment is precisely the main result 1. According to remark 3 and criterion 4, for $c \geqslant 4$, we do not have any conclusion.

4.1. Averages

In table 3, for each fixed c, we averaged the distances $D^n_{\omega,c}$, $n = 0, 1, 2, \ldots, 200$, of all our realizations. For those averaged distances, we determined the re-scaling parameters $\tilde{a}$, as well as $\tilde{y}_c$ and $\tilde{L}_c$ in analogy. The significance of our findings is that the re-scaling factors $\tilde{a}$ are 'roughly' decreasing and rather well-behaved for $c \leqslant 1.5$. For larger disorder, $\tilde{a}$ becomes even less stable, and cannot even be found for large enough disorder.

In table 4 we document the analogous quantities for the $n = 500$ trials. Note that there is no rescaling factor for $c = 20$, while there is for that c in the $n = 200$ trials. The data sets are not related to each other, aside from sharing the same disorder c.

4.2. Comparing n = 200 with n = 500

The $n = 500$ data gave better results than the $n = 200$ data. The probability of finding a useable rescaling factor for $n = 500$ was higher than that of $n = 200$ for all but two values of c. The average rescaling factor $\tilde{a}$ was similar between the two data sets. Finally, $y_{\omega,c} - L_{\omega,c}$

Table 3. Data for the averaged distances for $n = 200$. The re-scaling factors $\tilde{a}$ are 'roughly' decreasing and well-behaved for $c \leqslant 1.5$. For larger disorder, $\tilde{a}$ becomes less stable.

c	0	0.1	0.2	0.3	0.4	0.5	0.6
$\tilde{a}$	2	1.9	1.65	1.5	1.3	1.1	0.95
$\tilde{y}_c$	0.958 69	0.958 69	0.958 65	0.958 62	0.958 53	0.958 46	0.958 43
$\tilde{L}_c$	0.958 63	0.958 62	0.958 57	0.958 54	0.958 46	0.958 39	0.958 35
c	0.7	0.8	0.9	1	1.5	2	2.5
$\tilde{a}$	1.3	1	0.8	0.9	1	0.6	0.85
$\tilde{y}_c$	0.9582	0.9584	0.9579	0.9578	0.9571	0.9539	0.9544
$\tilde{L}_c$	0.9581	0.9583	0.9576	0.9576	0.9566	0.9532	0.9530
c	3	3.5	4	4.5	5	10	15
$\tilde{a}$	0.65	0.3	0.55	0.65	0.5	0.5	0.3
$\tilde{y}_c$	0.9485	0.9466	0.9345	0.9414	0.9217	0.8332	0.5312
$\tilde{L}_c$	0.9461	0.9369	0.9255	0.9362	0.9077	0.6567	−0.2021
c	20	25	30	35	40		
$\tilde{a}$	0.1	0.85	0.45	N/A	N/A		
$\tilde{y}_c$	−0.3300	0.2928	−.0722	-3.0990	−5.1751		
$\tilde{L}_c$	−3.2230	−0.1794	−0.8701	−16.7931	−17.1925		

Table 4. Data for the averaged distances when taking $n = 500$.

c	0	0.1	0.2	0.3	0.4	0.5	0.6
$\tilde{a}$	2	1.75	1.35	1.2	1.1	1	1.05
$\tilde{y}_c$	0.958 69	0.958 68	0.958 64	0.958 61	0.958 55	0.958 47	0.958 23
$\tilde{L}_c$	0.958 68	0.958 67	0.958 63	0.958 60	0.958 53	0.958 44	0.958 19
c	0.7	0.8	0.9	1	1.5	2	2.5
$\tilde{a}$	1.05	1.45	1.15	0.65	0.7	0.6	0.9
$\tilde{y}_c$	0.958 06	0.957 86	0.958 05	0.957 95	0.9558	0.9543	0.9561
$\tilde{L}_c$	0.958 01	0.957 81	0.957 99	0.957 87	0.9551	0.9531	0.9548
c	3	3.5	4	4.5	5	10	15
$\tilde{a}$	0.55	0.6	0.65	0.35	0.15	0.25	0.3
$\tilde{y}_c$	0.9506	0.9571	0.9479	0.9390	0.9244	0.7053	0.4991
$\tilde{L}_c$	0.9481	0.9416	0.9406	0.9287	0.8987	0.2854	−0.8246
c	20	25	30	35	40		
$\tilde{a}$	N/A	0.1	0.8	N/A	N/A		
$\tilde{y}_c$	0.0569	−1.9658	−1.1264	0.0906	−2.1984		
$\tilde{L}_c$	−1.2457	−9.1777	−5.9900	−0.4750	−9.3414		

was smaller for the $n = 500$ data for small c, suggesting that the approximation given by $y_{\omega,c}$ is better.

5. Further validation of the method and the numerical experiments

We have conducted the following tests. Most important is the *a posteriori* test of orthogonality in the Lanczos algorithm in subsection 5.4.

5.1. Free discrete three-dimensional Schrödinger operator

When we apply the free discrete Schrödinger operator $H = H_0$ to the vector δ_{000}, it immediately becomes clear that $H\delta_{000}$ as well as all vectors $H^n\delta_{000}$, $n \in \mathbb{N} \cup \{0\}$, are symmetric with respect

to the origin. In dimension $d = 3$, it is not hard to see that the distance between δ_{111} and the orbit of δ_{000} under H is at least $\frac{\sqrt{7}}{2\sqrt{2}} \approx 0.9354$. Indeed, we have

$$\operatorname{dist}(\delta_{111}, \operatorname{clos\,span}\{H^n\delta_{000} : n \in \mathbb{N} \cup \{0\}\}) > \min_x \operatorname{dist}(u_x, \delta_{111}) = \frac{\sqrt{7}}{2\sqrt{2}},$$

where

$$u_x = x\delta_{-1-1-1} + x\delta_{1-1-1} + x\delta_{-11-1} + x\delta_{-1-11}x\delta_{-1-11} + x\delta_{-11-1} + x\delta_{1-1-1} + x\delta_{111},$$

the eight vertices of the length 2 cube centered at $(0, 0, 0)$.

In the experiments for the free discrete 3D Schrödinger operator we obtained a y-intercept of the approximating line approximately equals 0.958 6936. The re-scaled graph of distances still had a convex shape, so the actual distance as $n \to \infty$ would be bigger. In fact, we have extracted our data an upper estimate of $0.958\,6939 \approx D^{500}_{\omega,0}$. Therefore, the distance must lie in the interval [0.958 6936, 0.958 6939].

5.2. Orthogonalization process

The $c = 0$ case shows a decrease in distance on only every other step. The symmetry caused by the absence of random perturbations means the 3-tensor after orthogonalization has alternating diamonds of zero and nonzero entries radiating from the origin, meaning the distance decreases every second application of the operator, when there is a nonzero entry in the $(1, 1, 1)$-position.

5.3. Evolution under H_ω of the bulk for small values of c

We observe the bulk distribution which determines the distance from the origin where we are most likely to find an electron. Here, distance is measured by the taxicab method, so elements of the same distance form a diamond in the 3D integer lattice. The bulk at this distance is the Euclidean norm of the elements constituting the diamond.

To be precise, we consider the elements of the vector m_{500} and define

$$E(l, n) = \sqrt{\sum_{|i|+|j|+|k|=l} (m_n)^2_{i,j,k}} \tag{4}$$

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

Figure 5 is the result of averaging four sets of data for values of c ranging from 0.1 to 1. As expected, the energy remains closer to the origin as disorder increases.

5.4. Lanczos and orthogonality

The Lanczos algorithm is known to lose orthogonality in many instances, which could cast doubt on our distance calculations. To test the accuracy for our problem, we stored the entire Krylov subspace generated on a smaller problem instance ($n = 150$) and stored these as columns of a matrix K. The quantity $Q = \|K^T K - I\|_\infty$ should deviate with zero in proportion to the loss of orthogonality. In table 5, we measure the matrix ∞ norm for realizations for several cases of c. We see that the Krylov vectors in these cases are in fact quite close to orthogonal especially for $c \leqslant 2.0$, although the orthogonality seems to decrease as c grows.

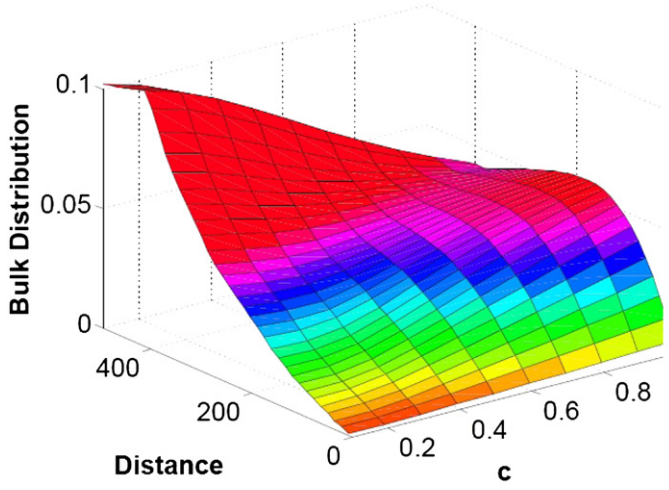

Figure 5. Bulk distribution of m_{500} for the disorders $c = 0.1 : 0.1 : 1$, averaged over the four realizations for each value of c.

Table 5. Measuring the loss of orthogonality in Lancozs' algorithm via the ∞ norm of the matrix of orthogonalized vectors when $n = 150$.

c	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	
Q	2.01e-11	7.2e-11	4.8e-11	4.4e-11	6.9e-11	9.6e-11	3.1e-11	4.8e-11	
c	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0
Q	4.5e-11	2.4e-11	6.1e-9	1.3e-10	4.0e-10	9.2e-11	1.2e-9	7.3e-8	5.1e-9

6. On computing and memory requirements

Using methodology similar to that in [14], all of the information contained in the 3-tensor is stored in one information vector. For this method, because of how the Hamiltonian acts, it is important for computing purposes that each point in the 3-tensor is stored in a position such that its neighbors along a coordinate axis are a consistent distance from that point in the vector. This methodology allows the vector to be half the size necessary for containing every point in a 3-tensor, but still approximately twice as large as is necessary. In order to explore localization in higher dimensions, a more efficient method is needed since a generalization of this code for dimension d has time complexity $\mathcal{O}(n^d)$.

We began with an implementation of the three-term Lanczos recurrence using this special storage pattern. At each step, we compute the new orthogonal iterate and hence the distance of $\delta_{1,1,1}$ to the updated subspace. Because of the 3D nature of our calculations, the memory requirements are quite large. With our storage pattern, each of our three vectors in the Lanczos recurrence requires, to leading term, $4n^3$ double-precision numbers. We also require a vector of the same size to store the random potential ω, giving minimal storage requirements of about $16n^3$ double-precision numbers or, equivalently, $128n^3$ bytes. When $n = 200$, this corresponds to just over a gigabyte of memory, and about 16 GB when $n = 500$.

Our simulations were run on a Dell Precision workstation with dual eight-core Intel Xeon E5-2680 processors running at 2.7 GHz with 128GB of RAM. We used gfortran version 4.4.7 with flags `-O3 -ftree-vectorizer-verbose=2 -msse2 -funroll-loops`

`-ffast-math`, which, among other optimizations, enables instruction-level superscalar parallelism. After prototyping our approach in MATLAB, we translated the code into FORTRAN90, which provided much higher performance. Still, each realization for $n = 500$ required over an hour to perform. To make use of multiple processing cores, we wrapped our FORTRAN90 routine into Python using the `f2py` package [16] and dispatched multiple jobs using Python's `multiprocessing` module. Due to the very large memory requirements, we were only able to run about four jobs concurrently.

Both Python and MATLAB use a standard Mersenne Twister pseudorandom number generator. Although we have not particularly checked for possible correlation between sites, the Mersenne Twister has a period of $2^{19\,937} - 1$ and passes numerous tests for statistical randomness. The number of sites, even in three dimensions with $n = 500$, is far lower than the Mersenne Twister's period.

7. Further projects

An immediate area for further exploration would be to consider various geometries, rather than simply the n-dimensional lattice. One geometry of interest is the Sierpinski gasket, starting at one corner and building the various triangles as n increases. Preliminary results indicate that a program modeling the free random Schrödinger operator on this geometry should run with time complexity $\mathcal{O}(n^{\frac{\ln(3)}{\ln(2)}})$.

Unfortunately, it is not possible to conduct this numerical experiment for the Bethe lattice due to memory requirements. Time complexity for the Bethe lattice with site number K being $\mathcal{O}(K^n)$ means that with our current computers, we can roughly accomplish only $n = 30$ iterations in the binary tree ($K = 2$). One should note that even increasing the number of cores by a factor of K (via parallelization) merely increases the number of steps by 1.

It should be possible to adapt the algorithm in subsection 5.3 to compare with various notions of dynamical localization.

Acknowledgments

The authors would like to thank the reviewers for adding value to the current work via useful comments and suggestions. The work of Liaw and King was supported by DMS-1261687.

References

[1] Abakumov E, Liaw C and Poltoratskiĭ A 2013 Cyclic vectors for rank-one perturbations and Anderson-type Hamiltonians *J. Lond. Math. Soc.* **88** 523–37
[2] Abrahams E, Anderson P W, Licciardello D C and Ramakrishnan T V 1979 Scaling theory of localization: absence of quantum diffusion in two dimensions *Phys. Rev. Lett.* **42** 673–6
[3] Aizenman M and Molchanov S 1993 Localization at large disorder and at extreme energies: an elementary derivation *Commun. Math. Phys.* **157** 245–78
[4] Anderson P W 1958 Absence of diffusion in certain random lattices *Phys. Rev.* **109** 1492–505
[5] Erdös L and Knowles A 2011 Quantum diffusion and eigenfunction delocalization in a random band matrix model *Commun. Math. Phys.* **303** 509–54
[6] Fröhlich J and Spencer T 1983 Absence of diffusion in the tight binding model for large disorder of low energy *Commun. Math. Phys.* **88** 151–84
[7] Fydorov Y V and Mirlin A D 1991 Scaling properties of localization in random band matrices: a σ-model approach *Phys. Rev. Lett.* **67** 2405–9
[8] Hundertmark D 2008 A short introduction to Anderson localization *Analysis and Stochastics of Growth Processes and Interface Models* (Oxford: Oxford University Press) pp 194–218

[9] Jakšić V and Last J 2006 Simplicity of singular spectrum in Anderson-type Hamiltonians *Duke Math. J.* **133** 185–204
[10] Kirsh W 2008 An invitation to random Schrödinger operators *Panoramas et synthèses* **25** 1–119
[11] Lagendijk A, Tigglen B van and Wiersma D S 2009 Fifty years of Anderson localization *Phys. Today* **82** 24–9 (featured article, august 2009)
[12] Lanczos C 1950 An iteration method for the solution of the eigenvalue problem of linear differential and integral operators *J. Res. Nat. Bureau Stand.* **45** paper 2133
[13] Last Y 2007 Exotic spectra: a review of Barry Simon's central contributions *Proc. Symp. in Pure Mathematics* vol 76.2 pp 697–712
[14] Liaw C 2013 Approach to the extended states conjecture *J. Stat. Phys.* **153** 1022–38
[15] Liaw C 2010 Deterministic spectral properties of Anderson-type Hamiltonians arXiv:1009.1353
[16] Peterson P 2009 F2PY: a tool for connecting Fortran and Python programs *Int. J. Comput. Sci. Eng.* **4** 296–305
[17] Schenk O, Bollhöfer M and Römer R A 2008 On large-scale diagonalization techniques for the Anderson model of localization *SIAM Rev.* **50** 91–112
[18] Schenker J 2013 How large is large? Estimating the critical disorder for the Anderson model arXiv-math:1305.6987v1
[19] Simon B 2005 Spectral analysis of rank-one perturbations and applications *Trace Ideals and Their Applications (Mathematical Surveys and Monographs)* 2nd edn (Providence, RI: Amer. Math. Soc.)
[20] Stein J and Krey U 1979 Numerical studies on the Anderson localization problem *Zeitschrift Für Phys.* B **34** 287–96
[21] Trefethen L N and Bau D 1997 *Numerical Linear Algebra* (Philadelphia: SIAM)
[22] Warzel A 2012 Surprises in the phase diagram of the Anderson model on the Bethe lattice *Plenary lecture given at the International Congress on Mathematical Physics, AalborgAug. 6-11* arXiv:1212.4367