TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 372, Number 5, 1 September 2019, Pages 3619–3667
https://doi.org/10.1090/tran/7832
Article electronically published on April 29, 2019

LOCALIZATION FOR THE ONE-DIMENSIONAL ANDERSON MODEL VIA POSITIVITY AND LARGE DEVIATIONS FOR THE LYAPUNOV EXPONENT

VALMIR BUCAJ, DAVID DAMANIK, JAKE FILLMAN, VITALY GERBUZ, TOM VANDENBOOM, FENGPENG WANG, AND ZHENGHE ZHANG

Abstract. We provide a complete and self-contained proof of spectral and dynamical localization for the one-dimensional Anderson model, starting from the positivity of the Lyapunov exponent provided by Fürstenberg's theorem. That is, a Schrödinger operator in $\ell^2(\mathbb{Z})$ whose potential is given by independent, identically distributed (i.i.d.) random variables almost surely has pure point spectrum with exponentially decaying eigenfunctions, and its unitary group exhibits exponential off-diagonal decay, uniformly in time. We also explain how to obtain analogous statements for extended CMV matrices whose Verblunsky coefficients are i.i.d., as well as for half-line analogues of these models.

Contents

1. Introduction

1.1. **The goal in a nutshell.** This paper is centered around the following fundamental result.

Received by the editors November 30, 2018 and, in revised form, February 19, 2019.
2010 *Mathematics Subject Classification.* Primary 35J10; Secondary 81Q10.
The first, second, fourth, and fifth authors were supported in part by NSF grant DMS-1361625.
The main idea of the new proof of the LDT in Section 3 was communicated to the second and seventh authors by Artur Avila while they were visiting IMPA, Rio de Janeiro. They would like to thank Artur Avila for sharing his idea, and IMPA for the hospitality.
The third author was supported in part by an AMS-Simons travel grant, 2016–2018.
The sixth author was supported by CSC (No. 201606330003) and NSFC (No. 11571327).
The seventh author was supported in part by an AMS-Simons travel grant, 2014–2016.

Theorem 1.1 (Spectral localization for the one-dimensional (1D) Anderson model)**.** *Consider the family $\{H_\omega\}_{\omega\in\Omega}$ of random Schrödinger operators, acting in $\ell^2(\mathbb{Z})$ via*

$$[H_\omega\psi](n) = \psi(n+1) + \psi(n-1) + V_\omega(n)\psi(n),$$

where the potential V_ω is given by independent, identically distributed random variables. It is assumed that the common distribution has a compact support that contains at least two elements. Then, almost surely, H_ω is spectrally localized; that is, it has pure point spectrum with exponentially decaying eigenfunctions.

We will provide a complete and relatively elementary derivation of this result, starting from the classical Fürstenberg theorem about products of random matrices. In particular, as opposed to all previously published proofs of this result, we will not appeal to multiscale analysis as a black box.

1.2. **Background and context.** Let us describe the context and the relevance of Theorem 1.1. The Nobel Prize winning work of Philip Warren Anderson suggested that randomness leads to the localization of quantum states in suitable energy regions that depend on the strength of the randomness. A particular signature of such a localization effect is a spectral localization statement, which asserts that the spectral type of the associated Schrödinger operator is pure point in suitable energy regions, and the eigenfunctions corresponding to the eigenvalues in these energy regions decay exponentially.

For the sake of concreteness, let us consider the standard Anderson model, which is just the d-dimensional generalization of the operator family considered in Theorem 1.1. That is, given a probability measure $\widetilde{\mu}$ on $\mathbb{R}$ whose topological support is compact and contains at least two points, we consider the product space $\Omega = (\operatorname{supp}\widetilde{\mu})^{\mathbb{Z}^d}$ and the product measure $\mu = \widetilde{\mu}^{\mathbb{Z}^d}$. For every $\omega\in\Omega$ and $n\in\mathbb{Z}^d$, we set $V_\omega(n) = \omega_n$. This defines, for $\omega\in\Omega$, a potential $V_\omega : \mathbb{Z}^d \to \mathbb{R}$, and in turn a Schrödinger operator

$$[H_\omega\psi](n) = \sum_{|m-n|_1=1} \psi(m) + V_\omega(n)\psi(n)$$

in $\ell^2(\mathbb{Z}^d)$. Standard ergodicity arguments show that the spectrum and the spectral type of H_ω are almost surely independent of ω; that is, there exist sets $\Sigma, \Sigma_{\mathrm{pp}}, \Sigma_{\mathrm{sc}}, \Sigma_{\mathrm{ac}}$, and a set $\Omega_0\subseteq\Omega$ of full μ-measure such that, for every $\omega\in\Omega_0$, we have $\sigma(H_\omega) = \Sigma$ and $\sigma_\bullet(H_\omega) = \Sigma_\bullet$, $\bullet\in\{\mathrm{pp},\mathrm{sc},\mathrm{ac}\}$. It is not too hard to show that

$$\Sigma = [-2d, 2d] + \operatorname{supp}\widetilde{\mu}.$$

The assumption that $\operatorname{supp}\widetilde{\mu}$ is compact ensures that these operators are bounded; the real-valuedness of the potential ensures that they are also self-adjoint. The boundedness is not crucial, and one could in fact consider probability measures $\widetilde{\mu}$ with unbounded support. However, the phenomenon of Anderson localization already occurs in the bounded case, and many authors limit their attention to this case—as do we. Furthermore, the assumption that $\operatorname{supp}\widetilde{\mu}$ contain more than one point excludes the trivial case of a constant potential, for which the Anderson localization phenomenon is obviously impossible (the spectrum is purely absolutely continuous in this case).

The spectral signature of Anderson localization is now the following. There exists a set $\Sigma_{\mathrm{AL}}\subseteq\Sigma$, which is a finite union of nondegenerate intervals such that $\Sigma_{\mathrm{AL}}\subseteq\Sigma_{\mathrm{pp}}$ (or, really, "=") and $\operatorname{int}\Sigma_{\mathrm{AL}}\cap\Sigma_{\mathrm{sc}} = \operatorname{int}\Sigma_{\mathrm{AL}}\cap\Sigma_{\mathrm{ac}} = \emptyset$. This

means that almost surely the spectrum of H_ω is pure point on Σ_{AL}, and hence H_ω has a set of eigenvalues that is dense in Σ_{AL}. The additional feature is that the associated eigenfunctions decay exponentially. The size of Σ_{AL} relative to the size of Σ depends on the dimension and the strength of the randomness. It is expected that $\Sigma_{\mathrm{AL}} = \Sigma$ for $d = 1$ and $d = 2$, and that for $d \geq 3$ we only have $\Sigma_{\mathrm{AL}} \subseteq \Sigma$. However, $\Sigma_{\mathrm{AL}} = \Sigma$ does hold in the case $d \geq 3$ when the randomness is strong enough. More specifically, each of the connected components of Σ_{AL} is a neighborhood of a boundary point of Σ. The length of these intervals grows with increasing randomness, up to a point where they cover all of Σ. From this strength of randomness onward, we have $\Sigma_{\mathrm{AL}} = \Sigma$.

Notice that Theorem 1.1 is precisely the expected statement $\Sigma_{\mathrm{AL}} = \Sigma$ for the case $d = 1$. The localization part of the expected statement for the case $d \geq 3$ is known for sufficiently regular $\widetilde{\mu}$ as well, but proving that $\overline{\Sigma \setminus \Sigma_{\mathrm{AL}}} = \Sigma_{\mathrm{ac}} \neq \emptyset$ for sufficiently small randomness in the case $d \geq 3$ is the main open problem in the study of random Schrödinger operators. The expected statement for $d = 2$ is not known. The best known result in the case $d = 2$ is the same as the best known result for $d \geq 3$, but no better. That is, one knows localization (under suitable assumptions on $\widetilde{\mu}$) only in neighborhoods of the boundary points of Σ, not in all of Σ, as is expected. This problem ("prove spectral localization for the two-dimensional Anderson model throughout the whole spectrum for any strength of randomness") is the second main open problem in the field. The third main open problem is to establish the localization result in dimensions $d \geq 2$, which, as pointed out above, is known under suitable assumptions on $\widetilde{\mu}$, for *any* single-site measure $\widetilde{\mu}$. For example, the case where $\widetilde{\mu}$ has a nonzero pure point part is for the most part not covered by known localization results in higher dimensions yet, though there has been some recent progress due to Ding and Smart in dimension 2 [22].

Proofs of spectral localization results come in two general flavors. Some of them use methods that are strictly 1D, and others work in any dimension. Among the strictly 1D proofs, we mention the Kunz–Souillard approach [17, 21, 41] and the approach via spectral averaging [49, 52]. Proofs that work in arbitrary dimension include those based on multiscale analysis (MSA) [23–26, 30] and the fractional moment method [2–4]. However, most of these approaches are limited in terms of the single-site distributions to which they apply. Except for the MSA approach, all of the approaches require $\widetilde{\mu}$ to have a nontrivial absolutely continuous component, or to even be purely absolutely continuous. This leaves MSA as the only method heretofore available in cases where $\widetilde{\mu}$ is purely singular. The special case in which $\widetilde{\mu}$ is supported on precisely two points is the hardest; this is commonly referred to as the Bernoulli case, and the operator family in this case is called the Bernoulli–Anderson model. The approach based on MSA is the most complex among the available approaches. This makes the treatment of the Bernoulli case difficult, even in one dimension.

On the other hand, the case of one dimension is special in that Fürstenberg's theorem about products of random matrices provides compelling evidence that spectral localization holds for all single-site distributions in this case. It shows that, for all $\widetilde{\mu}$ and throughout Σ, solutions of the generalized eigenvalue equation have a strong tendency to be either exponentially increasing or exponentially decreasing. Coupled with the general fact that all spectral measures are supported on the

set of energies that admit polynomially bounded solutions, this should imply that all such polynomially bounded solutions are in fact exponentially decreasing, and hence are eigenfunctions, and hence the spectrum is pure point because the spectral measures are supported on such energies, which turn out to be eigenvalues by the argument above. The starting point of this argument, the exponential behavior of solutions, has no analogue in higher dimensions, and this is the reason why spectral localization is known at all energies in one dimension, is not known in two dimensions, and is in fact expected to fail in general for dimensions greater than 2.

Alas, the argument outlined in the previous paragraph has a flaw which has to do with exceptional sets and uncountable unions of zero-measure sets. Localization proofs in one dimension that are based on the output of Fürstenberg's theorem (i.e., all proofs other than those using the Kunz–Souillard method) must address this flaw. That is, they do implement the general strategy, but they address the complications that arise when uncountable unions of exceptional sets of zero measure are taken.

When $\widetilde{\mu}$ has a nontrivial absolutely continuous component, this is taken care of in a very elegant way by spectral averaging. One of the fundamental properties of spectral averaging (namely, that the average of spectral measures turns out to be Lebesgue measure) allows one to simply ignore sets of zero Lebesgue measure, and this shows in effect that the flaw is a nonissue in this case.

On the other hand, when $\widetilde{\mu}$ is singular, the only option up to this point has been to verify the assumptions that are necessary to start an MSA throughout the spectrum, which consists of proving a Wegner estimate and establishing an initial length scale estimate. The former is difficult to establish in the Bernoulli case, and the latter follows from the positive Lyapunov exponents provided by Fürstenberg's theorem. Thus, the work necessary to deal with the 1D case in full generality (i.e., including the Bernoulli case) mainly focused on establishing a Wegner estimate. This was accomplished, by different methods, in two papers: by Carmona, Klein, and Martinelli in 1987 [12] and by Shubin, Vakilian, and Wolff in 1998 [47]. Once all of the ingredients are in place, the MSA machine produces the desired spectral localization statement, and hence Theorem 1.1.

Because of the challenges involved in verifying the Wegner estimate in order to run an MSA, as well as the preponderance of additional tools available in one dimension, it has therefore been a well-recognized problem in the random operator community to find a more direct way of going from the positive Lyapunov exponents provided by Fürstenberg's theorem to the spectral localization statement contained in Theorem 1.1, that is, to find a one-dimensional proof of this one-dimensional result. This is precisely what we accomplish in this paper.

1.3. **What this paper accomplishes.** We provide a complete derivation of Theorem 1.1 from Fürstenberg's theorem without having to appeal to the MSA machine as a black box. Our proof not only is more direct and conceptually tight, it is naturally also substantially shorter than the previous proofs, obtained by combining either [12] or [47] with the full-blown proof that running an MSA yields localization.[1]

[1]On a related note, two of us wanted to include the localization result for the 1D Bernoulli–Anderson model in the forthcoming monograph [15], and for this we needed a self-contained proof.

To give the reader an initial impression of one of the key realizations in this paper, let us discuss the interplay between Lyapunov behavior and the existence of polynomially bounded solutions in more detail.

The difference equation

$$u(n+1)+u(n-1)+V_\omega(n)u(n)=Eu(n) \tag{1.1}$$

admits a two-dimensional solution space, as any two consecutive values of u determine all other values. Fixing $(u(0),u(-1))^\top$ as the point of reference, the linear map taking this vector to $(u(n),u(n-1))^\top$ is given by the so-called transfer matrix $M_n^E(\omega)$. Ergodicity of the full shift implies that, for each E, there are $L(E)\geq 0$ and $\Omega_-^E,\Omega_+^E\subseteq\Omega$ with $\mu(\Omega_-^E)=\mu(\Omega_+^E)=1$ such that

$$L(E)=\begin{cases}\lim_{n\to\infty}\frac{1}{n}\log\|M_n^E(\omega)\| & \text{for } \omega\in\Omega_+^E,\\ \lim_{n\to-\infty}\frac{1}{|n|}\log\|M_n^E(\omega)\| & \text{for } \omega\in\Omega_-^E.\end{cases}$$

The number $L(E)$ or the function $L(\cdot)$ are called the *Lyapunov exponent*. Fürstenberg's theorem implies that in fact $L(E)>0$ for every E. Now if $\omega\in\Omega_+^E$ (resp., $\omega\in\Omega_-^E$), then, due to a result of Oseledec, there is a 1D subspace of the solution space in which every element decays exponentially at ∞ (resp., $-\infty$), while every linearly independent solution grows exponentially at ∞ (resp., $-\infty$). More precisely, the rate is given by the Lyapunov exponent; that is, the decaying solutions obey

$$\lim_{n\to\pm\infty}\frac{1}{|n|}\log(|u(n)|^2+|u(n-1)|^2)^{1/2}=-L(E),$$

while growing solutions obey

$$\lim_{n\to\pm\infty}\frac{1}{|n|}\log(|u(n)|^2+|u(n-1)|^2)^{1/2}=L(E).$$

These decay/growth statements are sometimes referred to as *Lyapunov behavior*.

Thus, for every E, we have Lyapunov behavior at both $\pm\infty$ for almost every ω. Suppose that we could turn this around and claim, for almost every ω, Lyapunov behavior at both $\pm\infty$ for every E. Let us call this statement the *Anderson localizer's dream*. Then spectral localization would immediately follow. Indeed, fix an ω from this full-measure set. Since spectrally almost all E's admit a polynomially bounded solution and each solution is either exponentially increasing or decreasing at ∞, and either exponentially increasing or decreasing at $-\infty$, it follows that all polynomially bounded solutions must decay exponentially (at the Lyapunov rate) at both $\pm\infty$ and must hence be genuine eigenfunctions. Thus, spectrally all energies are genuine eigenvalues, and spectral localization (pure point spectrum with exponentially decaying eigenfunctions for almost all ω's) follows.

Alas, things are not so easy. First, switching the order of the quantifiers only gives Lyapunov behavior at both $\pm\infty$ for almost every ω and (for example) Lebesgue almost every E, as a consequence of Fubini's theorem (applied to the product measure $\mu\times\mathrm{Leb}$). As briefly alluded to above, this is already sufficient to establish spectral localization thanks to spectral averaging if the single-site distribution has an absolutely continuous component. But it is decidedly not sufficient in the singular case. Second and more importantly, the Anderson localizer's dream is even known to fail. Namely, Gorodetski and Kleptsyn have shown in [35] that, for

almost all ω's, Lyapunov behavior fails for energies E from a dense G_δ subset of the spectrum Σ in the sense that

$$0 = \liminf_{n\to\infty} \frac{1}{n} \log \|M_n^E(\omega)\| < \limsup_{n\to\infty} \frac{1}{n} \log \|M_n^E(\omega)\| = L(E)$$

(and a similar statement on the left half line).[2] In other words, the Anderson localizer's dream is a mirage that should not be chased.

However, we will formulate the appropriate modification of the Anderson localizer's dream. To state it, let us denote the set of energies E for which the difference equation (1.1) admits a nontrivial solution u satisfying a linear upper bound,

$$|u(n)| \le C_u(1 + |n|), \tag{1.2}$$

with C_u a u-dependent constant, by $\mathcal{G}(H_\omega)$. Energies E in $\mathcal{G}(H_\omega)$ are called *generalized eigenvalues* of H_ω, and the corresponding linearly bounded solutions u are called *generalized eigenfunctions*. It is a classical theorem of Shnol [46] and Simon [48] that

$$\mathcal{G}(H_\omega) \subseteq \sigma(H_\omega)$$

and

$$\chi_{\mathbb{R}\setminus\mathcal{G}(H_\omega)}(H_\omega) = 0.$$

In particular, $\mathcal{G}(H_\omega)$ supports all spectral measures of H_ω.

Theorem 1.2 (The modified Anderson localizer's dream)**.** *For μ-almost every $\omega \in \Omega$ and every $E \in \mathcal{G}(H_\omega)$, one has*

$$\lim_{n\to\infty} \frac{1}{n} \log \|M_n^E(\omega)\| = \lim_{n\to-\infty} \frac{1}{|n|} \log \|M_n^E(\omega)\| = L(E).$$

That is, the desired Lyapunov behavior holds for all generalized eigenvalues, and this is precisely the set of energies for which one needs Lyapunov behavior to be able to deduce spectral localization. In particular, Theorem 1.2 implies Theorem 1.1.

Notice also that one gets as a natural byproduct the result that the decay rate of the eigenfunctions is given by the Lyapunov exponent, which is certainly to be expected. To summarize, in this paper, we have found the natural statement and what we believe to be the natural proof of the phenomenon of spectral localization for the 1D Anderson model.

Our approach also allows us to prove dynamical localization,

$$\sup_{t\in\mathbb{R}} |\langle \delta_n, e^{-itH_\omega}\delta_m\rangle| \lesssim e^{\epsilon|m|}e^{-\beta|n-m|};$$

see Theorem 6.4 for the precise statement and [18, 29–31] for some previous results on dynamical localization for the Anderson model. Here, ϵ is arbitrarily small and β is arbitrarily close to the best possible decay rate, which is given by the minimum of the Lyapunov exponent on the almost-sure spectrum. Thus, our spectral and dynamical localization results are both established with the correct decay rate.

Furthermore, we can also use our method to prove localization for CMV matrices with random coefficients. A CMV matrix arises in the representation of the map $f(z) \mapsto zf(z)$ in $L^2(\partial\mathbb{D}, \mathrm{d}\mu)$ relative to a suitable basis, where μ denotes a

[2] As we were completing the work on this paper, we learned that [35] also contains a new proof of spectral localization for the 1D Anderson model, which arises as a byproduct of their extension of the classical Fürstenberg theorem. Additionally, while the present paper was under review, Jitomirskaya and Zhu also wrote a proof of localization based on positive Lyapunov exponents and large deviations [37].

probability measure on the unit circle that does not admit a finite support. It is a five-diagonal semi-infinite matrix that is determined by a sequence of *Verblunsky coefficients* $\{\alpha_n\}_{n\in\mathbb{Z}_+} \subset \mathbb{D}$, which arise as the recursion coefficients of the orthogonal polynomials associated with μ, and the derived quantities $\rho_n = \left(1 - |\alpha_n|^2\right)^{1/2}$:

$$\mathcal{C} = \begin{bmatrix} \overline{\alpha_0} & \overline{\alpha_1}\rho_0 & \rho_1\rho_0 & & & & & \\ \rho_0 & -\overline{\alpha_1}\alpha_0 & -\rho_1\alpha_0 & & & & & \\ & \overline{\alpha_2}\rho_1 & -\overline{\alpha_2}\alpha_1 & \overline{\alpha_3}\rho_2 & \rho_3\rho_2 & & & \\ & \rho_2\rho_1 & -\rho_2\alpha_1 & -\overline{\alpha_3}\alpha_2 & -\rho_3\alpha_2 & & & \\ & & & \overline{\alpha_4}\rho_3 & -\overline{\alpha_4}\alpha_3 & \overline{\alpha_5}\rho_4 & \rho_5\rho_4 & \\ & & & \rho_4\rho_3 & -\rho_4\alpha_3 & -\overline{\alpha_5}\alpha_4 & -\rho_5\alpha_4 & \\ & & & & \ddots & \ddots & \ddots & \end{bmatrix}. \tag{1.3}$$

This matrix defines a unitary operator in $\ell^2(\mathbb{Z}_+)$, and the spectral measure corresponding to $\mathcal{C}$ and the vector δ_0 is given by μ.

This sets up a one-to-one correspondence between measures μ and coefficient sequences $\{\alpha_n\}_{n\in\mathbb{Z}_+}$, which has been extensively studied in recent years, mainly due to the infusion of ideas from Simon's monographs [50, 51].

Similarly, an *extended CMV matrix* is a unitary operator on $\ell^2(\mathbb{Z})$ defined by a bi-infinite sequence $\{\alpha_n\}_{n\in\mathbb{Z}} \subset \mathbb{D}$ in an analogous way:

$$\mathcal{E} = \begin{bmatrix} \ddots & \ddots & \ddots & & & & & \\ \overline{\alpha_0}\rho_{-1} & -\overline{\alpha_0}\alpha_{-1} & \overline{\alpha_1}\rho_0 & \rho_1\rho_0 & & & & \\ \rho_0\rho_{-1} & -\rho_0\alpha_{-1} & -\overline{\alpha_1}\alpha_0 & -\rho_1\alpha_0 & & & & \\ & & \overline{\alpha_2}\rho_1 & -\overline{\alpha_2}\alpha_1 & \overline{\alpha_3}\rho_2 & \rho_3\rho_2 & & \\ & & \rho_2\rho_1 & -\rho_2\alpha_1 & -\overline{\alpha_3}\alpha_2 & -\rho_3\alpha_2 & & \\ & & & & \overline{\alpha_4}\rho_3 & -\overline{\alpha_4}\alpha_3 & \overline{\alpha_5}\rho_4 & \rho_5\rho_4 \\ & & & & \rho_4\rho_3 & -\rho_4\alpha_3 & -\overline{\alpha_5}\alpha_4 & -\rho_5\alpha_4 \\ & & & & & \ddots & \ddots & \ddots \end{bmatrix}. \tag{1.4}$$

From the point of view of orthogonal polynomials, the study of $\mathcal{C}$ is more natural, but when the Verblunsky coefficients are generated by an invertible ergodic map (such as, for example, the full shift of primary interest in this paper), the study of $\mathcal{E}$ is more natural.

We are interested in the case of random Verblunsky coefficients. As in the Schrödinger case, we fix a single-site distribution that is compactly supported inside the open unit disk $\mathbb{D}$ (rather than $\mathbb{R}$ as above). This induces random sequences $\{\alpha_n(\omega)\}_{n\in\mathbb{Z}_+}$ and $\{\alpha_n(\omega)\}_{n\in\mathbb{Z}}$, as well as random CMV matrices $\mathcal{C}_\omega$ and random extended CMV matrices $\mathcal{E}_\omega$.

The following theorem is the CMV analogue of Theorem 1.1.

Theorem 1.3 (Spectral localization for random extended CMV matrices)**.** *Consider the family $\{\mathcal{E}_\omega\}_{\omega\in\Omega}$ of random extended CMV matrices, acting in $\ell^2(\mathbb{Z})$, where the Verblunsky coefficents are given by independent, identically distributed random variables. It is assumed that the topological support of the common distribution is a compact subset of $\mathbb{D}$ that contains at least two elements. Then, almost surely, $\mathcal{E}_\omega$ is spectrally localized; that is, it has pure point spectrum with exponentially decaying eigenfunctions. Moreover, the rate of decay at energy z is exactly $L(z)$.*

Due to the special interest to the orthogonal polynomial community, we also state explicitly the half-line version of the previous result.

Theorem 1.4 (Spectral localization for random CMV matrices)**.** *Consider the family $\{\mathcal{C}_\omega\}_{\omega\in\Omega}$ of random CMV matrices, acting in $\ell^2(\mathbb{Z}_+)$, where the Verblunsky coefficents are given by independent, identically distributed random variables. It is assumed that the topological support of the common distribution is a compact subset of $\mathbb{D}$ that contains at least two elements. Then, almost surely, $\mathcal{C}_\omega$ is spectrally localized; that is, it has pure point spectrum with exponentially decaying eigenfunctions.*

As above, one could ask about the case in which the single-site distribution is not compactly supported in $\mathbb{D}$, and our method should extend to these cases under the appropriate assumptions (ensuring, e.g., the existence of the Lyapunov exponent, which is absolutely fundamental to this approach). Theorem 1.4 was first discovered in situations where spectral averaging was applicable: the single-site distribution was assumed to be absolutely continuous with respect to either Lebesgue measure on $\mathbb{D}$ or arc length on a circle centered at the origin and of radius smaller than 1; compare this to [51, Theorem 12.6.3], [53]. There is some related work on another class of random unitary operators [36, 38]; their results also require the absolute continuity of the single-site distribution. For singular single-site distributions, Ahlbrecht, Scholz, and Werner worked out a version of MSA for 1D quantum walks in [1]. Then, using the connection between spatially inhomogeneous quantum walks and CMV matrices, one can use the results of [1] to deduce Theorem 1.3. Since the necessary calculations may not be known to all of our readers, we show how to make this connection between localization for quantum walks and localization for CMV matrices in the appendix.

Moreover, as in the Schrödinger case, our method also yields suitable results concerning exponential dynamical localization; see Theorem 7.3 below for the precise statement.

For the sake of completeness, we mention that our approach provides a half-line version of Theorem 1.1 as well.

1.4. **Strategy of the proof.** The basic strategy of our localization proof follows the ideas of Bourgain and Schlag's localization proof for half-line Schrödinger operators whose potentials are generated by the doubling map on the circle [10], who in turn were inspired by the paper [9] of Bourgain and Goldstein on quasi-periodic potentials. In particular, the main ingredients of this approach are uniform positivity of the Lyapunov exponent, a uniform large deviation theorem (LDT) for the same, and a version of a lemma regarding the *elimination of double resonances*. Here, uniformity is with respect to the spectral parameter E.

Let us also mention the papers [6, 7] by Binder, Goldstein, and Voda, who implemented a similar strategy for the Anderson model on the strip under the more restrictive assumption of a single-site distribution with a bounded density.

The remainder of the paper is organized as follows. In Section 2, we show continuity and uniform positivity of the Lyapunov exponent as a function of E, starting with verifying the conditions of Fürstenberg's theorem for $\mathrm{SL}(2,\mathbb{R})$, formulated as Theorem 2.1. This section is largely expository, and details are provided for the convenience of the reader.

In Section 3, starting again with the conditions of Fürstenberg's theorem, we supply a relatively simple proof of a uniform LDT for the Lyapunov exponent. Such results are known for i.i.d. random matrices [42] and one-parameter families of i.i.d. matrices [55]. However, we wish to emphasize that this proof of the LDT is new. It makes use of the independence of the underlying dynamics more transparent

and may have the potential to be useful elsewhere. In fact, by the author of [57], a version of the LDT is currently being worked out for the model in [57], where the potentials are generated by some strongly mixing dynamics. Some of the arguments in Section 3 may be directly used there.

In Section 4, we supply a simple proof of Hölder continuity of the Lyapunov exponent and the integrated density of states. This result is well known (compare this to [43, Théorème 3]), but the modern proof via the avalanche principle is simpler.

In Section 5, we obtain suitable upper bounds on the norms of the transfer matrices and relate them to estimates on the Green functions of finite-volume truncations of the full-line model.

In Section 6, we complete the proofs of Theorems 1.1, 1.2, and 6.4; the starting point is Proposition 6.1, which is an appropriate formulation of the elimination of double resonances in this setting.

We wish to emphasize that the tools we develop in Sections 5 and 6 are more general than the similar ones in [10], and this eventually enables us to show exponential dynamical localization for the Bernoulli–Anderson model.

Finally, in Section 7, following the arguments from the proofs of Theorems 1.1 and 6.4, we prove Theorems 1.3 and 1.4, and the CMV version of an exponential dynamical localization result, Theorem 7.3.

2. Positivity and continuity of the Lyapunov Exponent

In this section, we will introduce several classical results concerning positivity and continuity of the Lyapunov exponent for products of i.i.d. random $\mathrm{SL}(2,\mathbb{R})$ matrices, which will be instrumental in our proof of Anderson localization.

Consider a probability space $(\mathcal{A},\widetilde{\mu})$, and let (Ω,T,μ) be the full shift space generated by $(\mathcal{A},\widetilde{\mu})$: in other words, $\Omega=\mathcal{A}^{\mathbb{Z}}$, $\mu=\widetilde{\mu}^{\mathbb{Z}}$, and

$$(T\omega)_n=\omega_{n+1},\qquad \omega\in\Omega,\ n\in\mathbb{Z}.$$

Consider a map $M:\mathcal{A}\to\mathrm{SL}(2,\mathbb{R})$. For simplicity, we assume that this map is bounded, that is,

$$\sup_{\alpha\in\mathcal{A}}\|M(\alpha)\|<\infty.$$

This generates a map $\Omega\to\mathrm{SL}(2,\mathbb{R})$, which we also denote by M, via

$$M(\omega)=M(\omega_0),\qquad \omega\in\Omega,$$

which in turn induces an $\mathrm{SL}(2,\mathbb{R})$-cocycle over T in a canonical way:

$$(T,M):\Omega\times\mathbb{R}^2\to\Omega\times\mathbb{R}^2,\qquad (T,M)(\omega,\vec{v})=(T\omega,M(\omega)\vec{v}). \tag{2.1}$$

We define the iterates of M over the skew product by $(T,M)^n=(T^n,M_n)$ for $n\in\mathbb{Z}$. One can check that

$$M_n(\omega)=\begin{cases} M(T^{n-1}\omega)\cdots M(\omega), & n>0,\\ \mathbb{I}, & n=0,\\ \left[M_{-n}(T^n\omega)\right]^{-1}, & n<0.\end{cases}$$

We can (and do) view M_n as the product of n i.i.d. random $\mathrm{SL}(2,\mathbb{R})$ matrices with common distribution $\widetilde{\mu}$ when $n\in\mathbb{Z}_+$.

The Lyapunov exponent of the cocycle (2.1) is defined by

$$L = L(T, M) := \lim_{n\to\infty} \frac{1}{n} \int_\Omega \log \|M_n\| \, \mathrm{d}\mu = \inf_{n\geq 1} \frac{1}{n} \int_\Omega \log \|M_n\| \, \mathrm{d}\mu \geq 0.$$

By Kingman's subadditive ergodic theorem, one also has

$$L = \lim_{n\to\infty} \frac{1}{n} \log \|M_n(\omega)\|$$

for μ-almost every $\omega \in \Omega$.

In the current setting, it is convenient to consider the Lyapunov exponent as a function of the probability measure on $\mathrm{SL}(2,\mathbb{R})$. Concretely, through the probability space $(\mathcal{A}, \widetilde{\mu})$ and the map M, we obtain a probability measure on $\mathrm{SL}(2,\mathbb{R})$ via $\nu = M_*\widetilde{\mu}$, that is, the push-forward measure of $\widetilde{\mu}$ under the map M. Thus, in our setting, we also sometimes write the Lyapunov exponent as $L = L(M_*\widetilde{\mu})$.

We denote by $\mathbb{RP}^1$ the real projective line; that is, $\mathbb{RP}^1$ is the set of lines in $\mathbb{R}^2$ that pass through the origin. It clear that each $M \in \mathrm{SL}(2,\mathbb{R})$ induces a map on $\mathbb{RP}^1$, which will again be denoted by M. We say that a subgroup $G \subset \mathrm{SL}(2,\mathbb{R})$ is *strongly irreducible* if there is no finite nonempty set $\mathcal{F} \subset \mathbb{RP}^1$ such that $M(\mathcal{F}) = \mathcal{F}$ for all $M \in G$.

The following (special case of a) deep theorem of Fürstenberg is essential for our analysis.

Theorem 2.1 (Fürstenberg [27, Theorem 8.6])**.** *Let ν be a probability measure on* $\mathrm{SL}(2,\mathbb{R})$ *that satisfies*

$$\int \log \|M\| \, \mathrm{d}\nu(M) < \infty.$$

Denote by G_ν the smallest closed subgroup of $\mathrm{SL}(2,\mathbb{R})$ *that contains* $\operatorname{supp} \nu$.

Assume the following:

(i) *G_ν is not compact.*
(ii) *G_ν is strongly irreducible.*

Then $L > 0$.

Remark 2.2. Under condition (i), strong irreducibility of G_ν is equivalent to the following:

(ii′) There is no set $\mathcal{F} \subseteq \mathbb{RP}^1$ of cardinality 1 or 2 such that $M(\mathcal{F}) = \mathcal{F}$ for all $M \in G_\nu$.

From this, one can immediately deduce global positivity of the Lyapunov exponent for the Anderson model as soon as the single-site distribution has at least two points in its support. We now precisely define the Anderson model and set up our notation.

Definition 2.3. Suppose that our probability space $(\mathcal{A}, \widetilde{\mu})$ consists of a compact set of real numbers; that is, we assume henceforth that

$$\mathcal{A} = \operatorname{supp} \widetilde{\mu} \subset \mathbb{R},$$

and that $\mathcal{A}$ is compact. Then, for each $\omega \in \Omega$,

$$[H_\omega \psi](n) = \psi(n-1) + \psi(n+1) + \omega_n \psi(n), \qquad n \in \mathbb{Z},\ \psi \in \ell^2(\mathbb{Z}),$$

defines a bounded self-adjoint operator on $\ell^2(\mathbb{Z})$. For each $E \in \mathbb{R}$, we define the map $M^E : \mathcal{A} \to \mathrm{SL}(2,\mathbb{R})$ via

$$M^E(\alpha) = \begin{bmatrix} E-\alpha & -1 \\ 1 & 0 \end{bmatrix}, \qquad \alpha \in \mathcal{A},$$

which may be extended to Ω as above. Then it is straightforward to verify that $M_n^E(\omega)$ is the n-step transfer matrix of the eigenvalue equation $H_\omega \phi = E\phi$. More specifically, $H_\omega \phi = E\phi$ if and only if

$$\begin{bmatrix} \phi(n) \\ \phi(n-1) \end{bmatrix} = M_n^E(\omega) \begin{bmatrix} \phi(0) \\ \phi(-1) \end{bmatrix} \qquad \text{for all } n \in \mathbb{Z}.$$

The induced measure on $\mathrm{SL}(2,\mathbb{R})$ will be denoted by $\nu_E = M_*^E \widetilde{\mu}$, and the Lyapunov exponent at energy E is then defined and denoted by $L(E) := L(\nu_E)$.

Defining

$$\Sigma := \mathcal{A} + [-2,2] = \{x + y : x \in \mathcal{A},\ y \in [-2,2]\}, \tag{2.2}$$

one can check that $\sigma(H_\omega) = \Sigma$ for μ-a.e. $\omega \in \Omega$, e.g., by using generalized eigenfunctions.

Clearly, if $\#\mathcal{A} = 1$, then the theory is quite trivial. Concretely, if $\mathcal{A}$ consists of the single point $a \in \mathbb{R}$, then Ω contains only the constant sequence $\omega_n \equiv a$; in this case, $\sigma(H_\omega) = \Sigma = [a-2, a+2]$, and the spectral type is purely absolutely continuous. Henceforth, we adopt the standing nontriviality assumption that $\#\mathcal{A} \geq 2$.

Theorem 2.4. *In the Anderson model, $\nu_E = M_*^E \widetilde{\mu}$ satisfies assumptions* (i) *and* (ii) *from Theorem* 2.1 *for every $E \in \mathbb{R}$. In particular, we have $L(E) > 0$ for every $E \in \mathbb{R}$.*

Proof. Fix $E \in \mathbb{R}$. Since the support of the single-site distribution has cardinality at least 2, it follows that ν_E also has at least two points in its support. Thus, G_{ν_E} contains at least two distinct elements of the form

$$M_x = \begin{bmatrix} x & -1 \\ 1 & 0 \end{bmatrix},$$

say, M_a and M_b with $a \neq b$. Note that

$$A = M_a M_b^{-1} = \begin{bmatrix} 1 & a-b \\ 0 & 1 \end{bmatrix} \in G_{\nu_E}.$$

Taking powers of the matrix A, we see that G_{ν_E} is not compact.

Now consider $V_1 := \mathrm{span}(\vec{e}_1)$, the projection of $\vec{e}_1 := (1,0)^\top$ to $\mathbb{RP}^1$. Then $AV_1 = V_1$ and, for every $V \in \mathbb{RP}^1$, $A^n V$ converges to V_1. Thus, if there is a nonempty finite invariant set of directions $\mathcal{F} \subseteq \mathbb{RP}^1$, one must have $\mathcal{F} = \{V_1\}$. However, we also have

$$A' = M_a^{-1} M_b = \begin{bmatrix} 1 & 0 \\ a-b & 1 \end{bmatrix} \in G_\nu$$

and $A'V_1 \neq V_1$. Thus, G_{ν_E} is strongly irreducible, so conditions (i) and (ii) of Theorem 2.1 are met. Consequently, $L(E) = L(\nu_E) > 0$ by Fürstenberg's theorem. □

The main application of positive Lyapunov exponents is that one obtains precise asymptotic statements about orbits under cocycle iterates. The following deterministic theorem supplies what we need.

Theorem 2.5 (Ruelle [45]). *Suppose that* $A^{(n)} \in \mathrm{SL}(2,\mathbb{R})$ *obeys*

$$\lim_{n\to\infty} \frac{1}{n} \log \|A^{(n)}\| = 0$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log \|A^{(n)} \cdots A^{(1)}\| = L > 0.$$

Then there exists a 1D subspace $V \in \mathbb{RP}^1$ *such that*

$$\lim_{n\to\infty} \frac{1}{n} \log \|A^{(n)} \cdots A^{(1)} \vec{v}\| = \begin{cases} -L & \vec{v} \in V \setminus \{0\}, \\ L & \vec{v} \in \mathbb{R}^2 \setminus V. \end{cases}$$

For a proof of Theorem 2.5, see [45] or [14, Theorem 2.8].

For the Anderson model, we will also want to know that $L(E)$ is a continuous function of E, which follows easily from a theorem that goes back to Fürstenberg and Kifer in the 1980s. Given a sequence of Borel probability measures $\{\nu_k\}$ supported in $\mathrm{SL}(2,\mathbb{R})$, we say that ν_k converges to ν *weakly and boundedly* if

$$\int_{\|M\|\geq N} \log^+ \|M\| \, \mathrm{d}\nu_k(M) + \int_{\|M\|\geq N} \log^+ \|M\| \, \mathrm{d}\nu(M) \to 0 \tag{2.3}$$

as $N \to \infty$, uniformly in k and

$$\int f \, \mathrm{d}\nu_k \to \int f \, \mathrm{d}\nu$$

for all $f \in C_c(\mathrm{SL}(2,\mathbb{R}),\mathbb{C})$, the space of continuous complex-valued functions on $\mathrm{SL}(2,\mathbb{R})$ with compact support. In (2.3), we use $\log^+(x) = \max(\log x, 0)$ to denote the positive part of $\log x$. We first state a (special case of a) theorem of Fürstenberg and Kifer, which is another cornerstone for our analysis.

Theorem 2.6 (Fürstenberg and Kifer [28, Theorem B]). *Let* ν *be a probability measure on* $\mathrm{SL}(2,\mathbb{R})$ *for which*

$$\mathrm{Fix}(G_\nu) := \left\{ V \in \mathbb{RP}^1 : MV = V \text{ for every } M \in G_\nu \right\}$$

contains at most one element. Then, if $\nu_k \to \nu$ *weakly and boundedly, it holds that*

$$\lim_{k\to\infty} L(\nu_k) = L(\nu).$$

Applying this theorem to the Anderson model, we obtain continuity of L as a function of $E \in \mathbb{R}$.

Theorem 2.7. *In the Anderson model,* $L(E)$ *is continuous as a function of* E*. In particular,* L *is uniformly positive in the sense that*

$$\gamma := \inf_{E\in\mathbb{R}} L(E) > 0. \tag{2.4}$$

Proof. Let $E \in \mathbb{R}$ be given, and, as above, let $\nu_E = M_*^E \widetilde{\mu}$ denote the induced measure on $\mathrm{SL}(2,\mathbb{R})$. By Theorem 2.4, G_{ν_E} is strongly irreducible; in particular, $\mathrm{Fix}(G_{\nu_E})$ is empty, so ν_E satisfies the assumption of Theorem 2.6. Consequently, to prove continuity of $L(E)$ in E, it suffices to show that $\nu_{E_n} \to \nu_E$ weakly and boundedly whenever $E_n \to E$.

Given a sequence $E_n \to E$, one can verify that there is a uniform compact subset of $\mathrm{SL}(2,\mathbb{R})$ that simultaneously supports ν_E and every ν_{E_n}, so (2.3) follows. Thus, we only need to show the "weakly" part. By dominated convergence, we obtain

$$\lim_{n\to\infty} \int f \, \mathrm{d}\nu_{E_n} = \lim_{n\to\infty} \int f \circ M^{E_n} \, \mathrm{d}\widetilde{\mu} = \int f \circ M^{E} \, \mathrm{d}\widetilde{\mu} = \int f \, \mathrm{d}\nu_E$$

for every $f \in C(\mathrm{SL}(2,\mathbb{R}),\mathbb{C})$, which concludes the proof of continuity.

Combining the continuity of L with Theorem 2.4, we see that $L(E)$ is uniformly bounded from below away from zero on any compact set. On the other hand, one can check that

$$\lim_{|E|\to\infty} L(E) = +\infty,$$

so (2.4) follows. □

Remark 2.8. We have defined γ to be the global minimum of $L(E)$ over $E \in \mathbb{R}$. However, it turns out that

$$\gamma = \min_{E\in\Sigma} L(E); \tag{2.5}$$

that is, L achieves its minimum value on Σ. To see this, one may use the Thouless formula [5, 13, 54]:

$$L(E) = \int_{\mathbb{R}} \log|E-x| \, \mathrm{d}N(x). \tag{2.6}$$

In (2.6), $\mathrm{d}N$ denotes the *density of states* (DOS) measure associated with the family $\{H_\omega\}_{\omega\in\Omega}$, which is defined by

$$\int g(E) \, \mathrm{d}N(E) = \int_\Omega \langle \delta_0, g(H_\omega)\delta_0\rangle \, \mathrm{d}\mu(\omega)$$

for bounded measurable functions g. The DOS is a Borel probability measure with $\operatorname{supp} \mathrm{d}N = \Sigma$ [5]. Then, if $E \notin \Sigma$, L is differentiable at E with

$$\frac{\mathrm{d}L}{\mathrm{d}E}(E) = \int_{\mathbb{R}} \frac{\mathrm{d}N(x)}{E-x}.$$

This follows by using dominated convergence and noting that $E \notin \Sigma$ allows one to uniformly bound $h^{-1}(\log|E+h-x| - \log|E-x|)$ over $x \in \Sigma$ for sufficiently small h. Applying this argument again, one gets

$$\frac{\mathrm{d}^2L}{\mathrm{d}E^2}(E) = -\int_{\mathbb{R}} \frac{\mathrm{d}N(x)}{(E-x)^2} < 0.$$

Then, since L is continuous on $\mathbb{R}$ and $L(E) \to \infty$ as $|E| \to \infty$, (2.5) follows. The reader should note that the argument used to deduce differentiability of L no longer works if $E \in \Sigma = \operatorname{supp} \mathrm{d}N$, so we only get smoothness of L outside of Σ.

The last piece of information that we will need to run our arguments is a statement to the effect that, for any sequence of unit vectors $v_n \in \mathbb{R}^2$, $\|M_n^E(\omega)v_n\|$ grows like $e^{nL(E)}$ as $n \to \infty$ (μ-almost surely); this statement will enable us to prove an initial scale estimate that we can then use to inductively prove an LDT. To do this, we must verify one final technical hypothesis on the group G_{ν_E}.

Definition 2.9. Given a subset G of $\mathrm{GL}(d,\mathbb{R})$, we say that G is *contracting* if there exists a sequence $\{g_n\}_{n=1}^\infty$ in G for which $\|g_n\|^{-1}g_n$ converges to a rank-1 operator.

Proposition 2.10. *In the Anderson model, G_{ν_E} is contracting for every E.*

Proof. Given E, let a, b, and A be as in the proof of Theorem 2.4, and take $g_n = A^n$. It is easy to check that

$$\|g_n\|^{-1} g_n \to \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

which concludes the proof. □

Proposition 2.11. *For any E and any convergent sequence $\{\vec{v}_n\}_{n=1}^{\infty}$ of unit vectors in $\mathbb{R}^2$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| M_n^E(\omega)\vec{v}_n \right\| = L(E)$$

for almost every $\omega \in \Omega$.

Proof. This follows from [8, pp. 53 and 54, Corollary 3.4], which may be invoked because G_{ν_E} is strongly irreducible and contracting. □

Proposition 2.11 represents the starting point of our new proof of the LDT. We want to say a few words about this proposition since it is already a highly nontrivial result. First, Proposition 2.11 trivially holds true if the Lyapunov exponent is zero; that is, when $L = 0$. Consequently, the nontrivial part of this proposition lies in the case in which $L > 0$.

From Oseledec's multiplicative ergodic theorem, we know that if $L > 0$, the cocycle (2.1) has two invariant sections (i.e., measurable maps $\Lambda^s, \Lambda^u : \Omega \to \mathbb{RP}^1$ with $A\Lambda^\bullet = \Lambda^\bullet \circ T$ for $\bullet \in \{s, u\}$) that are called *stable* and *unstable* sections. For almost every ω, vectors drawn from the stable subspace $\Lambda^s(\omega)$ will contract exponentially fast in forward time (under the cocycle map) with the rate L, while vectors drawn from the unstable subspace $\Lambda^u(\omega)$ contract in backward time. Away from the stable direction, every vector grows exponentially in forward time with the rate L. Thus, Propsition 2.11 is a more general and sophisticated version of the following statement: for products of i.i.d. random $\mathrm{SL}(2, \mathbb{R})$ matrices obeying the conditions of Theorems 2.1 and 2.6, every nonzero vector in $\mathbb{R}^2$ is not in the stable direction of the phase ω with probability 1.

3. Large deviation estimates for products of i.i.d. matrices

For all $E \in \mathbb{R}$, Theorem 2.4 and Proposition 2.10 imply that $G = G_{\nu_E}$ is noncompact, strongly irreducible, and contracting. From Theorem 2.7, we know that L is a continuous function of E (which is a consequence of continuity of the cocycle map and the contracting property). We will use these properties to deduce a suitable uniform (in E) LDT.

Henceforth, define

$$\hat{\Sigma} := [-\kappa, \kappa], \qquad \kappa := 2 + \max_{\alpha\in\mathcal{A}} |\alpha|.$$

In particular, $\hat{\Sigma}$ is a compact interval containing the almost-sure spectrum, Σ, defined in (2.2). The goal of this section is to supply a simple proof of the following uniform LDT using strong irreducibility and contractivity of G_{ν_E}.

Theorem 3.1. *For any $\varepsilon > 0$, there exist $C = C(\varepsilon) > 0$, $\eta = \eta(\varepsilon) > 0$ such that*

$$\mu\left\{\omega \in \Omega : \left| \frac{1}{n} \log \left\| M_n^E(\omega) \right\| - L(E) \right| \geq \varepsilon \right\} \leq Ce^{-\eta n}$$

for all $n \in \mathbb{Z}_+$ and all $E \in \hat{\Sigma}$.

We begin by proving a first step toward this estimate. Let us denote the set of all unit vectors in $\mathbb{R}^2$ by

$$\mathbb{S}^1 = \left\{\vec{v} \in \mathbb{R}^2 : |v_1|^2 + |v_2|^2 = 1\right\}.$$

Proposition 3.2. *For any $\varepsilon > 0$, $0 < \delta < 1$, and $E \in \hat{\Sigma}$, there exists an $N = N(\varepsilon, \delta, E)$ such that*

$$\mu\left\{\omega : \left|\frac{1}{n}\log\left\|M_n^E(\omega)\vec{v}\right\| - L(E)\right| \geq \varepsilon\right\} < \delta \tag{3.1}$$

for all $n \geq N$ and $\vec{v} \in \mathbb{S}^1$.

Proof. Suppose, for the sake of establishing a contradiction, that the statement of the proposition is false; that is, suppose that there exist $E_0 \in \hat{\Sigma}$, $\delta_0 \in (0,1)$, $\varepsilon_0 > 0$, a sequence of integers $n_k \to \infty$, and unit vectors $\vec{v}_k \in \mathbb{S}^1$ such that

$$\mu\left\{\omega : \left|\frac{1}{n_k}\log\left\|M_{n_k}^{E_0}(\omega)\vec{v}_k\right\| - L(E_0)\right| \geq \varepsilon_0\right\} \geq \delta_0 \qquad \text{for every } k \in \mathbb{Z}_+. \tag{3.2}$$

Since E_0 plays no role in the argument, we suppress it from the notation for the remainder of the proof. By passing to a subsequence, we may assume that the vectors $\vec{v}_k$ converge to a vector $\vec{v}_\infty \in \mathbb{S}^1$. Consequently, since $n_k^{-1}\log\|M_{n_k}\vec{v}_k\| - L$ is uniformly bounded on Ω, we may apply Proposition 2.11 and dominated convergence to get

$$\lim_{k\to\infty}\int_\Omega \left|\frac{1}{n_k}\log\|M_{n_k}(\omega)\vec{v}_k\| - L\right| \mathrm{d}\mu(\omega) = 0. \tag{3.3}$$

On the other hand, letting A_k denote the set on the left-hand side of (3.2), we get

$$\int_\Omega \left|\frac{1}{n_k}\log\|M_{n_k}\vec{v}_k\| - L\right| \mathrm{d}\mu \geq \int_{A_k} \left|\frac{1}{n_k}\log\|M_{n_k}\vec{v}_k\| - L\right| \mathrm{d}\mu \geq \varepsilon_0\delta_0$$

for every $k \in \mathbb{Z}_+$, which contradicts (3.3). □

Our next goal is to extend this proposition to a neighborhood of E, so we need to estimate how changes in E perturb M_n. To do this, we define

$$F_n(\omega, E) = \frac{1}{|n|}\log\|M_n^E(\omega)\|, \qquad n \in \mathbb{Z},\ \omega \in \Omega,\ E \in \mathbb{R}, \tag{3.4}$$

where we adopt the convention $F_0 \equiv 1$. Throughout the paper, there will be various uniform bounds that can be controlled in terms of bounds on the single-step matrices. Hence, we introduce

$$\Gamma := \sup\left\{\left\|M^E(\alpha)\right\| : E \in \hat{\Sigma},\ \alpha \in \mathcal{A}\right\}. \tag{3.5}$$

Since Ω and $\hat{\Sigma}$ are compact and $M^E(\omega)$ is continuous as a function of $(E,\omega) \in \hat{\Sigma}\times\Omega$, it follows that Γ is finite. From the definitions, it is easy to see that

$$|F_n(\omega, E)| \leq \log\Gamma \qquad \text{for all } n \in \mathbb{Z},\ \omega \in \Omega,\ E \in \hat{\Sigma}.$$

Combining this with (2.4), we get the uniform bounds

$$\gamma \leq L(E) \leq \log\Gamma \qquad \text{for all } E \in \hat{\Sigma}. \tag{3.6}$$

The following lemma is a straightforward calculation.

Lemma 3.3. *We have*

$$\left\|M_n^E(\omega) - M_n^{E'}(\omega')\right\| \leq n\Gamma^{n-1}\left(|E-E'| + \max_{0\leq j<n}|\omega_j - \omega_j'|\right) \tag{3.7}$$

for all $E, E' \in \Sigma$, $\omega, \omega' \in \Omega$, *and* $n \in \mathbb{Z}_+$. *Therefore,*

$$|F_n(\omega, E) - F_n(\omega', E')| \leq \Gamma^{n-1}\left(|E-E'| + \max_{0\leq j<n}|\omega_j - \omega_j'|\right) \tag{3.8}$$

for all E, E', ω, ω', *and* n. *Furthermore, with*

$$L_n(E) := \int_\Omega F_n(\omega, E)\,\mathrm{d}\mu(\omega),$$

one has

$$|L_n(E) - L_n(E')| \leq \Gamma^{n-1}|E-E'|. \tag{3.9}$$

Proposition 3.4. *For any* $\varepsilon > 0$, $0 < \delta < 1$, *and* $E \in \hat{\Sigma}$, *there exists an* $N' = N'(\varepsilon, \delta, E)$ *such that, for every integer* $n \geq N'$, *there exists a* $\rho = \rho(n) > 0$ *with the property that*

$$\mu\left\{\omega : \left|\frac{1}{n}\log\left\|M_n^{E'}(\omega)\vec{v}\right\| - L(E')\right| \geq \varepsilon\right\} < \delta, \tag{3.10}$$

for all unit vectors $\vec{v} \in \mathbb{S}^1$, *whenever* $|E-E'| < \rho$ *and* $E' \in \hat{\Sigma}$.

Proof. Fix $\varepsilon > 0$, $0 < \delta < 1$, and $E \in \hat{\Sigma}$, and put $N' = N(\varepsilon/2, \delta, E)$ from Proposition 3.2. Let $n \geq N$ be given, and let $\rho = \rho(n) > 0$ be chosen such that

$$|L(E) - L(E')| < \varepsilon/4 \qquad \text{whenever } E' \in \hat{\Sigma} \text{ and } |E-E'| < \rho,$$

which can be done by continuity of the Lyapunov exponent and compactness of $\hat{\Sigma}$. If necessary, shrink ρ to ensure that

$$\rho \leq \frac{\varepsilon}{4\Gamma^{n-1}},$$

where Γ is as in (3.5). For any $\vec{v} \in \mathbb{S}^1$, we have

$$\left\{\omega : \left|\frac{1}{n}\log\|M_n^{E'}(\omega)\vec{v}\| - L(E')\right| \geq \varepsilon\right\} \subset \left\{\omega : \left|\frac{1}{n}\log\|M_n^{E}(\omega)\vec{v}\| - L(E)\right| \geq \varepsilon/2\right\} \tag{3.11}$$

whenever $E' \in \hat{\Sigma}$ and $|E-E'| < \rho$. Concretely, if ω lies in the complement of the right-hand side of (3.11), we have

$$\begin{aligned}\left|\frac{1}{n}\log\|M_n^{E'}(\omega)\vec{v}\| - L(E')\right| &\leq \Gamma^{n-1}|E-E'| + \left|\frac{1}{n}\log\|M_n^{E}(\omega)\vec{v}\| - L(E')\right| \\ &\leq \frac{\varepsilon}{4} + \left|\frac{1}{n}\log\|M_n^{E}(\omega)\vec{v}\| - L(E)\right| + |L(E) - L(E')| \\ &< \varepsilon,\end{aligned}$$

where the first inequality uses Lemma 3.3. Thus, (3.11) holds and the conclusion of the proposition follows from Proposition 3.2 and our choice of N. □

In what follows, we will need to use the following discrete Chebyshev-type inequality.

Lemma 3.5. *Let $P \in \mathbb{Z}_+$ and $a_1, \dots, a_P \in \mathbb{R}$ be given. Given $L < \max_j a_j \le B$ and $\delta > 0$, define*

$$\mathcal{J}_{L,\delta} = \{j : a_j > L + \delta\}.$$

If

$$\frac{1}{P}\sum_{j=1}^{P} a_j \ge L + \varepsilon \tag{3.12}$$

for some $\varepsilon > 0$, then

$$\#\mathcal{J}_{L,\delta} \ge P\frac{\varepsilon - \delta}{B - L - \delta} \tag{3.13}$$

for every $0 < \delta < B - L$.

Proof. Assume that (3.12) holds, and let $m := \#\mathcal{J}_{L,\delta}$. Using (3.12) and splitting the sum over $\mathcal{J}_{L,\delta}$ and its complement, one obtains

$$P(L+\varepsilon) \le \sum_{j=1}^{P} a_j = \sum_{j\in\mathcal{J}_{L,\delta}} a_j + \sum_{j\notin\mathcal{J}_{L,\delta}} a_j \le mB + (P-m)(L+\delta).$$

Solving for m, we get (3.13). □

We may now combine our foregoing work to fashion the final stepping stone before proving the main LDT: a vectorwise uniform LDT. Compare this to [55, Theorem 4].

Proposition 3.6. *For every $\varepsilon > 0$, there exist constants $C, \eta > 0$ such that*

$$\mu\left\{\omega : \left|\frac{1}{n}\log\left\|M_n^E(\omega)\vec{v}\right\| - L(E)\right| \ge \varepsilon\right\} \le Ce^{-\eta n}$$

for every $n \in \mathbb{Z}_+$, $\vec{v} \in \mathbb{S}^1$, and $E \in \hat{\Sigma}$.

Proof. Let Γ be defined as in (3.5), put $B = \log\Gamma + 1$, fix $\varepsilon \in (0,1)$, and notice that

$$\varepsilon < B - L(E) \qquad \text{for every } E \in \hat{\Sigma}$$

by (3.6). Motivated by Lemma 3.5, we define $\xi > 0$ by

$$\xi := \frac{\varepsilon/4}{B - \gamma - \frac{\varepsilon}{4}},$$

and we observe that

$$\xi \le \frac{\varepsilon/4}{B - L(E) - \frac{\varepsilon}{4}} \qquad \text{for every } E \in \hat{\Sigma} \tag{3.14}$$

by (3.6). In particular, ξ depends on ε, but not on E. Now fix $\delta > 0$ small enough that $1 + \delta e \le e^{\xi/2}$, and let $E \in \hat{\Sigma}$ be given. Then, put $N = N'(\varepsilon/4, \delta; E)$ and $\rho = \rho(N)$ as in Proposition 3.4, and let $E' \in (E-\rho, E+\rho)$ and $n \in \mathbb{Z}_+$ be given. Writing $n = NP + r$ with $P \in \mathbb{Z}_+$ and $0 \le r < N$, one can check that

$$\log\|M_n^{E'}(\omega)\vec{v}\| = \sum_{p=0}^{P-1}\log\left\|M_N^{E'}(T^{pN}\omega)\vec{v}_p\right\| + \log\left\|M_r^{E'}(T^{PN}\omega)\vec{v}_P\right\|,$$

where

$$\vec{v}_p = \vec{v}_p(\omega, E') := \frac{M_{pN}^{E'}(\omega)\vec{v}}{\|M_{pN}^{E'}(\omega)\vec{v}\|}, \qquad 0 \le p \le P.$$

We first deal with the set

$$\mathcal{B}_n^+ = \mathcal{B}_n^+(E', \varepsilon, \vec{v}) := \left\{ \omega : \frac{1}{n} \log \left\| M_n^{E'}(\omega)\vec{v} \right\| - L(E') \geq \varepsilon \right\}.$$

For n large enough, one has

$$\mathcal{B}_n^+ \subseteq \left\{ \omega : \frac{1}{P} \sum_{p=0}^{P-1} \frac{1}{N} \log \|M_N^{E'}(T^{pN}\omega)\vec{v_p}(\omega, E')\| \geq L(E') + \frac{\varepsilon}{2} \right\}.$$

In light of this, Lemma 3.5, and our choice of ξ in (3.14), we obtain

$$\mathcal{B}_n^+ \subseteq \bigcup_{\substack{\mathcal{J} \subset [0, P-1] \cap \mathbb{Z} \\ \#\mathcal{J} \geq \xi P}} \bigcap_{p \in \mathcal{J}} A_p, \tag{3.15}$$

where $A_p = A_p(E', \varepsilon, \vec{v})$ is given by

$$A_p := \left\{ \omega : \frac{1}{N} \log \|M_N^{E'}(T^{pN}\omega)\vec{v_p}(\omega, E')\| \geq L(E') + \frac{\varepsilon}{4} \right\}.$$

Thus, it remains to bound the measure of sets of the form

$$A_{\mathcal{J}} = \bigcap_{p \in \mathcal{J}} A_p,$$

with $\mathcal{J} \subset [0, P-1] \cap \mathbb{Z}$ being a set of cardinality at least ξP. To that end, we notice that whether or not $\omega \in A_p$ depends only on the coordinates $(\omega_0, \omega_1, \ldots, \omega_{(p+1)N-1})$. To capture this dependence, we introduce the following grouping of coordinates. Suppose that $\#\mathcal{J} = m \geq \xi P$, write $\mathcal{J} = \{p_1 < p_2 < \cdots < p_m\}$, and define

$$\vec{\omega}_j := \left(\omega_{(p_{j-1}+1)N} \ldots \omega_{(p_j+1)N-1}\right), \qquad 1 \leq j \leq m,$$

where we take $p_0 = -1$ by convention. Thus, we obtain the following grouping of the coordinates of ω:

$$\omega = (\ldots, \underbrace{\omega_0, \ldots, \omega_{(p_1+1)N-1}}_{\vec{\omega}_1 \in \Omega_1 := \mathcal{A}^{(p_1+1)N}}, \ldots, \underbrace{\omega_{(p_{m-1}+1)N}, \ldots, \omega_{(p_m+1)N-1}}_{\vec{\omega}_m \in \Omega_m := \mathcal{A}^{(p_m - p_{m-1})N}}, \ldots).$$

Denoting the $(p_j - p_{j-1})N$-fold product of $\widetilde{\mu}$ with itself on Ω_j by $\widetilde{\mu}_j$, (3.15) gives us

$$\begin{aligned} \mu\left(A_{\mathcal{J}}\right) &= \int_{\Omega} \prod_{p \in \mathcal{J}} \chi_{A_p}(\omega)\, \mathrm{d}\mu(\omega) \\ &= \int_{\Omega_1} \cdots \int_{\Omega_m} \left[\prod_{j=1}^{m} \chi_{A_{p_j}}(\vec{\omega}_1, \ldots, \vec{\omega}_j) \right] \mathrm{d}\widetilde{\mu}_m(\vec{\omega}_m) \cdots \mathrm{d}\mu_1(\vec{\omega}_1). \end{aligned} \tag{3.16}$$

In the innermost integral, the "important variables" that govern the growth of $\|M_N \vec{v}_{p_m}\|$ correspond to the last N coordinates of $\vec{\omega}_m$. So we further split $\vec{\omega}_m$ into its last N entries and first $(p_m - p_{m-1} - 1)N$ entries. Write $\vec{\sigma} = (\sigma_1, \ldots, \sigma_N)$ for the terminal N coordinates of $\vec{\omega}_m$, and let $\vec{\omega}_m^1$ denote the remaining initial

$(p_m - p_{m-1} - 1)N$ coordinates of $\vec{\omega}_m$. Then our goal is to estimate

$$\begin{aligned}(3.17)\qquad &\int_{\Omega_m} \chi_{A_{p_m}}(\vec{\omega}_1,\dots,\vec{\omega}_m)\,\mathrm{d}\widetilde{\mu}_m(\vec{\omega}_m)\\ &\quad= \int_{\Omega_m^1}\int_{\mathcal{A}^N} \chi_{A_{p_m}}(\vec{\omega}_1,\dots,\vec{\omega}_m^1,\vec{\sigma})\,\mathrm{d}\widetilde{\mu}^N(\vec{\sigma})\,\mathrm{d}\widetilde{\mu}^{(p_m-p_{m-1}-1)N}(\vec{\omega}_m^1)\end{aligned}$$

uniformly over $\vec{\omega}_1,\dots,\vec{\omega}_{m-1}$. Note that, in the definition of A_{p_m}, $M_N(T^{p_m N}\omega)$ depends only on $\vec{\sigma}$, and $\vec{v}_{p_m}$ depends only on $(\vec{\omega}_1,\vec{\omega}_2,\dots,\vec{\omega}_m^1)$. Consequently,

$$\int_{\mathcal{A}^N} \chi_{A_{p_m}}(\vec{\omega}_1,\vec{\omega}_2,\dots,\vec{\omega}_m^1,\vec{\sigma})\,\mathrm{d}\widetilde{\mu}^N(\vec{\sigma}) = \widetilde{\mu}^N\left\{\vec{\sigma} : \frac{1}{N}\log\|M_N^{E'}(\vec{\sigma})\vec{v}_{p_m}\| \ge L + \frac{\varepsilon}{4}\right\} < \delta$$

for our choice of N by Proposition 3.4. Now plugging back into (3.17) gives us

$$\int_{\Omega_m} \chi_{A_{p_m}}(\vec{\omega}_1,\vec{\omega}_2,\dots,\vec{\omega}_m)\,\mathrm{d}\widetilde{\mu}_m(\vec{\omega}_m) < \delta.$$

Inductively applying the same argument m times, we get

$$(3.18)\qquad \mu(A_{\mathcal{J}}) \le \delta^m \qquad \text{whenever } \mathcal{J}\subset[0,P-1]\cap\mathbb{Z} \text{ and } \#\mathcal{J} = m.$$

Bounding the measure of $\mathcal{B}_n^+$ is now a matter of counting and our choice of δ. Namely, if we write $I = [0,P-1]\cap\mathbb{Z}$, (3.15) and (3.18) imply that

$$\mu(\mathcal{B}_n^+) \le \sum_{\substack{\mathcal{J}\subseteq I\\ \#\mathcal{J}\ge\xi P}} \delta^{\#\mathcal{J}} \le e^{-\xi P}\sum_{\mathcal{J}\subseteq I}(\delta e)^{\#\mathcal{J}} = e^{-\xi P}(1+\delta e)^P \le e^{-\xi P/2}$$

by our choice of δ. Taking

$$\eta_0 = \eta_0(E,\varepsilon) := \frac{\xi}{3N},$$

we find that

$$\mu(\mathcal{B}_n^+) \le e^{-\xi P/2} \le e^{-\eta_0 n}$$

for all sufficiently large n (note that the largeness condition depends solely on E and ε). One bounds the μ-measure of

$$\mathcal{B}_n^- = \mathcal{B}_n^-(E',\varepsilon,\vec{v}) := \left\{\omega : \frac{1}{n}\log\|M_n^{E'}(\omega)\vec{v}\| - L(E') \le -\varepsilon\right\}$$

similarly and obtains

$$\mu(\mathcal{B}_n^-) \le e^{-\eta_0 n}$$

for all sufficiently large n by following the argument used to estimate $\mu(\mathcal{B}_n^+)$. Naturally, this yields

$$(3.19)\qquad \mu\left\{\omega : \left|\frac{1}{n}\log\left\|M_n^{E'}(\omega)\vec{v}\right\| - L(E')\right| \ge \varepsilon\right\} = \mu(\mathcal{B}_n^+\cup\mathcal{B}_n^-) \le 2e^{-\eta_0 n}$$

for n sufficiently large.

Thus, for each E in $\hat{\Sigma}$, we find $\eta_0 = \eta_0(E,\varepsilon)$, $\rho = \rho(E,\varepsilon)$, and $N'' = N''(E,\varepsilon)$, so (3.19) holds for all $E'\in(E-\rho,E+\rho)$, $n\ge N''$, and $\vec{v}\in\mathbb{S}^1$. Then we obtain the conclusion of the theorem via a straightforward compactness argument (using compactness of $\hat{\Sigma}$). □

Proof of Theorem 3.1. Let $E \in \hat{\Sigma}$ and $\varepsilon > 0$ be given. Notice that

$$\left\{\omega : \left|\frac{1}{n}\log\|M_n^E(\omega)\| - L\right| \geq \varepsilon\right\} = \mathcal{B}_n^+ \cup \mathcal{B}_n^-, \tag{3.20}$$

where $\mathcal{B}_n^{\pm} = \mathcal{B}_n^{\pm}(E,\varepsilon)$ are given by

$$\mathcal{B}_n^{\pm} := \left\{\omega : \pm\left(\frac{1}{n}\log\|M_n^E(\omega)\| - L\right) \geq \varepsilon\right\}.$$

In terms of the sets from Proposition 3.6, one has

$$\mathcal{B}_n^+(E,\varepsilon) = \bigcup_{\vec{v}\in\mathbb{S}^1} \mathcal{B}_n^+(E,\varepsilon,\vec{v}), \qquad \mathcal{B}_n^-(E,\varepsilon) = \bigcap_{\vec{v}\in\mathbb{S}^1} \mathcal{B}_n^-(E,\varepsilon,\vec{v}).$$

Thus, denoting the standard basis in $\mathbb{R}^2$ by $\{\vec{e}_1,\vec{e}_2\}$, we get $\mathcal{B}_n^-(E,\varepsilon) \subseteq \mathcal{B}_n^-(E,\varepsilon,\vec{e}_1)$, and hence

$$\mu\left(\mathcal{B}_n^-(E,\varepsilon)\right) \leq \mu\left(\mathcal{B}_n^-(E,\varepsilon,\vec{e}_1)\right) \leq Ce^{-\eta_1 n}$$

by Proposition 3.6.

It remains to estimate the measure of $\mathcal{B}_n^+(E,\varepsilon)$. Since

$$\|M_n(\omega)\| \leq \sqrt{2}\max_{j=1,2}\|M_n(\omega)\vec{e}_j\|,$$

we get

$$\mathcal{B}_n^+(E,\varepsilon) \subseteq \mathcal{B}_n^+(E,\varepsilon/2,\vec{e}_1) \cup \mathcal{B}_n^+(E,\varepsilon/2,\vec{e}_2)$$

for all n large enough that $n\varepsilon \geq \log 2$, which in turn gives

$$\mu(\mathcal{B}_n^+(E,\varepsilon)) \leq Ce^{-\eta_2 n} + Ce^{-\eta_3 n}.$$

Choosing $\eta = \min\{\eta_1,\eta_2,\eta_3\}$, the theorem follows. □

4. Hölder continuity of the Lyapunov Exponent

The purpose of this section is to supply a simple proof of the Hölder continuity of the Lyapunov exponent of the Anderson model. In essence, once we have a uniform lower bound on the Lyapunov exponent as in (2.4) and a uniform LDT as in Theorem 3.1, Hölder continuity of L follows from the approach developed by Goldstein and Schlag [34].

We will concentrate on the regularity of L on the interval $\hat{\Sigma}$. By Remark 2.8, L is already a smooth function of E away from $\hat{\Sigma}$.

Theorem 4.1. *There exist constants $C > 0, \beta > 0$ depending solely on $\widetilde{\mu}$ such that*

$$|L(E) - L(E')| \leq C|E - E'|^{\beta} \tag{4.1}$$

for all $E, E' \in \hat{\Sigma}$.

Remark 4.2. Another important object in the spectral analysis of the Anderson model is the accumulation function of the DOS, called the *integrated density of states* (IDS):

$$N(E) = \int_{(-\infty,E]} \mathrm{d}N.$$

By the Thouless formula, the IDS is (almost) the Hilbert transform of L; consequently, one can deduce quantitative continuity estimates for the IDS from such estimates on L. In particular, one can deduce Hölder continuity of N as a function of E from Theorem 4.1 with the same choice of β.

Corollary 4.3 ([43, Thèorème 3]). *There exists a constant $C > 0$ such that*

$$|N(E) - N(E')| \leq C|E - E'|^{\beta}$$

for all $E, E' \in \mathbb{R}$, where β is as in Theorem 4.1.

Proof. This follows from the uniform positivity of L, Theorem 4.1, and standard arguments using the Thouless formula, (2.6). See, e.g., [34, proof of Theorem 6.1, particularly the argument on pp. 175 and 176]. □

The key ingredient in the approach of Goldstein and Schlag is the *avalanche principle* [34]. Basically, the avalanche principle permits us good control on the norm of a product of $\mathrm{SL}(2, \mathbb{R})$ matrices provided that we have suitable estimates on consecutive pairwise products. A bit more precisely, if we consider a product like

$$A = \prod_{j=n}^{1} A^{(j)} = A^{(n)} A^{(n-1)} \cdots A^{(1)},$$

then, if $\|A^{(j+1)}A^{(j)}\|$ is not too small compared to $\|A^{(j+1)}\| \cdot \|A^{(j)}\|$, the most contracted direction of $A^{(j+1)}$ is not too close to the most contracted direction of $(A^{(j)})^{-1}$. If this holds for each $1 \leq j < n$, then one has good norm control for the product A.

The avalanche principle allows one to move from small scales of matrix products to large scales in an inductive fashion. In particular, in Lemma 4.5, we are able to relate finite-step LEs at different scales using the positivity of L and the LDT.

Lemma 4.4 (Avalanche principle). *Let $A^{(1)}, \ldots, A^{(n)}$ be a finite sequence in $\mathrm{SL}(2, \mathbb{R})$ satisfying the following conditions:*

$$\min_{1 \leq j \leq n} \|A^{(j)}\| \geq \lambda > n, \tag{4.2}$$

$$\max_{1 \leq j < n} \left| \log \|A^{(j+1)}\| + \log \|A^{(j)}\| - \log \|A^{(j+1)}A^{(j)}\| \right| < \frac{1}{2} \log \lambda. \tag{4.3}$$

Then

$$\left| \log \|A^{(n)} \ldots A^{(1)}\| + \sum_{j=2}^{n-1} \log \|A^{(j)}\| - \sum_{j=1}^{n-1} \log \|A^{(j+1)}A^{(j)}\| \right| \leq C \frac{n}{\lambda}. \tag{4.4}$$

See [34, Proposition 2.2] for a proof of Lemma 4.4.

Lemma 4.5. *There are constants $c, C > 0$ that depend only on $\widetilde{\mu}$ with the property that*

$$|L(E) + L_n(E) - 2L_{2n}(E)| \leq Ce^{-cn} \tag{4.5}$$

for all $n \in \mathbb{Z}_+$ and every $E \in \hat{\Sigma}$.

Proof. Let $E \in \hat{\Sigma}$ be given; all estimates in the present argument will be uniform over $E \in \hat{\Sigma}$, so we will suppress it from the notation, writing L, L_n, M_n in place of $L(E), L_n(E), M_n^E$. Throughout the argument, we let C denote a large $\hat{\Sigma}$-dependent constant. It is straightforward to verify that the value of C increases only finitely many times as the argument progresses, and that it can indeed be chosen uniformly over $E \in \hat{\Sigma}$. Now choose $\varepsilon > 0$ small enough that

$$0 < \frac{4\varepsilon}{\gamma - \varepsilon} < \frac{1}{2}, \tag{4.6}$$

where γ is the constant from (2.4), and pick $\eta > 0$ small enough that $\eta < 4(\gamma - \varepsilon)$ and that the conclusion of Theorem 3.1 holds for this choice of η and ε. Given $n \in \mathbb{Z}_+$ large, choose $\ell \in \mathbb{Z}_+$ such that

$$e^{\frac{\eta}{5}\ell} \leq n \leq e^{\frac{\eta}{4}\ell} < e^{(L-\varepsilon)\ell}, \tag{4.7}$$

where the final inequality follows from our choice of η. For each $\omega \in \Omega$, consider

$$A^{(j)}(\omega) := M_\ell\left(T^{(j-1)\ell}\omega\right), \qquad 1 \leq j \leq n.$$

From Theorem 3.1, by choosing n, and hence ℓ, large, we obtain that there is an exceptional set $\mathcal{B} = \mathcal{B}(n)$ with

$$\mu(\mathcal{B}) \leq e^{-\frac{1}{4}\eta\ell}$$

such that

$$\left|\frac{1}{\ell}\log\|A^{(j)}(\omega)\| - L\right|, \qquad \left|\frac{1}{2\ell}\log\|A^{(j+1)}(\omega)A^{(j)}(\omega)\| - L\right| < \varepsilon \tag{4.8}$$

whenever $\omega \notin \mathcal{B}$ and $1 \leq j \leq n$. Consequently, we have

$$\|A^{(j)}(\omega)\| \geq e^{(L-\varepsilon)\ell} > n$$

for all $\omega \in \Omega \setminus \mathcal{B}$ and $1 \leq j \leq n$. Moreover, for $1 \leq j < n$, we get

$$\left|\log\|A^{(j+1)}(\omega)\| + \log\|A^{(j)}(\omega)\| - \log\|A^{(j+1)}(\omega)A^{(j)}(\omega)\|\right| < 4\varepsilon\ell$$

from (4.8). Thus, for $\omega \in \Omega\setminus\mathcal{B}$, conditions (4.2) and (4.3) of Lemma 4.4 are fulfilled upon taking $\lambda = \exp((L-\varepsilon)\ell)$, where we have used (4.6) to verify that (4.3) holds true. Consequently, we obtain the conclusion of (4.4) for each $\omega \notin \mathcal{B}$, which reads

$$\left|\log\|M_{\ell n}(\omega)\| + \sum_{j=2}^{n-1}\log\|M_\ell(T^{(j-1)\ell}\omega)\| - \sum_{j=1}^{n-1}\log\|M_{2\ell}(T^{(j-1)\ell}\omega)\|\right| \leq C\frac{n}{\lambda}.$$

Dividing the inequality above by ℓn, integrating it against μ, and splitting the domain of integration into $\Omega \setminus \mathcal{B}$ and $\mathcal{B}$, we get

$$\left|L_{\ell n} + \frac{n-2}{n}L_\ell - \frac{2(n-1)}{n}L_{2\ell}\right| \leq C\left(\lambda^{-1}\ell^{-1} + \mu(\mathcal{B})\right) \leq Ce^{-\frac{1}{4}\eta\ell}.$$

Thus, we obtain

$$|L_{\ell n} + L_\ell - 2L_{2\ell}| \leq \frac{C}{n} + Ce^{-\frac{1}{4}\eta\ell} \leq Ce^{-\frac{1}{5}\eta\ell}. \tag{4.9}$$

Notice that (4.9) holds for any sufficiently large n and ℓ that are related via (4.7). Thus, we now can build up estimates on large scales inductively. To make this precise, let us define a relation $\gg$ by declaring

$$y \gg x \iff e^{\frac{\eta}{5}x} \leq \frac{y}{x} \leq \frac{1}{2}e^{\frac{\eta}{4}x}.$$

For the first scale, we choose $n_1 \in \mathbb{Z}_+$ large and $n_2 \gg n_1$ a multiple of n_1; applying (4.9) with $\ell = n_1$ and $n = n_2/n_1$ yields

$$|L_{n_2} + L_{n_1} - 2L_{2n_1}| \leq Ce^{-\frac{1}{5}\eta n_1} \tag{4.10}$$

and

$$|L_{2n_2} + L_{n_1} - 2L_{2n_1}| \leq Ce^{-\frac{1}{5}\eta n_1}. \tag{4.11}$$

Combining (4.10) and (4.11) gives us

$$|L_{2n_2} - L_{n_2}| \le Ce^{-\frac{1}{5}\eta n_1}. \tag{4.12}$$

Inductively, choosing $n_{s+1} \gg n_s$ such that $n_s | n_{s+1}$, we get (4.10) and (4.12) with the pair (n_1, n_2) replaced by (n_s, n_{s+1}), which in turn yields

$$|L_{n_{s+1}} - L_{n_s}| \le Ce^{-\frac{1}{5}\eta n_{s-1}} \qquad \text{for every } s \ge 2.$$

Putting these estimates together, we obtain

$$\begin{aligned} 0 \le L_{n_2} - L &\le \sum_{s=2}^{\infty} |L_{n_{s+1}} - L_{n_s}| \\ &\le \sum_{s=1}^{\infty} Ce^{-\frac{1}{5}\eta n_s} \\ &\le Ce^{-\frac{1}{5}\eta n_1}. \end{aligned}$$

Consequently, we obtain (4.5) upon replacing L_{n_2} with L in (4.10). □

Now we are ready to prove Theorem 4.1.

Proof of Theorem 4.1. Let $E, E' \in \hat{\Sigma}$ be given. Recall that

$$|L_n(E) - L_n(E')| \le \Gamma^n |E - E'|$$

by (3.9). Combining this with (4.5), we get

$$|L(E) - L(E')| \le C^n |E - E'| + Ce^{-cn} \tag{4.13}$$

for all large n, where $C, c > 0$ are suitable constants. Hölder continuity of L then follows by choosing n well. More precisely, with

$$n = \left\lfloor \frac{1}{3 \log C} \log \frac{1}{|E - E'|} \right\rfloor, \qquad \beta = \min\left\{ \frac{2}{3}, \frac{c}{3 \log C} \right\},$$

(4.13) yields

$$|L(E) - L(E')| \le C|E - E'|^{\beta},$$

which proves Theorem 4.1. □

5. Estimating transfer matrices and Green functions

In the present section, we work out the next main thrust of the localization proof—namely, suitable upper bounds on Green functions of finite-volume truncations of H_ω. In the 1D setting, the truncated Green functions are intimately connected with the transfer matrices; hence, the main technical result of the section is actually an estimate on the transfer matrices. We use the LDT to prove bounds on transfer matrices on blocks of length n on a full-measure subset of Ω, at the price of averaging over n^2 consecutive blocks. In fact, we prove a "centered" version of this result that allows us to shift the result to the center of localization (once we know that the eigenfunctions are localized, that is). We then parlay this result into an upper bound on the Green functions. Recall the function F_n defined in (3.4), that is,

$$F_n(\omega, E) = \frac{1}{|n|} \log \|M_n^E(\omega)\|.$$

Lemma 5.1. *For any $\varepsilon > 0$, there exists an $n_0 = n_0(\varepsilon)$ such that*

$$\mu\left\{\omega : \left|L(E) - \frac{1}{r}\sum_{s=0}^{r-1} F_n(T^{sn+\ell}\omega, E)\right| \geq \varepsilon\right\} \leq e^{-\frac{\varepsilon r}{2}} \tag{5.1}$$

for every $\ell \in \mathbb{Z}$, $E \in \hat{\Sigma}$, $r \in \mathbb{Z}_+$, and every $n \geq n_0$.

Proof. Since μ is T-invariant, we need to deal only with $\ell = 0$. Fix $E \in \hat{\Sigma}$, fix $\varepsilon > 0$, and suppress E from the notation. By the triangle inequality and Chebyshev's inequality, we have

$$\begin{aligned}\mu\left\{\omega : \left|L - \frac{1}{r}\sum_{s=0}^{r-1} F_n(T^{sn}\omega)\right| \geq \varepsilon\right\} &\leq \mu\left\{\omega : \frac{1}{r}\sum_{s=0}^{r-1} |L - F_n(T^{sn}\omega)| \geq \varepsilon\right\} \\ &\leq e^{-r\varepsilon}\int_\Omega \exp\left(\sum_{s=0}^{r-1} |L - F_n \circ T^{sn}|\right) \mathrm{d}\mu.\end{aligned}$$

Since $F_n \circ T^{sn}$ and $F_n \circ T^{s'n}$ depend on disjoint sets of coordinates whenever $s \neq s'$, it follows that $e^{|L-F_n\circ T^{sn}|}$ are independent random variables on Ω for distinct s. Thus,

$$\int_\Omega \exp\left(\sum_{s=0}^{r-1} |L - F_n \circ T^{sn}|\right) \mathrm{d}\mu = \prod_{s=0}^{r-1}\int_\Omega e^{|L-F_n\circ T^{sn}|}\,\mathrm{d}\mu,$$

which in turn implies that

$$\begin{aligned}\mu\left\{\omega : \left|L - \frac{1}{r}\sum_{s=0}^{r-1} F_n(T^{sn}\omega)\right| \geq \varepsilon\right\} &\leq e^{-r\varepsilon}\prod_{s=0}^{r-1}\int_\Omega e^{|L-F_n\circ T^{sn}|}\,\mathrm{d}\mu \\ &= \left(e^{-\varepsilon}\int_\Omega e^{|L-F_n|}\,\mathrm{d}\mu\right)^r,\end{aligned} \tag{5.2}$$

where the second line follows from the T-invariance of μ.

It remains to bound the integral on the right-hand side of (5.2). To that end, take $\delta < \varepsilon/2$, split Ω into the regions where $|L - F_n| \geq \delta$ and $|L - F_n| < \delta$, and then apply Theorem 3.1 to estimate the measure of the former region:

$$\int_\Omega e^{|L-F_n(\omega)|}\,\mathrm{d}\mu(\omega) \leq C\Gamma^2 e^{-\eta n} + e^{\delta}. \tag{5.3}$$

Since $\delta < \varepsilon/2$, we may choose $n_0 = n_0(\varepsilon)$ large enough that

$$C\Gamma^2 e^{-\eta n_0} + e^{\delta} < e^{\frac{\varepsilon}{2}}. \tag{5.4}$$

In view of (5.2), (5.3), and (5.4), we have

$$\mu\left\{\omega : \left|L - \frac{1}{r}\sum_{s=0}^{r-1} F_n(T^{sn}\omega)\right| \geq \varepsilon\right\} \leq e^{-\frac{\varepsilon r}{2}}$$

for all $n \geq n_0$ and all $r \in \mathbb{Z}_+$. □

Using the lemma, we can get a full-measure set of $\omega \in \Omega$ on which one can control transfer matrices across blocks of length n if one is willing to average over n^2 consecutive blocks.

Proposition 5.2. *For any $0 < \varepsilon < 1$, there exists a subset $\Omega_+ = \Omega_+(\varepsilon) \subset \Omega$ of full μ-measure such that the following statement holds true. For every $\varepsilon \in (0,1)$ and every $\omega \in \Omega_+(\varepsilon)$, there exists an $\widetilde{n}_0 = \widetilde{n}_0(\omega, \varepsilon)$ such that*

$$\left| L(E) - \frac{1}{n^2} \sum_{s=0}^{n^2-1} F_n\left(T^{\zeta_0+sn}(\omega), E\right) \right| < \varepsilon \tag{5.5}$$

for all $n, \zeta_0 \in \mathbb{Z}$ with $n \geq \max\left(\widetilde{n}_0, (\log(|\zeta_0|+1))^{2/3}\right)$ and all $E \in \hat{\Sigma}$.

Proof. The key realization in this proof is that one can obtain good control of the averages of the functions $F_n(\omega, E)$ globally over $E \in \hat{\Sigma}$ by controlling F_n on suitable finite subsets of $\hat{\Sigma}$ whose cardinality can in turn be bounded via the perturbative estimates from Lemma 3.3 and the compactness of $\hat{\Sigma}$.

For a given large enough n (we will determine largeness later), we first consider sets $\mathcal{B}_{n,\zeta_0} = \mathcal{B}_{n,\zeta_0}(\varepsilon)$ where (5.5) fails to hold:

$$\mathcal{B}_{n,\zeta_0} := \left\{ \omega : \sup_{E \in \hat{\Sigma}} \left| L(E) - \frac{1}{n^2} \sum_{s=0}^{n^2-1} F_n(T^{\zeta_0+sn}\omega, E) \right| \geq \varepsilon \right\}.$$

Given $0 < \delta \leq 1/2$, define the naïve grid

$$\Sigma_0 := \left[\hat{\Sigma} \cap (2\delta\,\mathbb{Z})\right] \cup \{\pm\kappa\}.$$

It is straightforward to check that Σ_0 is δ-dense in $\hat{\Sigma}$ in the sense that

$$\hat{\Sigma} \subseteq \bigcup_{t \in \Sigma_0} [t-\delta, t+\delta]. \tag{5.6}$$

Moreover, we may estimate the cardinality of Σ_0 via

$$\#\Sigma_0 \leq \frac{\kappa}{\delta} + 3 \leq \frac{2\kappa}{\delta};$$

note that we used $\delta \leq 1/2$ and $\kappa \geq 2$ in the second step. Now fix $0 < \varepsilon < 1$, and let Γ denote the uniform bound on $\|M\|$ from (3.5). Taking $\delta = \varepsilon(3\Gamma^n)^{-1}$ in the discussion above, we may produce a finite set $\Sigma_0 \subset \hat{\Sigma}$ which is $\varepsilon(3\Gamma^n)^{-1}$-dense in $\hat{\Sigma}$ in the sense of (5.6), with cardinality bounded above by

$$\#\Sigma_0 \leq \frac{6\kappa\Gamma^n}{\varepsilon}. \tag{5.7}$$

If necessary, enlarge n to ensure that

$$C\left(\frac{\varepsilon}{3\Gamma^n}\right)^{\beta} < \frac{\varepsilon}{3}, \tag{5.8}$$

where C and β are from Theorem 4.1. Then (3.8) and Theorem 4.1 yield

$$\mathcal{B}_{n,\zeta_0} \subset \bigcup_{E \in \Sigma_0} \left\{ \omega : \left| L(E) - \frac{1}{n^2} \sum_{s=0}^{n^2-1} F_n\left(T^{\zeta_0+sn}\omega, E\right) \right| \geq \frac{\varepsilon}{3} \right\}.$$

Consequently, by taking n large enough that $n \geq n_0(\varepsilon/3)$ and that (5.8) holds and using the T-invariance of μ, one obtains

$$\mu(\mathcal{B}_{n,\zeta_0}) \leq \frac{6\kappa\Gamma^n}{\varepsilon} \exp\left(-\frac{\varepsilon n^2}{6}\right) \tag{5.9}$$

by Lemma 5.1. Now, with

$$\mathcal{B}_n := \bigcup_{|\zeta_0|\leq e^{n^{3/2}}} \mathcal{B}_{n,\zeta_0},$$

it is clear that $\mu(\mathcal{B}_n) \leq e^{-c\varepsilon n^2}$ for n large, so $\Omega_+ := \Omega \setminus \limsup \mathcal{B}_n$ satisfies $\mu(\Omega_+) = 1$ by the Borel–Cantelli lemma. Naturally, for each $\omega \in \Omega_+$, we can find $\widetilde{n}_0 = \widetilde{n}_0(\omega, \varepsilon)$ large enough that $\omega \notin \mathcal{B}_n$ whenever $n \geq \widetilde{n}_0$. In other words,

$$\omega \notin \mathcal{B}_{n,\zeta_0} \qquad \text{whenever } n \geq \widetilde{n}_0(\omega,\varepsilon) \text{ and } |\zeta_0| \leq e^{n^{3/2}}.$$

Changing the order of n and ζ_0, the statement above clearly implies that

$$\omega \notin \mathcal{B}_{n,\zeta_0} \qquad \text{whenever } \zeta_0 \in \mathbb{Z} \text{ and } n \geq \max\left(\widetilde{n}_0, (\log(|\zeta_0|+1))^{2/3}\right).$$

By the definition of B_{n,ζ_0}, we obtain the statement of the proposition. □

We are now in a position to estimate the finite-volume Green functions. Before stating the estimate, we fix some notation. Let $\Lambda = [a,b] \cap \mathbb{Z}$ be a finite subinterval of $\mathbb{Z}$, and denote by $P_\Lambda : \ell^2(\mathbb{Z}) \to \ell^2(\Lambda)$ the canonical projection. We denote the restriction of H_ω to Λ by

$$H_{\omega,\Lambda} := P_\Lambda H_\omega P_\Lambda^*.$$

For any $E \notin \sigma(H_{\omega,\Lambda})$, define

$$G^E_{\omega,\Lambda} := (H_{\omega,\Lambda} - E)^{-1}$$

to be the resolvent operator associated with $H_{\omega,\Lambda}$. Like $H_{\omega,\Lambda}$, $G^E_{\omega,\Lambda}$ has a representation as a finite matrix; denote its matrix elements by $G^E_{\omega,\Lambda}(m,n)$, that is,

$$G^E_{\omega,\Lambda}(m,n) := \left\langle \delta_m, G^E_{\omega,\Lambda}\delta_n \right\rangle, \qquad n,m \in \Lambda.$$

Additionally, for $N \in \mathbb{Z}_+$, let us define $H_{\omega,N} := H_{\omega,[0,N)}$ to be the restriction of H_ω to the box $\Lambda_N := [0,N) \cap \mathbb{Z}$. We will likewise use the same notation for the associated resolvent $G^E_{\omega,N}$.

Using Cramer's rule, we know that

$$G^E_{\omega,N}(j,k) = \frac{\det[H_{\omega,j} - E]\det[H_{T^{k+1}\omega,N-k-1} - E]}{\det[H_{\omega,N} - E]}, \tag{5.10}$$

for any $0 \leq j \leq k \leq N-1$ and $E \notin \sigma(H_{\omega,N})$, where one interprets $\det[H_{\omega,0} - E] = 1$.

Another relation that will be important in what follows is

$$M^E_N(\omega) = \begin{bmatrix} \det(E - H_{\omega,N}) & -\det(E - H_{T\omega,N-1}) \\ \det(E - H_{\omega,N-1}) & -\det(E - H_{T\omega,N-2}) \end{bmatrix}, \qquad N \geq 2. \tag{5.11}$$

This is a standard fact which one may prove inductively.

In particular, since the norm of a matrix majorizes the absolute value of any of its entries, we obtain

$$\left|G^E_{\omega,N}(j,k)\right| \leq \frac{\|M^E_j(\omega)\|\|M^E_{N-k}(T^k\omega)\|}{|\det[H_{\omega,N} - E]|} \tag{5.12}$$

for all $0 \leq j \leq k \leq N-1$ by combining (5.10) and (5.11). Thus, it is straightforward to transform estimates on transfer matrices into estimates on Green functions of truncations of H_ω; to complete our goal of estimating Green functions, we will use Proposition 5.2 to estimate transfer matrix norms and then apply (5.12).

Corollary 5.3. *Given $\varepsilon \in (0,1)$ and $\omega \in \Omega_+(\varepsilon)$, there exists an $\widetilde{n}_1 = \widetilde{n}_1(\omega,\varepsilon)$ such that the following statements hold true. For all $E \in \hat{\Sigma}$, we have*

$$\frac{1}{n}\log\left\|M_n^E(T^{\zeta_0}\omega)\right\| \le L(E) + 2\varepsilon \tag{5.13}$$

whenever $n, \zeta_0 \in \mathbb{Z}$ satisfy $n \ge \max\left(\widetilde{n}_1, \log^2(|\zeta_0|+1)\right)$.

Moreover, for all $n, \zeta_0 \in \mathbb{Z}$ with $n \ge \varepsilon^{-1}\max\left(\widetilde{n}_1, 2\log^2(|\zeta_0|+1)\right)$, we have

$$\left|G^E_{T^{\zeta_0}\omega,n}(j,k)\right| \le \frac{\exp[(n-|j-k|)L(E) + C_0\varepsilon n]}{|\det[H_{T^{\zeta_0}\omega,n} - E]|} \tag{5.14}$$

for all $E \in \hat{\Sigma} \setminus \sigma(H_{T^{\zeta_0}\omega,n})$ and all $j,k \in [0,n)$, where C_0 is a constant that depends only on $\widetilde{\mu}$.

Proof. Fix $\varepsilon \in (0,1)$, $\omega \in \Omega_+(\varepsilon)$, and $E \in \hat{\Sigma}$. As usual, our estimates are independent of the energy, so we suppress E from the notation. Choose $\widetilde{n}_1 \in \mathbb{Z}_+$ large enough that

$$\widetilde{n}_1 \ge \max\left(\widetilde{n}_0(\omega,\varepsilon)^3, 4\varepsilon^{-1}, 2986\right) \quad \text{and} \quad \frac{12\log\Gamma}{\widetilde{n}_1^{1/3} - 3} < \varepsilon. \tag{5.15}$$

Given $n, \zeta_0 \in \mathbb{Z}$ with $n \ge \max\left(\widetilde{n}_1, \log^2(|\zeta_0|+1)\right)$, we want to apply Proposition 5.2 with $m = \lceil n^{1/3} \rceil$. Notice that $0 \le m^3 - n \le 3m^2$. Thus, by submultiplicativity of the matrix norm and unimodularity of the transfer matrices, we have

$$\|M_n(T^{\zeta_0}\omega)\| \le \Gamma^{3m^2} \prod_{s=0}^{m^2-1} \|M_m(T^{\zeta_0+sm}\omega)\|. \tag{5.16}$$

By our choice of $\widetilde{n}_1$, we have $m \ge \widetilde{n}_0$ and $m \ge (\log(|\zeta_0|+1))^{2/3}$. Thus, combining (5.16) with Proposition 5.2 and (5.15), a direct computation shows that

$$\begin{aligned} \frac{1}{n}\log\|M_n(T^{\zeta_0}\omega)\| &\le \frac{3m^2\log\Gamma}{n} + \frac{m^3}{n}(L+\varepsilon) \\ &\le L + \frac{6\log\Gamma}{m-3} + \frac{m}{m-3}\varepsilon \\ &\le L + 2\varepsilon, \end{aligned} \tag{5.17}$$

which yields (5.13).

Now suppose that $n \ge \varepsilon^{-1}\max\left(\widetilde{n}_1, 2\log^2(|\zeta_0|+1)\right)$, and put $h := \lceil \varepsilon n \rceil$. For $j \ge 0$, we have

$$\left\|M_j(T^{\zeta_0}\omega)\right\| \le \left\|M_{j+h}(T^{\zeta_0-h}\omega)\right\| \left\|[M_h(T^{\zeta_0-h}\omega)]^{-1}\right\|.$$

Clearly, $j \ge 0$ and $h \ge \widetilde{n}_1$. Moreover, by our choice of $\widetilde{n}_1$ and the relation between n and ζ_0, a direct computation shows that $h \ge \log^2(|\zeta_0 - h| + 1)$. Thus, we can apply (5.17) to estimate the norms on the right-hand side and obtain

$$\left\|M_j(T^{\zeta_0}\omega)\right\| \le e^{(j+2h)L(E)+2\varepsilon(j+2h)} \le e^{jL(E)+C_0\varepsilon n}, \tag{5.18}$$

where C_0 is a constant that depends only on $\widetilde{\mu}$.

Naturally, one can also apply the analysis above to estimate the transfer matrix $M_{n-k}(T^{\zeta_0+k}\omega)$ with $j \le k \le n-1$. Using (5.15) and the relationships among h,

ζ_0, and k, a direct computation shows that $\log^2(|\zeta_0 - k + h| + 1) \leq h$. Thus, (5.17) yields

$$\begin{aligned} \|M_{n-k}(T^{\zeta_0+k}\omega)\| &\leq \left\|M_{n-k+h}(T^{\zeta_0+k-h}\omega)\right\| \left\| \left[M_h(T^{\zeta_0+k-h}\omega)\right]^{-1} \right\| \\ &\leq \exp\left[(n-k)L(E) + C_0\varepsilon n\right]. \end{aligned} \tag{5.19}$$

Combining equations (5.18) and (5.19) with the observation (5.12), one sees that, for a suitable choice of C_0,

$$\begin{aligned} |G^E_{T^{\zeta_0}\omega,n}(j,k)| &\leq \frac{\|M_j(T^{\zeta_0}\omega)\|\|M_{n-k}(T^{\zeta_0+k}\omega)\|}{|\det[H_{T^{\zeta_0}\omega,n} - E)]|} \\ &\leq \frac{\exp[(n-|j-k|)L(E) + C_0\varepsilon n]}{|\det[H_{T^{\zeta_0}\omega,n} - E)]|} \end{aligned}$$

for all $E \in \hat{\Sigma} \setminus \sigma(H_{T^{\zeta_0}\omega,n})$ and all $0 \leq j \leq k < n$. The case $j \geq k$ follows because H is self-adjoint and E is real. □

6. Proof of spectral and exponential dynamical localization

In the present section, we conclude the proof of Theorem 1.2, which, as discussed in the Introduction, implies Theorem 1.1 by standard reasoning. We also state and prove Theorem 6.4, which contains our exponential dynamical localization result.

The key remaining cornerstone is supplied by Proposition 6.1, which is a version of an argument usually referred to as the elimination of double resonances. We note that double resonances appear frequently in, and are one of the most subtle parts of, the mathematical analysis of Anderson localization. In particular, from the proof of the present paper, or the proof of [9, 10], one may see that, for 1D ergodic Schrödinger operators, uniform positivity and uniform LDT of the Lyapunov exponent, together with the elimination of double resonances, imply Anderson localization for suitable parameters. Moreover, the set of parameters (phases, frequencies) for which Anderson localization holds true depends exactly on the bad set that is eliminated by the exclusion of double resonances.

It is also interesting to note that, in some sense, the elimination of double resonances detects the randomness of the base dynamics in a very sensitive manner. For instance, a version of Proposition 6.1 was developed by Bourgain and Goldstein for real-analytic quasi-periodic potentials in [9], where the authors invoked suitable complexity bounds for *semialgebraic* sets. Bourgain and Schlag [10] considered operators with strongly mixing potentials, where the elimination becomes a bit easier due to the strong mixing property. Obviously the Anderson model considered in the present paper is closer to the one in [10], and the elimination process is even more transparent because of the independence of the potential values.

After we prove Proposition 6.1, this result is then used to supply estimates that enable us to run the avalanche principle and prove positivity (and existence) of *nonaveraged* LEs

$$L(E,\omega) := \lim_{n\to\infty} \frac{1}{n} \log \|M^E_n(\omega)\|.$$

Once we have existence and positivity of $L(E,\omega)$ for a full-measure set of $\omega \in \Omega$ and a suitably rich (ω-dependent) set of energies, we are able to deduce the modified Anderson localizer's dream. Thus, the avalanche principle in some sense plays the role of MSA in our arguments. As discussed in the Introduction, the Anderson localizer's dream itself is false; that is, one does not have positivity (or

even existence) of $L(E,\omega)$ for all E and a uniform full-measure set of $\omega \in \Omega$. After proving the modified Anderson localizer's dream, we can then make a second pass through the localization argument to obtain better estimates; in particular, we obtain centered versions of the localization estimates with constants that depend only on ω and the center of localization. This supplies a sufficient input to deduce almost-sure exponential dynamical localization.

For $N \in \mathbb{Z}_+$, we define

$$\overline{N} := \left\lfloor N^{\log N} \right\rfloor = \left\lfloor e^{(\log N)^2} \right\rfloor,$$

which is a superpolynomially and subexponentially growing function of N.

We now introduce the set of double resonances. Given $\varepsilon > 0$ and $N \in \mathbb{Z}_+$, let $\mathcal{D}_N = \mathcal{D}_N(\varepsilon)$ denote the set of all $\omega \in \Omega$ such that

$$\|G^E_{T^\zeta\omega,[-N_1,N_2]}\| \geq e^{K^2} \tag{6.1}$$

and

$$F_m(T^{\zeta+r}\omega, E) \leq L(E) - \varepsilon \tag{6.2}$$

for some choice of $\zeta \in \mathbb{Z}$, $K \geq \max(N, \log^2(|\zeta|+1))$, $0 \leq N_1, N_2 \leq K^9$, $E \in \Sigma$, $K^{10} \leq r \leq \overline{K}$, and $m \in \{K, 2K\}$ (F_m is as defined in (3.4)).

Proposition 6.1. *For all $0 < \varepsilon < 1$, there exist constants $C > 0$ and $\widetilde{\eta} > 0$ such that*

$$\mu(\mathcal{D}_N(\varepsilon)) \leq Ce^{-\widetilde{\eta}N}$$

for all $N \in \mathbb{Z}_+$.

Proof. Define auxiliary "bad sets" for fixed ζ and K:

$$\mathcal{D}_{K,\zeta} = \{\omega : (6.1), (6.2) \text{ are satisfied for some choice of } E, N_1, N_2, r, m \text{ as above}\}.$$

Fix $\varepsilon \in (0,1)$, and begin by noticing that

$$\mathcal{D}_{K,\zeta} \subset \bigcup_{K^{10} \leq r \leq \overline{K}} \bigcup_{0 \leq N_1, N_2 \leq K^9} \widetilde{\mathcal{D}}_1(N_1, N_2, r, \zeta) \cup \widetilde{\mathcal{D}}_2(N_1, N_2, r, \zeta), \tag{6.3}$$

where $\widetilde{\mathcal{D}}_j(N_1, N_2, r, \zeta)$ denotes the collection of all $\omega \in \Omega$ for which there exists $E \in \Sigma$ such that (6.1) and (6.2) hold with $m = jK$. We will estimate $\mu(\widetilde{\mathcal{D}}_1)$. The estimates for $\widetilde{\mathcal{D}}_2$ are completely analogous. To that end, suppose that $\omega \in \widetilde{\mathcal{D}}_1(N_1, N_2, r, \zeta)$, i.e., (6.1) and (6.2) hold for some $E \in \Sigma$. By the spectral theorem, there exists an $E_0 \in \sigma(H_{T^\zeta\omega,[-N_1,N_2]})$ with

$$|E - E_0| \leq \left\| G^E_{T^\zeta\omega,[-N_1,N_2]} \right\|^{-1} \leq e^{-K^2}. \tag{6.4}$$

On the other hand, choosing K large enough that $\Gamma^K e^{-K^2} \leq \frac{\varepsilon}{4}$ and $Ce^{-\beta K^2} \leq \frac{\varepsilon}{4}$ (where C, β are from (4.1)), we get

$$\begin{aligned} F_K(T^{\zeta+r}\omega, E_0) &\leq F_K(T^{\zeta+r}\omega, E) + \frac{\varepsilon}{4} \\ &\leq L(E) - \varepsilon + \frac{\varepsilon}{4} \\ &\leq L(E_0) - \frac{\varepsilon}{2}, \end{aligned}$$

where we have used Lemma 3.3 in the first line, (6.2) in the second line, and Theorem 4.1 in the final line. Thus, when K is large enough, we get

$$\widetilde{\mathcal{D}}_1(N_1,N_2,r,\zeta)\subseteq\hat{\mathcal{D}}_1(N_1,N_2,r,\zeta)\qquad\text{for all } N_1,N_2,r,\text{ and }\zeta, \tag{6.5}$$

where $\hat{\mathcal{D}}_1=\hat{\mathcal{D}}_1(N_1,N_2,r,\zeta)$ denotes the set of all $\omega\in\Omega$ such that

$$F_K(T^{\zeta+r}\omega,E_0)\le L(E_0)-\frac{\varepsilon}{2}$$

for some $E_0\in\sigma\left(H_{T^\zeta\omega,[-N_1,N_2]}\right)$.

Now the conditions (6.1) and (6.2) depend only on a finite number of entries of ω. Concretely, we notice that $G^E_{T^\zeta\omega,[-N_1,N_2]}$ depends on $\vec{\omega}':=(\omega_{\zeta-N_1},\dots,\omega_{\zeta+N_2})$, while $M^E_K(T^{\zeta+r}\omega)$ depends only on $\vec{\omega}'':=(\omega_{\zeta+r},\dots,\omega_{\zeta+r+K-1})$. In particular, (6.1) and (6.2) depend on independent sets of random variables. Consequently, we obtain

$$\int_{\mathcal{A}^{[\zeta+r,\zeta+r+K)}}\chi_{\hat{\mathcal{D}}_1}(\omega)\,\mathrm{d}\widetilde{\mu}^K(\vec{\omega}'')\le C(N_1+N_2+1)e^{-\eta_1K}$$

for each fixed choice of $\vec{\omega}'\in\mathcal{A}^{[-N_1+\zeta,N_2+\zeta]}$. Then, since membership in $\hat{\mathcal{D}}_1$ is determined entirely by coordinates in $\zeta+[-N_1,N_2]$ and $\zeta+[r,r+K)$, we have

$$\begin{aligned}\mu(\hat{\mathcal{D}}_1)&=\int_\Omega\chi_{\hat{\mathcal{D}}_1}(\omega)\,\mathrm{d}\mu(\omega)\\&=\int\limits_{\mathcal{A}^{\zeta+[-N_1,N_2]}}\ \int\limits_{\mathcal{A}^{\zeta+[r,r+K)}}\chi_{\hat{\mathcal{D}}_1}(\omega)\,\mathrm{d}\widetilde{\mu}^K(\vec{\omega}'')\,\mathrm{d}\widetilde{\mu}^{N_1+N_2+1}(\vec{\omega}')\\&\le CK^9e^{-\eta_1K}\\&\le Ce^{-\eta_2K}.\end{aligned}$$

Thus, we obtain $\mu(\widetilde{\mathcal{D}}_1(N_1,N_2,r,\zeta))\le Ce^{-\eta_2K}$ by applying (6.5). Applying similar reasoning to $\widetilde{\mathcal{D}}_2$, one can estimate $\mu(\widetilde{\mathcal{D}}_2(N_1,N_2,r,\zeta))\le Ce^{-\eta_3K}$; putting everything together yields

$$\begin{aligned}\mu(\mathcal{D}_{K,\zeta})&\le\sum_{0\le N_1,N_2\le K^9}\ \sum_{K^{10}\le r\le\overline{K}}\Big(\mu(\widetilde{\mathcal{D}}_1(N_1,N_2,r))+\mu(\widetilde{\mathcal{D}}_2(N_1,N_2,r))\Big)\\&\le CK^{18}\overline{K}e^{-\eta_4K}\\&\le Ce^{-2\widetilde{\eta}K}\end{aligned}$$

for some suitable choice of $\widetilde{\eta}$. Changing the order of K and ζ, we have

$$\mathcal{D}_N=\bigcup_{\zeta\in\mathbb{Z}}\ \bigcup_{K\ge\max\{N,\log^2(|\zeta|+1)\}}\mathcal{D}_{K,\zeta}\subseteq\bigcup_{K\ge N}\ \bigcup_{|\zeta|\le e^{\sqrt{K}}}\mathcal{D}_{K,\zeta}.$$

Then the estimates above yield

$$\mu(\mathcal{D}_N)\le\sum_{K\ge N}(2e^{\sqrt{K}}+1)Ce^{-2\widetilde{\eta}K}\le Ce^{-\widetilde{\eta}N}$$

for large enough N. Adjusting the constants to account for small N concludes the proof. □

By Proposition 6.1, the set

$$\Omega_-=\Omega_-(\varepsilon):=\Omega\setminus\limsup\mathcal{D}_N(\varepsilon)$$

has full μ-measure. We are now in a position to define the full-measure set upon which Anderson localization holds. First, put

$$\Omega_0 = \Omega_\Sigma \cap \bigcap_{\varepsilon\in(0,1)} \Omega_+(\varepsilon) \cap \Omega_-(\varepsilon),$$

where $\Omega_+(\varepsilon)$ is as defined in Proposition 5.2, and Ω_Σ denotes the set of $\omega \in \Omega$ for which $\sigma(H_\omega) = \Sigma$. In essence, $\omega \in \Omega_+$ gives us upper bounds on $F_N(\omega, E)$ on $\hat{\Sigma}$, while $\omega \in \Omega_-$ will give us lower bounds on F_N provided that we can prove an estimate like (6.1). Then we define

$$\Omega_* = \Omega_0 \cap \mathcal{R}[\Omega_0],$$

where $\mathcal{R}$ denotes the reflection $[\mathcal{R}\omega]_n = \omega_{-1-n}$. It is straightforward to verify that Ω_* has full μ-measure.

We will need the following standard formula that relates solutions of the difference equation (1.1) and truncated Green functions at energy E. Suppose that $a \le b$ are integers, and that $n \in \Lambda := [a, b] \cap \mathbb{Z}$. If u is a solution of the difference equation $Hu = Eu$ in the sense of (1.1) and $E \notin \sigma(H_\Lambda)$, then

$$u(n) = -G^E_\Lambda(n, a)u(a - 1) - G^E_\Lambda(n, b)u(b + 1). \tag{6.6}$$

We now have all of the necessary pieces to do the following two things:

(1) Prove the modified Anderson localizer's dream. That is, we will show that generalized eigenfunctions exhibit Lyapunov behavior for almost every $\omega \in \Omega$ and every $E \in \mathcal{G}(H_\omega)$.
(2) The following version of SULE (semiuniformly localized eigenfunctions), from which we will deduce exponential dynamical localization for μ-a.e. ω.

Theorem 6.2 (SULE)**.** *For every $\delta > 0$ and $\omega \in \Omega_*$, there exist constants C_δ and $C_{\omega,\delta}$ such that, for every eigenfunction u of H_ω and every $n \in \mathbb{Z}$,*

$$|u(\bar{\zeta} + n)| \le C_{\omega,\delta}\|u\|_\infty e^{C_\delta \log^{22}(|\bar{\zeta}|+1)} e^{-(1-\delta)L(E)|n|} \tag{6.7}$$

for some $\bar{\zeta}$ that depends on u.

It turns out that these two can be done via two passes through the same argument. We begin with a generalized eigenfunction u of H_ω at energy $E \in \mathcal{G}(H_\omega)$, obeying (1.2). Since u cannot vanish identically, we can pick $\zeta \in \mathbb{Z}$ such that $u(\zeta) \ne 0$ and normalize u so that $u(\zeta) = 1$. In fact, at this stage, we may choose $\zeta = 0$ or 1. Since $\omega \in \Omega_0$, we have $\omega \in \widetilde{\Omega}^{(N)} := \Omega_+(\varepsilon) \setminus \mathcal{D}_N(\varepsilon)$ for all $N \ge N_0$, where N_0 is sufficiently large.

We will then use Proposition 6.1, the bounds on the norms of the Green functions, and the avalanche principle in a somewhat subtle fashion. Initially, since we are dealing with ω in the good set $\widetilde{\Omega}^{(N)}$, our goal is to establish that there are some $0 \le N_1, N_2 \le K^9$ such that (6.1) holds; that is,

$$\|G^E_{T^\zeta\omega,[-N_1,N_2]}\| \ge e^{K^2}. \tag{6.8}$$

Having established (6.8), it then follows that expression (6.2) fails for all choices of $K^{10} \le r \le \bar{K}$ and $m \in \{K, 2K\}$. This supplies lower bounds on F_K and F_{2K} that we can then use to run the avalanche principle. This will lead to the modified Anderson localizer's dream, and hence to exponential decay of u. In this process, the constants and largeness conditions depend not only on ω and ε but also on u. However, once we know that u is exponentially decaying, we can pass it through

the same argument centered around its global maximum to get much better control on the constants involved; in particular, we get uniformity of constants in u at the expense of introducing a constant that grows subexponentially in the center of localization.

We wish to emphasize that, to get spectral localization, it suffices to deal with $\zeta = 0$ or 1, from which we may get estimates like $|u(\zeta+n)| \leq C(\omega,u)e^{-\xi n}$; here, the constants depend not only on ω but also on the eigenfunction u. Thus, we cannot get any meaningful estimates of dynamical quantities from the "first pass" through the localization argument, since such quantities tend to involve all eigenfunctions at once.

To tackle this issue, we need the full strength of the statements from Section 5 and the present section to get all $\zeta \in \mathbb{Z}$ involved. In particular, we will see that the base scale for a function localized at ζ is $K \geq \varepsilon^{-1}\max(N, 2\log^2(|\zeta|+1))$ for a suitable choice of $\varepsilon > 0$ and some large N independent of ζ.

Proof of Theorem 1.2 (*Modified Anderson localizer's dream*). We will show that Ω_* is the desired full-measure subset of Ω. So let $\omega \in \Omega_*$ and $E \in \mathcal{G}(H_\omega)$ be given. Since Ω_* is $\mathcal{R}$-invariant and $\mathcal{G}(H_{\mathcal{R}\omega}) = \mathcal{G}(H_\omega)$, it suffices to show that

$$\lim_{N\to\infty} \frac{1}{N}\log\|M_N^E(\omega)\| = L(E). \tag{6.9}$$

Since $\omega \in \Omega_+(\varepsilon)$ for each $\varepsilon > 0$, Corollary 5.3 yields

$$\limsup_{N\to\infty} \frac{1}{N}\log\|M_N^E(\omega)\| \leq L(E);$$

so, to prove (6.9), it remains to show that

$$\liminf_{N\to\infty} \frac{1}{N}\log\|M_N^E(\omega)\| \geq L(E). \tag{6.10}$$

To that end, let $\varepsilon \in (0,1)$ be given, and let u denote a generalized eigenfunction of H_ω corresponding to energy E and satisfying (1.2). After normalizing, we may assume that $u(\zeta) = 1$ for some choice of $\zeta \in \{0,1\}$. Define

$$K = K(N,\zeta) := \left\lceil \frac{1}{\varepsilon}\max\left(N, 2\log^2(|\zeta|+1)\right)\right\rceil, \tag{6.11}$$

where $N \in \mathbb{Z}_+$ is sufficiently large. More specifically, we take $N \geq N_0$, where $N_0 = N_0(\omega,\varepsilon) := \max\{\widetilde{n}_0(\omega,\varepsilon), \widetilde{n}_1(\omega,\varepsilon), \widetilde{n}_2(\omega,\varepsilon)\}$, where $\widetilde{n}_0$ comes from Proposition 5.2, where $\widetilde{n}_1$ is from Corollary 5.3, and where $\widetilde{n}_2$ is chosen such that $\omega \in \widetilde{\Omega}^{(N)} := \Omega_+(\varepsilon) \setminus \mathcal{D}_N(\varepsilon)$ whenever $N \geq \widetilde{n}_2(\omega,\varepsilon)$. Of course, right now, $\zeta \in \{0,1\}$, so $K \sim \varepsilon^{-1}N$. We keep up with the dependence on ζ to facilitate the "second pass" through the argument with ζ chosen to be a location at which $|u|$ is maximized.

Let us begin by first proving the following claim.

Claim 1. There exist integers a_i, b_i, $i = 1,2$ such that

$$-K^9 \leq a_1 \leq -K^3 + 1, \tag{6.12}$$

$$0 \leq a_2 \leq K^9, \tag{6.13}$$

and $b_i \in \{a_i + K^3 - 2, a_i + K^3 - 1, a_i + K^3\}$ such that

$$\left|G^E_{T^\zeta\omega,\Lambda_i}(j,k)\right| \leq \exp\left(-|j-k|L(E) + C_0\varepsilon K^3\right), \tag{6.14}$$

for any $j,k \in \Lambda_i := [a_i, b_i)$, where C_0 is a constant that depends only on $\widetilde{\mu}$.

Proof. The claim follows by using Corollary 5.3 to estimate the Green function, (5.11), to exchange the characteristic polynomial for a norm of a transfer matrix, and Proposition 5.2 to find a suitable starting point for said transfer matrix.

More precisely, we begin by applying Proposition 5.2 with $n = K^3$ twice: once with $\zeta_0 = \zeta$ and once with $\zeta_0 = \zeta - K^9$. With $\zeta_0 = \zeta$, we get

$$\left| L(E) - K^{-6} \sum_{s=0}^{K^6-1} F_{K^3}(T^{\zeta+sK^3}\omega) \right| < \varepsilon;$$

with $\zeta_0 = \zeta - K^9$, we have

$$\left| L(E) - K^{-6} \sum_{s=0}^{K^6-1} F_{K^3}(T^{\zeta-K^9+sK^3}\omega) \right| < \varepsilon.$$

Note that the second invocation of Proposition 5.2 requires that $K^3 \geq \log^{2/3}(|\zeta_0 - K^9| + 1)$, which may be obtained by a direct computation.

Thus, a straightforward argument by contradiction yields s_i satisfying $-K^9 \leq s_1 \leq -K^3$, $0 \leq s_2 \leq K^9 - K^3$, and

$$L(E) - \frac{1}{K^3} \log \|M^E_{K^3}(T^{\zeta+s_i}\omega)\| < \varepsilon. \tag{6.15}$$

Since the norm of a 2×2 matrix can be dominated by 4 times its greatest entry, this yields

$$L(E) - \frac{1}{K^3} \log |\det[H_{T^{\zeta+a_i}\omega,k_i} - E]| \ < 2\varepsilon \tag{6.16}$$

for some choice of $a_i \in \{s_i, s_i + 1\}$ and $k_i \in \{K^3, K^3 - 1, K^3 - 2\}$ as long as N_0 is large enough that $N_0^{-3} \log 4 \leq \varepsilon$. Put $b_i = a_i + k_i$ and $\Lambda_i = [a_i, b_i)$.

Now, combining (5.14) with (6.16), we obtain

$$\begin{aligned}
\left| G^E_{T^\zeta\omega,\Lambda_i}(j,k) \right| &= \left| G^E_{T^{\zeta+a_i}\omega,k_i}(j - a_i, k - a_i) \right| \\
&\leq \frac{\exp\left((k_i - |j-k|)L(E) + C_0\varepsilon k_i\right)}{|\det[H_{T^{\zeta+a_i}\omega,k_i} - E]|} \\
&\leq \frac{\exp\left((K^3 - |j-k|)L(E) + C_0\varepsilon K^3\right)}{\exp\left(K^3(L(E) - 2\varepsilon)\right)} \\
&= \exp\left(-|j-k|L(E) + (C_0 + 2)\varepsilon K^3\right)
\end{aligned}$$

for all $j, k \in \Lambda_i$. Note here we may need to uniformly enlarge N_0 to ensure $\varepsilon k_i \geq \max(\widetilde{n}_1, 2\log^2(|\zeta + a_i| + 1))$. □

Claim 2. Recall that u satisfies $|u(n)| \leq C_u(1 + |n|)$ and $|u(\zeta)| = 1$. Then with Λ_i as in Claim 1, put

$$\ell_i = \left\lfloor \frac{a_i + b_i}{2} \right\rfloor.$$

Then we have

$$|u(\zeta + \ell_i)| \leq e^{-2K^2}, \qquad i = 1, 2.$$

whenever $N > N_0$ is large enough. Here, the size of N_0 depends on ω, ε, and u (with the dependence on u entering solely through C_u).

Proof. From (6.6), we obtain

$$|u(\zeta+\ell_i)| \le \big|G^E_{\omega,\Lambda_i}(\ell_i,a_i)\big|\big|u(\zeta+a_i-1)\big| + \big|G^E_{\omega,\Lambda_i}(\ell_i,b_i-1)\big|\big|u(\zeta+b_i)\big|.$$

Recall that $\log^2(|\zeta|+1) \le K$, which implies that $|\zeta| \le e^{\sqrt{K}}$. This yields

$$\begin{aligned}|u(\zeta+\ell_i)| &\le C_u(K^9+e^{\sqrt{K}})e^{C_0\varepsilon K^3}(e^{-|\ell_i-a_i|L(E)}+e^{-|\ell_i-(b_i-1)|L(E)})\\ &\le 2C_u(K^9+e^{\sqrt{K}})\exp\Big(-L(E)K^3/3+C_0\varepsilon K^3\Big).\end{aligned}$$

Then, for N large enough (depending only on ω,ε,u), we can estimate the last line by e^{-2K^2}, which concludes the proof of the claim. □

Remark 6.3. Claim 2 is the only place that requires the dependence of N_0 on u, and the dependence comes from the u-dependent constant C_u. In particular, if it is known that u is a normalized eigenvector with $\max|u(n)|=1$, then the largeness of N depends only on (ω,ε). In the argument after this remark, the largeness of N_0 will be independent of u. In particular, whenever we say "enlarge N_0 if necessary", the reader may verify that such an enlargement may be performed in a u-independent fashion.

Now we use $|u(\zeta)|=1$ and (6.6) with $a=\zeta+\ell_1+1$ and $b=\zeta+\ell_2-1$ to get

$$\begin{aligned}1&=|u(\zeta)|\\ &\le |G^E_{\omega,[a,b]}(\zeta,a)||u(\zeta+\ell_1)| + |G^E_{\omega,[a,b]}(\zeta,b)||u(\zeta+\ell_2)|\\ &\le \Big(G^E_{\omega,[a,b]}(\zeta,a)| + |G^E_{\omega,[a,b]}(\zeta,b)|\Big)\,e^{-2K^2}.\end{aligned}$$

From this, we deduce

$$\Big\|G^E_{\omega,[a,b]}\Big\| \ge \Big\|G^E_{\omega,[a,b]}\delta_\zeta\Big\| \ge \frac{\sqrt{2}}{2}e^{2K^2} \ge e^{K^2}.$$

Thus, (6.1) follows; since $0\le -(a-\zeta),(b-\zeta)\le K^9$, we can in turn conclude that expression (6.2) fails for every $K^{10}\le r\le \overline{K}$ and $m=K,2K$; that is, we have

$$\frac{1}{m}\log\|M^E_m(T^{\zeta+r}\omega)\| > L(E)-\varepsilon \tag{6.17}$$

whenever $K^{10}\le r\le\overline{K}$ and $m\in\{K,2K\}$.

Now use (6.17) to apply the avalanche principle. Concretely, choose $n\in\mathbb{Z}_+$ with $K^{10}\le n\le K^{-1}\overline{K}-K^9$, define

$$A^{(j)} := M^E_K(T^{\zeta+K^{10}+(j-1)K}\omega), \qquad 1\le j\le n.$$

With $\lambda:=\exp(K(L(E)-\varepsilon))$, (6.17) gives

$$\|A^{(j)}\| \ge \lambda \ge n$$

for all j, where the second inequality holds as long as N_0 is sufficiently large. Since $K\ge\widetilde{n}_1$ and $K\ge\log^2(|\zeta|+|\overline{K}|+1)$ (enlarge N_0 if necessary), we may use (5.13) to obtain

$$\|A^{(j)}\| \le \exp\big(K(L(E)+2\varepsilon)\big), \qquad 1\le j\le n.$$

Thus, (6.17) implies

$$\begin{aligned}\Big|\log\|A^{(j+1)}\| + \log\|A^{(j)}\| - \log\|A^{(j+1)}A^{(j)}\|\Big| &< 2K(L(E)+2\varepsilon) - 2K(L(E)-\varepsilon)\\ &= 6K\varepsilon\\ &\le \frac{1}{2}\log\lambda,\end{aligned}$$

where the final inequality needs ε to be sufficiently small; it is easy to see that this smallness condition depends only on $\widetilde{\mu}$ (through γ). Thus, taking $\hat{N} = nK$ and $r_0 = K^{10}$, we have $\hat{N} \in [K^{11}, \overline{K} - K^{10}]$, and the avalanche principle (Lemma 4.4) yields

$$\begin{aligned}\log\|M_{\hat{N}}(T^{\zeta+r_0}\omega)\| &= \log\|A^{(n)}\cdots A^{(1)}\|\\ &\ge \sum_{j=1}^{n-1}\log\|A^{(j+1)}A^{(j)}\| - \sum_{j=2}^{n-1}\log\|A^{(j)}\| - C\frac{n}{\lambda}\\ &\ge (n-1)2K(L(E)-\varepsilon) - (n-2)K(L(E)+2\varepsilon) - C\\ &\ge \hat{N}(L(E)-5\varepsilon)\end{aligned}$$

by choosing N_0 large.

Putting this together, we can control $\|M_\ell^E(T^\zeta\omega)\|$ for general $K^{11}+K^{10} \le \ell \le \bar{K}$ by interpolation. In particular, writing $\ell = nK + p$ with $0 \le p < K$ and $n \ge K^{10} + K^9$, we have

$$\begin{aligned}\|M_\ell^E(T^\zeta\omega)\| &\ge \frac{\|M_{\ell-K^{10}}(T^{\zeta+K^{10}}\omega)\|}{\|M_{K^{10}}(T^\zeta\omega)\|}\\ &\ge \Gamma^{-K^{10}-p}\|M_{nK-K^{10}}(T^{\zeta+K^{10}}\omega)\|\\ &\ge \Gamma^{-K^{10}-p}e^{(nK-K^{10})(L(E)-5\varepsilon)}\\ &\ge e^{\ell(L(E)-6\varepsilon)}\end{aligned} \tag{6.18}$$

as long as N is sufficiently large. Since the intervals $[K^{11}+K^{10}, \bar{K}]$ cover all sufficiently large integers, the foregoing estimates yield

$$\liminf_{n\to\infty}\frac{1}{n}\log\|M_n^E(T^\zeta\omega)\| \ge L(E) - 6\varepsilon. \tag{6.19}$$

Since $\zeta \in \{0,1\}$ and since (6.19) holds for all $\varepsilon > 0$, we obtain (6.10), as desired. □

Proof of Theorem 6.2 (*SULE*). Let $\omega \in \Omega_*$, $E \in \mathcal{G}(H_\omega)$, and $\delta > 0$ be given. By Theorem 1.2, the associated eigenvector u decays exponentially, and hence (after normalization), we may define the center of localization $\bar{\zeta}$ via $u(\bar{\zeta}) = \|u\|_\infty = 1$. There is an unimportant ambiguity here—namely, that $|u|$ can obtain its maximum value multiple times. However, since $u \in \ell^2$, it may do so only finitely many times, and it does not matter which of those occurrences we use for $\bar{\zeta}$. Then define $K = K(N, \bar{\zeta})$ as in (6.11), and suppose that N is large.

Now fix $\varepsilon > 0$ small; it will be apparent that "how small" depends only on $\widetilde{\mu}$ and δ. Then, running the proof of Theorem 1.2 with ζ replaced by $\bar{\zeta}$, we obtain everything from Claim 1 to (6.18) (with $\bar{\zeta}$ replacing ζ). However, this time, we

get $N \geq N_0(\omega, \varepsilon)$; that is, N_0 no longer depends on u. Then, for any ℓ with $K^{11} + K^{10} \leq \ell \leq \overline{K}$, we get

$$\frac{1}{\ell} \log \|M_\ell^E(T^{\bar{\zeta}}\omega)\| \geq L(E) - 6\varepsilon.$$

In particular, combining this with (5.11) and (5.14) implies that there exist $\bar{\zeta}' \in \{\bar{\zeta}, \bar{\zeta}+1\}$ and $\ell' \in \{\ell, \ell-1, \ell-2\}$ such that

$$\left|G^E_{T^{\bar{\zeta}'}\omega,\ell'}(j,k)\right| \leq \exp\left(-|j-k|L(E) + (C_0+6)\varepsilon\ell'\right)$$

for any $j, k \in [0, \ell')$. Pick $n \in [\frac{1}{4}\ell', \frac{1}{2}(\ell'-1)]$ (notice that $\ell' - n \geq n$). Then (6.6) and the normalization $\|u\|_\infty = 1$ yield

$$\begin{aligned}
|u(\bar{\zeta}'+n)| &\leq |G^E_{T^{\bar{\zeta}'}\omega,\ell'}(0,n)| + |G^E_{T^{\bar{\zeta}'}\omega,\ell'}(n,\ell'-1)| \\
&\leq \exp\left(-L(E)n + (C_0+6)\varepsilon\ell'\right) + \exp\left(-L(E)(\ell'-n) + (C_0+6)\varepsilon\ell'\right) \\
&\leq 2\exp\left(-L(E)n + 4(C_0+6)\varepsilon n\right) \\
&\leq e^{-(1-\delta)L(E)n},
\end{aligned}$$

where the last line needs ε sufficiently small, depending on the choice of $\delta > 0$. Shifting things back to $\bar{\zeta}$ if necessary, we get

$$|u(\bar{\zeta}+n)| \leq Ce^{-(1-\delta)L(E)n} \qquad \text{for every } n \in \left[\frac{1}{4}(K^{11}+K^{10}), \frac{1}{2}\overline{K-3}\right]. \tag{6.20}$$

These intervals are overlapping for N_0 large enough (and here the largeness does not depend on any parameters), so we conclude that (6.20) holds true for any

$$n \geq \frac{\varepsilon^{-11}}{4} \max(N_0, 2\log^2(|\bar{\zeta}|+1))^{11},$$

where N_0 depends only on ω (recall that ε depends on δ). For $0 \leq n \leq \frac{\varepsilon^{-11}}{4} \max(N_0, 2\log^2(|\bar{\zeta}|+1))^{11}$, we may estimate $|u(\bar{\zeta}+n)|$ trivially and adjust constants accordingly:

$$\begin{aligned}
|u(\zeta+n)| &\leq e^{(1-\delta)L(E)\frac{\varepsilon^{-11}}{4}\max(N_0, 2\log^2(|\bar{\zeta}|+1))^{11}} e^{-(1-\delta)L(E)n} \\
&\leq C_{\omega,\delta} e^{C_\delta \log^{22}(|\bar{\zeta}|+1)} e^{-(1-\delta)L(E)n},
\end{aligned}$$

where $C_{\omega,\delta}$ depends on ω, δ, and C_δ depends only on δ. This proves the estimates in (6.7) for all $n \geq 0$. To deal with $n < 0$, simply use the reflection-invariance of Ω_*. □

We can use the version of SULE from Theorem 6.2 to prove the following version of almost-sure exponential dynamical localization. Recall that γ denotes the global uniform lower bound on the Lyapunov exponent—compare this to (2.4)—and that it coincides with the minimum of the Lyapunov exponent on the almost-sure spectrum Σ—compare this to (2.5).

Theorem 6.4 (Exponential dynamical localization)**.** *For any $\omega \in \Omega_*$, $\epsilon > 0$, and $0 < \beta < \gamma$, there is a constant $\widetilde{C} = \widetilde{C}_{\omega,\beta,\epsilon} > 0$ such that*

$$\sup_{t\in\mathbb{R}} |\left\langle \delta_n, e^{-itH_\omega}\delta_m \right\rangle| \leq \widetilde{C} e^{\epsilon|m|} e^{-\beta|n-m|}$$

for all $m, n \in \mathbb{Z}$.

For the remainder of the present section, we fix $\omega \in \Omega_*$ and leave the dependence of various quantities on ω implicit. Let $\{u_\ell : \ell \in \mathbb{Z}\}$ denote an enumeration of the normalized eigenvectors of H_ω. By Theorem 6.2, each u_ℓ satisfies (6.7) for a suitable choice of localization center $\zeta_\ell \in \mathbb{Z}$. For the remainder of this section, fix an arbitrary $\beta \in (0, \gamma)$. We need the following proposition, which concerns the distribution of centers of localization.

Proposition 6.5. *There exists an $L_0 \in \mathbb{Z}_+$ large enough that*

$$\#\{\ell : |\zeta_\ell| \le L\} \le L^2$$

for all $L \ge L_0$.

Proof. Let $\mathcal{U}_L := \{\ell \in \mathbb{Z} : |\zeta_\ell| \le L\}$. When L is sufficiently large, (6.7) yields

$$|u_\ell(n)| \le e^{2\beta L} e^{-\beta|n|} \le e^{-\frac{1}{2}\beta|n|} \qquad \text{whenever } \ell \in \mathcal{U}_L \text{ and } |n| \ge 4L.$$

This in turn implies that

$$\sqrt{\sum_{|n|\ge 4L} |u_\ell(n)|^2} \le e^{-\beta L} \tag{6.21}$$

whenever L is sufficiently large. Let $u_{\ell,L} \in \mathbb{R}^{8L+1}$ be $P_{[-4L,4L]}u_\ell$. Using (6.21) and the fact that $\{u_\ell : \ell \in \mathcal{U}_L\}$ is an orthonormal set, a direct computation shows that

$$|\langle u_{\ell,L}, u_{\ell',L}\rangle| \begin{cases} \ge 1 - e^{-\beta L}, & \ell = \ell', \\ \le 3e^{-\beta L}, & \ell \ne \ell', \end{cases} \tag{6.22}$$

whenever L is sufficiently large. Now consider the Gram matrix associated with $\{u_{\ell,L} : \ell \in \mathcal{U}_L\}$, that is, the matrix M having entries

$$M_{\ell,\ell'} = \langle u_{\ell,L}, u_{\ell',L}\rangle, \qquad \ell, \ell' \in \mathcal{U}_L.$$

For sufficiently large L, (6.22) implies that M is strictly diagonally dominant and hence invertible. In particular, $\{u_{\ell,L} : \ell \in \mathcal{U}_L\}$ is a linearly independent set. Since these vectors are elements of $\mathbb{R}^{8L+1}$, it follows that $\#\mathcal{U}_L \le 8L+1$. Since $8L+1 < L^2$ whenever $L \ge 9$, the proposition follows. □

Now we are ready to prove Theorem 6.4. In fact, it is well known that SULE-type conditions imply almost-sure exponential dynamical localization. We supply the details to keep the paper self-contained, following the argument in [20, Theorem 7.5].

Proof of Theorem 6.4. Given β and ϵ, choose β' with $\beta < \beta' < \gamma$ and $\beta' - \beta =: \eta < \epsilon$.

Expanding δ_m in the basis of eigenfunctions of H_ω, we obtain

$$|\langle \delta_n, e^{-itH_\omega}\delta_m\rangle| \le \sum_{\ell\in\mathbb{Z}} |u_\ell(n)u_\ell(m)|.$$

Then, using (6.7), we get

$$\sum_\ell |u_\ell(n)u_\ell(m)| \le C_\omega^2 \sum_\ell e^{C_2 \log^{22}(|\zeta_\ell|+1)} e^{-\beta'(|n-\zeta_\ell|+|m-\zeta_\ell|)}.$$

Next, we have

$$\begin{aligned} e^{-\beta'(|n-\zeta_\ell|+|m-\zeta_\ell|)} &\le e^{-\eta|\zeta_\ell|} e^{\eta|m|} e^{-(\beta'-\eta)|m-n|} \\ &\le e^{-\eta|\zeta_\ell|} e^{\epsilon|m|} e^{-\beta|m-n|} \end{aligned}$$

by the triangle inequality. Consequently,

$$\sup_{t\in\mathbb{R}} \left|\langle \delta_n, e^{-itH_\omega}\delta_0\rangle\right| \leq C_\omega^2 \sum_\ell e^{C_2 \log^{22}(|\zeta_\ell|+1)} e^{-\eta|\zeta_\ell|} e^{\epsilon|m|} e^{-\beta|n-m|}.$$

But then, by Proposition 6.5, we have

$$A_0 := \sum_\ell e^{C_2 \log^{22}(|\zeta_\ell|+1)-\eta|\zeta_\ell|} = \sum_{L\geq 0} \sum_{|\zeta_\ell|=L} e^{C_2 \log^{22}(|\zeta_\ell|+1)-\eta|\zeta_\ell|} < \infty.$$

Thus, we obtain

$$\sup_{t\in\mathbb{R}} |\langle \delta_n, e^{-itH_\omega}\delta_0\rangle| \leq \widetilde{C} e^{\epsilon|m|} e^{-\beta|n-m|}$$

with $\widetilde{C} = A_0 C_\omega^2$. □

7. Localization for CMV matrices with random Verblunsky coefficients

In the present section, we will describe how to prove spectral and dynamical localization for CMV matrices with i.i.d. random Verblunsky coefficients. The overall outline of the proof is identical to the Schrödinger case. We will describe carefully the places where the proofs differ.

Let $\mathbb{D} \subset \mathbb{C}$ be the open unit disk. Suppose that our probability space $(\mathcal{A}, \widetilde{\mu})$ consists of a compact set of complex numbers in $\mathbb{D}$; that is, we assume henceforth that

$$\mathcal{A} = \operatorname{supp} \widetilde{\mu} \subset \mathbb{D},$$

and that $\mathcal{A}$ is compact. As in the Schrödinger case, we assume that $\#\mathcal{A} \geq 2$ to avoid trivialities. Let $(\Omega, \mu) = (\mathcal{A}^{\mathbb{Z}}, \widetilde{\mu}^{\mathbb{Z}})$ and $T : \Omega \to \Omega$ denote the left shift. As before, the function $\alpha : \Omega \to \mathbb{D}$ given by $\alpha(\omega) := \omega_0$ can be used to generate Verblunsky coefficients via

$$\alpha_\omega(n) := \alpha(T^n\omega) = \omega_n, \qquad n \in \mathbb{Z},$$

and we can (and do) view $\alpha_\omega = \{\alpha_\omega(n)\}_{n\in\mathbb{Z}}$ as a sequence of i.i.d. random variables on $\mathbb{D}$ with common distribution $\widetilde{\mu}$. For each $\omega \in \Omega$, let $\mathcal{C}_\omega$ denote the CMV matrix associated with the Verblunsky coefficients $\{\alpha_\omega(n)\}_{n=0}^\infty$, and let $\mathcal{E}_\omega$ be the extended CMV matrix associated with the coefficient sequence α_ω. Our immediate goal is to prove Theorems 1.3 and 1.4. We will discuss the proof of Theorem 1.3 in detail and then comment at the end of the section on the necessary changes for the proof of Theorem 1.4.

Theorem 1.3 will be a corollary of the following theorem from which we may see the difference between Schrödinger operators and CMV matrices. In short, there is an exceptional set $\mathcal{D} \subset \partial\mathbb{D}$ containing no more than three spectral parameters at which the hypotheses of Fürstenberg's theorem may fail; thus, we work on compact arcs away from $\mathcal{D}$, and we may exhaust $\partial\mathbb{D} \setminus \mathcal{D}$ by countably many such arcs.

In the theorems below, $L(z)$ denotes the Lyapunov exponent for the operator family $\{\mathcal{E}_\omega\}$, which shall be defined presently.

Theorem 7.1 (Anderson localization for CMV matrices)**.** *With α_ω as above, there exists a finite set $\mathcal{D}$ with $\#\mathcal{D} \leq 3$ such that the following holds true. For any compact interval $\mathcal{I} \subset \partial\mathbb{D} \setminus \mathcal{D}$, there is a full-measure set $\Omega_\mathcal{I} \subset \Omega$ such that $\mathcal{E}_\omega$ has pure point spectrum on $\mathcal{I}$ for every $\omega \in \Omega_\mathcal{I}$, and the eigenfunctions of $\mathcal{E}_\omega$ corresponding to any eigenvalue $z \in \mathcal{I}$ decay exponentially. Moreover, the rate of decay is exactly $L(z)$.*

Using the same two-pass approach as in Section 6, we can prove a CMV version of SULE.

Theorem 7.2 (SULE for CMV matrices)**.** *Let $\mathcal{I}$ and $\Omega_{\mathcal{I}}$ be as in Theorem* 7.1*, and suppose that $\omega \in \Omega_{\mathcal{I}}$. For every $\delta > 0$, there exist constants $C_\delta, C_{\omega,\delta}$ such that, for every eigenfunction of u of $\mathcal{E}_\omega$ having eigenvalue $z \in \mathcal{I}$, one has*

$$|u(\bar{\zeta}+n)| \le C_{\omega,\delta}\|u\|_\infty e^{C_\delta \log^{22}(|\bar{\zeta}|+1)} e^{-(1-\delta)L(z)|n|}$$

for all $n \in \mathbb{Z}$ and some $\bar{\zeta} = \bar{\zeta}(u)$.

Following the same arguments that led to Theorem 6.4, Theorem 7.2 implies a version of dynamical localization for CMV matrices.

Theorem 7.3 (Dynamical localization for CMV matrices)**.** *Let $\mathcal{I}$ and $\Omega_{\mathcal{I}}$ be as in Theorem* 7.1*. For any $\omega \in \Omega_{\mathcal{I}}$, $\epsilon > 0$, and any $\beta > 0$ with*

$$\beta < \gamma(\mathcal{I}) := \min_{z\in\mathcal{I}} L(z),$$

there is a constant $\widetilde{C} = \widetilde{C}_{\omega,\beta,\epsilon} > 0$ such that

$$\sup_{k\in\mathbb{Z}} \left|\left\langle \delta_m, \mathcal{E}_\omega^k P_{\mathcal{I},\omega}\delta_n\right\rangle\right| \le \widetilde{C} e^{\epsilon|m|} e^{-\beta|n-m|},$$

for all $m, n \in \mathbb{Z}$, where $P_{\mathcal{I},\omega}$ denotes the spectral projection of $\mathcal{E}_\omega$ to the interval $\mathcal{I}$.

Assume for the moment that Theorem 7.1 holds true.

Proof of Theorem 1.3. For $z \in \partial\mathbb{D}$, define

$$\Omega_z := \{\omega \in \Omega : z \text{ is not an eigenvalue of } \mathcal{E}_\omega\}.$$

By a standard argument, $\mu(\Omega_z) = 1$ for each $z \in \partial\mathbb{D}$ (for example, the arguments of [44] can easily be modified to the CMV setting). For each integer $n \ge 2$, let Ω_n be the full-measure set obtained from Theorem 7.1 for

$$\mathcal{I}_n := \partial\mathbb{D} \setminus \left(\bigcup_{z\in\mathcal{D}} \left\{ ze^{i\theta} : -1/n < \theta < 1/n \right\} \right).$$

Then take

$$\Omega_* := \left(\bigcap_{n\ge 2} \Omega_n \right) \cap \left(\bigcap_{z\in\mathcal{D}} \Omega_z \right).$$

Clearly, $\mu(\Omega_*) = 1$. Moreover, for each $\omega \in \Omega_*$, $\mathcal{E}_\omega$ has pure point spectrum on $\partial\mathbb{D} \setminus \mathcal{D}$, exponentially decaying eigenfunctions for all eigenvalues $z \in \partial\mathbb{D} \setminus \mathcal{D}$, and $\mathcal{D}$ contains no eigenvalue of $\mathcal{E}_\omega$. Hence, $\mathcal{E}_\omega$ exhibits Anderson localization for each $\omega \in \Omega_*$. □

Theorems 7.1 and 7.3 hold true for $\mathcal{C}_\omega$ as well, and the proofs are nearly identical. Once one has the half-line analogue of Theorem 7.1 in hand, the proof of Theorem 1.4 is exactly the same as the proof of Theorem 1.3. We will focus on the proof of Theorem 7.1 in the remainder of the present section and point out the differences between $\mathcal{E}_\omega$ and $\mathcal{C}_\omega$ in Remark 7.9 at the end of the paper.

7.1. **Uniform positivity, continuity, and LDT of the Lyapunov exponent.** For each $\alpha \in \mathcal{A}$ and $z \in \partial\mathbb{D}$, we define the corresponding *Szegő transfer matrices* by

(7.1)
$$S^z(\alpha) := \frac{1}{\rho_\alpha}\begin{bmatrix} z & -\bar{\alpha} \\ -\alpha z & 1 \end{bmatrix}, \qquad M^z(\alpha) := z^{-1/2}S^z(\alpha) = \frac{1}{\rho_\alpha}\begin{bmatrix} \sqrt{z} & -\frac{\bar{\alpha}}{\sqrt{z}} \\ -\alpha\sqrt{z} & \frac{1}{\sqrt{z}} \end{bmatrix},$$

where $\rho_\alpha := \sqrt{1-|\alpha|^2}$. For concreteness, we choose the branch of $\sqrt{\cdot}$ defined by

$$\sqrt{e^{i\theta}} = e^{i\theta/2}, \qquad -\pi < \theta \le \pi.$$

Notice that $\|S^z\| = \|M^z\|$ for all $z \in \partial\mathbb{D}$.

We note that $M^z(\alpha) \in \mathrm{SU}(1,1)$ for every $z \in \partial\mathbb{D}$ and $\alpha \in \mathbb{D}$, where $\mathrm{SU}(1,1)$ is defined by

(7.2)
$$\mathrm{SU}(1,1) := Q \cdot \mathrm{SL}(2,\mathbb{R}) \cdot Q^* = \{A \in \mathbb{C}^{2\times 2} : A = QBQ^* \text{ for some } B \in \mathrm{SL}(2,\mathbb{R})\}$$

and

$$Q := \frac{-1}{1+i}\begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \in \mathbb{U}(2).$$

Equivalently, $\mathrm{SU}(1,1)$ consists of all 2×2 unimodular matrices that preserve the standard quadratic form of signature $(1,1)$, that is, $\mathrm{SU}(1,1) = \{A \in \mathrm{SL}(2,\mathbb{C}) : A^*JA = J\}$ with $J = \vec{e}_1\vec{e}_1^{\top} - \vec{e}_2\vec{e}_2^{\top}$. We will freely use facts about the group $\mathrm{SU}(1,1)$ throughout this section; the interested reader is referred to [51, Section 10.4] for a thorough account. For $\omega \in \Omega$, we define an $\mathrm{SU}(1,1)$-cocycle via $M^z(\omega) = M^z(\omega_0)$ and

$$M_n^z(\omega) = M^z(T^{n-1}\omega)\cdots M^z(\omega) = M^z(\omega_{n-1})\cdots M^z(\omega_0)$$

for $n \in \mathbb{Z}_+$, as before. The Lyapunov exponent of the cocycle is then given by

$$L(z) = \lim_{n\to\infty} \frac{1}{n}\int_\Omega \log\|M_n^z(\omega)\|\,\mathrm{d}\mu(\omega).$$

Our first main goal is to obtain positivity and continuity of L and use those characteristics to deduce a suitable uniform LDT. We will deduce the positivity and continuity by appealing to the machinery of Section 2 and using the fact that $\mathrm{SU}(1,1)$ is unitarily conjugate to $\mathrm{SL}(2,\mathbb{R})$, as in (7.2). The Möbius transformation induced by Q maps the upper half-plane to the unit disk and sends the real line to the unit circle. Thus, in view of (7.2), the Möbius transformation induced by any element $\mathrm{SU}(1,1)$ preserves the unit circle and unit disk just like $\mathrm{SL}(2,\mathbb{R})$ preserves the real line and the upper half-plane. Henceforth, we use $\bar{A}$ to denote the Möbius transformation induced by an element $A \in \mathrm{GL}(2,\mathbb{C})$.

Thus, to verify all the necessary conditions, we may treat $\mathrm{SU}(1,1)$ matrices just as $\mathrm{SL}(2,\mathbb{R})$ matrices. More concretely, condition (i) of Theorem 2.1 and the contraction property may be verified directly (and these properties are clearly invariant under conjugation by Q). Since $\bar{Q}$ maps the real line to the unit circle, we will say that a subgroup $G \subset \mathrm{SU}(1,1)$ is strongly irreducible if there is no finite subset $\mathcal{F} \subset \partial\mathbb{D}$ such that $\bar{B}(\mathcal{F}) = \mathcal{F}$ for all $B \in G$; in particular, $G \subset \mathrm{SU}(1,1)$ is strongly irreducible in this sense if and only if Q^*GQ is a strongly irreducible subgroup of $\mathrm{SL}(2,\mathbb{R})$. In fact, once we know that G is noncompact, we need only to verify condition (ii′) as stated in the remark following Theorem 2.1. If ν is supported

in $\mathrm{SU}(1,1)$, the appropriate version of condition (ii′) is the statement: there is no $\mathcal{F} \subset \partial\mathbb{D}$ with cardinality 1 or 2 such that $\bar{B}(\mathcal{F}) = \mathcal{F}$ for every $B \in G_\nu$.

In view of the foregoing discussion, our goal is to show that G_{ν_z} is noncompact and contracting for every $z \in \partial\mathbb{D}$, and to find a finite set $\mathcal{D} = \mathcal{D}(\mathcal{A})$ such that G_{ν_z} is strongly irreducible (as a subgroup of $\mathrm{SU}(1,1)$) for every $z \in \partial\mathbb{D} \setminus \mathcal{D}$ (here, $\nu_z = M_*^z\widetilde{\mu}$ as before). Then Theorem 2.1 ensures that $L(z) > 0$ for all $z \notin \mathcal{D}$. Moreover, since G_{ν_z} is contracting, $L(z)$ is continuous on $\partial\mathbb{D}$. Consequently, we obtain $L(z) \geq \gamma > 0$ for all z in any compact interval $\mathcal{I} \subset \partial\mathbb{D} \setminus \mathcal{D}$. From there, the strong irreducibility condition and the contraction property ensure the uniform LDT (uniform over $z \in \mathcal{I}$). Then, Hölder continuity of $L(z)$ on $\mathcal{I}$ follows from the uniform positivity and uniform LDT of $L(z)$ and the fact that $\mathrm{SU}(1,1)$ is conjugate to $\mathrm{SL}(2,\mathbb{R})$ (which ensures the applicability of the avalanche principle).

First, the following proposition follows from [51, Lemma 10.4.14].

Proposition 7.4. *If* $[M^z(\alpha), M^z(\beta)] := M^z(\alpha)M^z(\beta) - M^z(\beta)M^z(\alpha) \neq 0$, *then the subgroup generated by* $\{M^z(\alpha), M^z(\beta)\}$ *contains a nonelliptic element.*

Proof. If either $M^z(\alpha)$ or $M^z(\beta)$ is nonelliptic, there is nothing to do; otherwise, both are elliptic, in which case one may apply [51, Lemma 10.4.14] to deduce the existence of a hyperbolic element in the subgroup of $\mathrm{SU}(1,1)$ that they generate. □

We first verify condition (i) of Theorem 2.1 and the contraction property.

Proposition 7.5. *For every* $z \in \partial\mathbb{D}$, *the group* G_{ν_z} *is noncompact and contracting.*

Proof. Let $z \in \partial\mathbb{D}$ be given, denote $G = G_{\nu_z}$, and let us note that it suffices to find a nonelliptic element $A \in G$ to obtain both noncompactness and the contraction property; more specifically, if $A \in \mathrm{SU}(1,1)$ is hyperbolic or parabolic, then $\|A^n\|$ becomes unbounded as $n \to \infty$ and $\|A^n\|^{-1}A^n$ converges to a rank-1 operator. There are two cases to consider:

Case 1. $\boldsymbol{z = 1}$. Since $\mathcal{A}$ contains at least two points, choose $\alpha \neq 0$ in $\mathcal{A}$. Then

$$M^z(\alpha) = \frac{1}{\rho_\alpha}\begin{bmatrix} 1 & -\bar{\alpha} \\ -\alpha & 1 \end{bmatrix} =: A$$

satisfies $\mathrm{tr}(A) = 2\rho_\alpha^{-1} > 2$, so A is a hyperbolic element of $\mathrm{SU}(1,1)$.

Case 2. $\boldsymbol{z \neq 1}$. In this case, choose $\alpha \neq \beta$ in $\mathcal{A}$. Then one can check that

(7.3)
$$M^z(\alpha)M^z(\beta) - M^z(\beta)M^z(\alpha) = \frac{1}{\rho_\alpha\rho_\beta}\begin{bmatrix} \bar{\alpha}\beta - \alpha\bar{\beta} & (\bar{\alpha} - \bar{\beta})(1 - z^{-1}) \\ (\alpha - \beta)(1 - z) & \alpha\bar{\beta} - \bar{\alpha}\beta \end{bmatrix} \neq 0.$$

Therefore, G_{ν_z} contains a nonelliptic element by Proposition 7.4, as desired. □

Next, we note that strong irreducibility condition (ii′) essentially follows from [51, Theorem 10.4.15].

Proposition 7.6. *Suppose that* $\#\mathcal{A} \geq 2$. *Then condition* (ii′) *fails at most on a finite set* $\mathcal{D} \subset \partial\mathbb{D}$ *with* $\#\mathcal{D} \leq 3$.

From the proof of Proposition 7.6, as in [51], we see that the set $\mathcal{D}$ arises from very particular geometric degeneracies, and it is in fact empty for many choices of a compact set $\mathcal{A} \subset \mathbb{D}$. More precisely, in view of [51, Theorem 10.4.19], we see that $\mathcal{D}$ is empty as soon as the following two conditions are met:

- $\mathcal{A}$ is not contained in a single circle or line that intersects $\partial\mathbb{D}$ orthogonally.
- The set
$$\left\{ \frac{|\mathrm{Im}(\bar{\alpha}\beta)|}{|\alpha - \beta|} : \alpha \neq \beta \text{ and } \alpha, \beta \in \mathcal{A} \right\}$$
contains at least two elements.

In particular, one can simply take $\mathcal{I} = \partial\mathbb{D}$ when these conditions hold true. Note this in particular implies that a full version of exponential dynamical localization in Theorem 7.3, i.e., there is no need to add $P_{\mathcal{I}}(\mathcal{E}_\omega)$. One may find more details regarding the first condition in [51]. To meet the second condition, $\mathcal{A}$ must contain at least three noncolinear points, and, if $\#\mathcal{A} = 3$, then the incenter of the triangle[3] with vertices at the points of $\mathcal{A}$ must be distinct from 0.

7.2. **Estimating transfer matrices and Green functions.** From now on, we will focus on an arbitrary fixed compact interval $\mathcal{I} \subset \partial\mathbb{D} \setminus \mathcal{D}$, on which a uniform LDT for $L(z)$ shall hold. Following the arguments of previous sections, one can show the following analogue of Proposition 5.2. The proof is nearly identical; one need only replace $E \in \hat{\Sigma}$ by $z \in \mathcal{I}$ and make small cosmetic modifications.

Proposition 7.7. *For any $0 < \varepsilon < 1$, there exists a subset $\Omega_+ = \Omega_+(\varepsilon) \subset \Omega$ of full μ-measure such that the following statement holds true. For every $\varepsilon \in (0,1)$ and every $\omega \in \Omega_+(\varepsilon)$, there exists $\widetilde{n}_0 = \widetilde{n}_0(\omega, \varepsilon)$ such that*

$$\left| L(z) - \frac{1}{n^2} \sum_{s=0}^{n^2-1} \frac{1}{n} \log \left\| M_n^z(T^{sn+\zeta}\omega) \right\| \right| < \varepsilon \tag{7.4}$$

for every $\zeta \in \mathbb{Z}$, every $n \geq \max\{\widetilde{n}_0, \log^{\frac{2}{3}}(|\zeta| + 1)\}$, and every $z \in \mathcal{I}$.

Now we consider the results in Section 5 for finite-volume truncations of CMV matrices. We will exploit the perspective on CMV Green functions developed in [40]. Here, it will helpful to use the following factorization of $\mathcal{E}_\omega$. Writing

(7.5)
$$\Theta(\alpha) = \begin{bmatrix} \overline{\alpha} & \sqrt{1-|\alpha|^2} \\ \sqrt{1-|\alpha|^2} & -\alpha \end{bmatrix}, \quad \mathcal{L}_\omega = \bigoplus_{j \in \mathbb{Z}} \Theta(\alpha_{2j}(\omega)), \quad \mathcal{M}_\omega = \bigoplus_{j \in \mathbb{Z}} \Theta(\alpha_{2j+1}(\omega)),$$

where $\Theta(\alpha_n(\omega))$ acts on coordinates n and $n+1$, one can confirm that $\mathcal{L}_\omega$ and $\mathcal{M}_\omega$ are unitary, and that $\mathcal{E}_\omega = \mathcal{L}_\omega \mathcal{M}_\omega$. Of course, $\mathcal{E}_\omega \psi = z\psi$ if and only if $(z\mathcal{L}_\omega^* - \mathcal{M}_\omega)\psi = 0$. Given $\tau_1, \tau_2 \in \overline{\mathbb{D}}$ and an interval $\Lambda = [a, b] \subseteq \mathbb{Z}$, we define $\mathcal{E}_\omega^{\tau_1,\tau_2}$ to be the CMV matrix whose Verblunsky coefficients coincide with those of $\mathcal{E}_\omega$, except for $\alpha_{a-1} = \tau_1$ and $\alpha_b = \tau_2$. We then define

$$\mathcal{E}_{\omega,\Lambda}^{\tau_1,\tau_2} = P_\Lambda \mathcal{E}_\omega^{\tau_1,\tau_2} P_\Lambda^*.$$

One can verify that $\mathcal{E}_{\omega,\Lambda}^{\tau_1,\tau_2}$ is unitary whenever $\tau_1, \tau_2 \in \partial\mathbb{D}$. Following the convention of [40], we use "$\bullet$" to indicate that the corresponding Verblunsky coefficient is

[3]That is, the point of intersection of the angle bisectors of that triangle.

unaltered, e.g.,

$$\mathcal{E}_{\omega,\Lambda}^{\bullet,\tau_2} = \mathcal{E}_{\omega,\Lambda}^{\alpha_{a-1},\tau_2}, \qquad \mathcal{E}_{\omega,\Lambda}^{\tau_1,\bullet} = \mathcal{E}_{\omega,\Lambda}^{\tau_1,\alpha_b}.$$

As before, we abbreviate $\mathcal{E}_{\omega,N}^{\tau_1,\tau_2} = \mathcal{E}_{\omega,[0,N)}^{\tau_1,\tau_2}$. The truncations $\mathcal{L}_{\omega,\Lambda}^{\tau_1,\tau_2}$ and $\mathcal{M}_{\omega,\Lambda}^{\tau_1,\tau_2}$ with $\tau_j \in \overline{\mathbb{D}} \cup \{\bullet\}$ are defined similarly. Then we define the polynomials

$$\varphi_{\omega,\Lambda}^{\tau_1,\tau_2}(z) = \frac{1}{\prod_{k\in\Lambda}\rho_k} \det\left[z - \mathcal{E}_{\omega,\Lambda}^{\tau_1,\tau_2}\right], \qquad z \in \mathbb{C},\ \tau_j \in \overline{\mathbb{D}} \cup \{\bullet\}.$$

The associated finite-volume Green functions are defined by

$$G_{\omega,\Lambda}^{\tau_1,\tau_2}(z) = \left(z\left[\mathcal{L}_{\omega,\Lambda}^{\tau_1,\tau_2}\right]^* - \mathcal{M}_{\omega,\Lambda}^{\tau_1,\tau_2}\right)^{-1}$$

and

$$G_{\omega,\Lambda}^{\tau_1,\tau_2}(j,k;z) = \langle \delta_j, G_{\omega,\Lambda}^{\tau_1,\tau_2}(z)\delta_k\rangle, \qquad j,k \in \Lambda.$$

By [40, Proposition 3.8], for $\tau_j \in \partial\mathbb{D}$, these objects are related via

$$\left|G_{\omega,\Lambda}^{\tau_1,\tau_2}(j,k;z)\right| = \frac{1}{\rho_j\rho_k}\left|\frac{\varphi_{\omega,[a,j-1]}^{\tau_1,\bullet}(z)\varphi_{\omega,[k+1,b]}^{\bullet,\tau_2}(z)}{\varphi_{\omega,[a,b]}^{\tau_1,\tau_2}(z)}\right|, \qquad a \le j \le k \le b,$$

which furnishes the CMV analogue of (5.10). To connect Green functions and transfer matrices à la (5.11), we use [40, Corollary 3.11 and the proof of Theorem 4.4], which gives

$$\left|\varphi_{\omega,[a,j-1]}^{\tau_1,\bullet}(z)\right| \le \sqrt{2}\|S_j^z(T^a\omega)\|, \qquad \left|\varphi_{\omega,[k+1,b]}^{\bullet,\tau_2}(z)\right| \le \sqrt{2}\|S_{b-k}^z(T^{k+1}\omega)\|,$$

as well as

$$\begin{bmatrix}\varphi_{\omega,[a,b]}^{\tau_1,\tau_2}(z) & \varphi_{\omega,[a,b]}^{-\tau_1,\tau_2}(z)\\ \varphi_{\omega,[a,b]}^{\tau_1,-\tau_2}(z) & \varphi_{\omega,[a,b]}^{-\tau_1,-\tau_2}(z)\end{bmatrix} = \begin{bmatrix}z & -\overline{\tau_2}\\ z & \overline{\tau_2}\end{bmatrix} S_{b-a}^z(T^a\omega)\begin{bmatrix}1 & 1\\ \tau_1 & -\tau_1\end{bmatrix}. \tag{7.6}$$

Consequently, the minor twist here is the following: when we want to deduce good Green function estimates from largeness of the transfer matrices, we wiggle the boundary condition instead of the interval. That is, if $\|S_n^z(\omega)\|$ is large, then (7.6) implies that $\varphi_{\omega,[0,n)}^{\tau_1,\tau_2}(z)$ is large for at least one choice of $\tau_j \in \{\pm 1\}$.

Proposition 7.8. *For any $0 < \varepsilon < 1$ and $\omega \in \Omega_+(\varepsilon)$, there exists an $\widetilde{n}_1 = \widetilde{n}_1(\omega,\varepsilon)$ large enough such that the following statements hold true.*

$$\frac{1}{N}\log\|M_N^z(T^\zeta\omega)\| \le L(z) + 2\varepsilon \tag{7.7}$$

for all $\zeta \in \mathbb{Z}$, $n \ge \max\{\widetilde{n}_1, \log^2(|\zeta|+1)\}$, $z \in \mathcal{I}$. Moreover, the following holds with C_0 dependent only on $\widetilde{\mu}$:

$$\left|G_{T^\zeta\omega,N}^{\tau_1,\tau_2}(j,k;z)\right| \le C_0\frac{e^{(N-|j-k|)L(z)+C_0\varepsilon N}}{\left|\varphi_{T^\zeta\omega,[0,N)}^{\tau_1,\tau_2}(z)\right|} \tag{7.8}$$

for all $j,k \in [0,n)$, $\zeta \in \mathbb{Z}$, $n \ge \frac{1}{\varepsilon}\max\{\widetilde{n}_1, 2\log^2(|\zeta|+1)\}$, $\tau_j \in \{\pm 1\}$, $z \in \mathcal{I} \setminus \sigma(\mathcal{E}_{T^\zeta\omega,N}^{\tau_1,\tau_2})$.

The proof of Proposition 7.8 is entirely analogous to that of Corollary 5.3. The C_0 in front comes from the factor $\rho_j^{-1}\rho_k^{-1}$ and hence depends only on the support of $\widetilde{\mu}$.

7.3. **Proof of Anderson localization.** The statement and proof of the CMV analogue of Proposition 6.1 (elimination of double resonances) is almost identical to the Schrödinger operator setting, with the twist that we need to allow for four different boundary conditions. Concretely, one defines $\mathcal{D}_N(\varepsilon)$ to denote those ω such that

$$|F_m(T^{\zeta+r}\omega, z)| \leq L(z) - \varepsilon$$

and

$$\|G^{\tau_1,\tau_2}_{T^\zeta\omega,[-N_1,N_2]}(z)\| \geq e^{K^2}$$

for some choice of m, ζ, r, K, N_j as before, some $z \in \mathcal{I}$, and some choice of $\tau_j \in \{\pm 1\}$.

Once we eliminate double resonances, we are ready to prove our Theorem 7.1. We need an appropriate version of Shnol's theorem to guarantee that spectrally almost every $z \in \partial\mathbb{D}$ is a generalized eigenvalue of $\mathcal{E}$, which is supplied by [16]. So, as before, we may work with ω in a full-measure set and with ξ a linearly bounded generalized eigenfunction of $\mathcal{E}_\omega$, normalized by $\xi_0 = 1$. We note the following difference between the proofs of Claims 1 and 2 in this setting.

In the proof of Claim 1, the appropriate CMV analogue of (6.15) is the same as in the Schrödinger case. Then, by (7.6), one can choose $\tau_1, \tau_2 \in \{\pm 1\}$ such that

$$\left|\varphi^{\tau_1,\tau_2}_{T^\zeta\omega,\Lambda_i}(z)\right| \geq e^{K^3(L(z)-2\varepsilon)}.$$

Combining this with (7.8), we obtain the CMV version of (6.14), i.e.,

$$\left|G^{\tau_1,\tau_2}_{T^\zeta\omega,\Lambda_i}(j,k;z)\right| \leq \exp\big(-|j-k|L(z) + C\varepsilon K^3\big), \tag{7.9}$$

for any $j, k \in \Lambda_i$ and this same choice of τ_i, which is enough for our purposes.

Finally, in the proof of Claim 2, we need the CMV version of (6.6), which is supplied by [40, Lemma 3.9]. Concretely, if u is a solution of the difference equation $\mathcal{E}u = zu$, define

$$\widetilde{\psi}(a) = \begin{cases} (z\overline{\tau_1} - \alpha_a)\, u(a) - \rho_a u(a+1), & a \text{ is even}, \\ (z\alpha_a - \tau_1)\, u(a) + z\rho_a u(a+1), & a \text{ is odd}, \end{cases}$$

and

$$\widetilde{\psi}(b) = \begin{cases} (z\overline{\tau_2} - \alpha_b)\, u(b) - \rho_b u(b-1), & b \text{ is even}, \\ (z\alpha_b - \tau_2)\, u(b) + z\rho_{b-1} u(b-1), & b \text{ is odd}. \end{cases}$$

Then we have

$$u(n) = G^{\tau_1,\tau_2}_{[a,b]}(n,a;z)\widetilde{\psi}(a) + G^{\tau_1,\tau_2}_{[a,b]}(n,b;z)\widetilde{\psi}(b) \tag{7.10}$$

for $a < n < b$, and, of course, $\widetilde{\psi}$ is linearly bounded whenever u is. Then, combining (7.10) with (7.9) and following the proof of Claim 2, we obtain, for all k near the center of Λ_i,

$$|\xi_k| \leq Ce^{-2K^2}. \tag{7.11}$$

Now, writing $\Lambda_i = [a_i, b_i]$, take $a = \lfloor (a_1 + b_1)/2 \rfloor$, $b = \lfloor (a_2 + b_2)/2 \rfloor$, and consider $G_{T^\zeta,[a,b]}$. By the foregoing arguments, ξ_ℓ satisfies (7.11) for ℓ near or at a, b and ξ is normalized so that $\xi_0 = 1$, we obtain

$$\|G^z_{T^\zeta\omega,[a,b]}\| \geq |G^z_{T^\zeta,[a,b]}(0,\ell)| \geq ce^{K^2}$$

for some ℓ near or at a or b.

Then the remainder of the proof of Theorem 7.1 is almost identical to the corresponding arguments for the Schrödinger case. In particular, we get

$$\lim_{|n|\to\infty} \frac{1}{|n|} \log \|M_n^z(\omega)\| = L(z) > 0.$$

To relate this back to quantitative exponential decay estimates for the generalized eigenfunctions of $\mathcal{E}_\omega$, we need to use the Gesztesy–Zinchenko transfer matrices [33], not the Szegő transfer matrices. However, this is not a big deal because there is a simple connection between these matrices [16]. Thus, we conclude that the generalized eigenfunctions of $\mathcal{E}_\omega$ are exponentially decaying at $\pm\infty$ (at the rate $L(z)$).

Finally, using the CMV tools, we can make a second pass through the argument and prove Theorem 7.2 (SULE), which in turn implies Theorem 7.3 (dynamical localization).

Remark 7.9. For the proof of the half-line version of Theorems 7.1 and 7.3, we note that $\mathcal{C}_{[a,b]} = \mathcal{E}_{[a,b]}$ whenever $1 \le a \le b$; when $a = 0$, we have $\mathcal{C}_{[0,b]} = \mathcal{E}_{[0,b]}$ with the modification $\alpha_{-1} = -1$. Moreover, the Szegő transfer matrices $M_n^z(\omega)$ remain the same as long as $n \ge 0$. We can then obtain the half-line analogues of the results in Sections 7.1 and 7.2 simply by following the arguments in those sections and suitably restricting the domains of n and $[a, b]$.

The main difference is in the statement of the elimination of double resonances. Here, we need to change $G^z_{T^\varsigma\omega,[-N_1,N_2]}$ to $G^z_{T^\varsigma\omega,[0,N_2]}$ in one of the conditions. Note that $G^z_{[0,N)}$ now refers to the Green function for $z-\mathcal{C}_{[0,N)}$. After that, the remainder of the proof follows the same argument as before.

Appendix A. CMV localization from quantum walk localization

One can deduce localization for CMV matrices with i.i.d. random Verblunsky coefficients from [1]. Since they prove localization for 1D quantum walks with i.i.d. random quantum coins, let us briefly describe the CGMV connection between CMV matrices and quantum walks in 1D, named for the contributions of Cantero, Grünbaum, Moral, and Velázquez [11].

A *quantum walk* on $\mathbb{Z}$ is a unitary operator on the Hilbert space $\mathcal{H}_{\mathrm{QW}} = \ell^2(\mathbb{Z}) \otimes \mathbb{C}^2$; $\mathcal{H}_{\mathrm{QW}}$ enjoys a basis consisting of vectors of the form $\delta_n^\pm = \delta_n \otimes e_\pm$, where $e_+ = (1, 0)^\top$ and $e_- = (0, 1)^\top$ denote the usual basis of $\mathbb{C}^2$. To specify a quantum walk, choose for each n a unitary matrix

$$Q_n = \begin{bmatrix} q_n^{11} & q_n^{12} \\ q_n^{21} & q_n^{22} \end{bmatrix} \in \mathrm{U}(2, \mathbb{C}).$$

Then the quantum walk operator takes the form $\mathbf{U} = \mathbf{SQ}$, where $\mathbf{S}$ denotes the conditional shift

$$\mathbf{S}\delta_n^\pm = \delta_{n\pm1}^\pm,$$

and $\mathbf{Q}$ denotes the direct sum of the coins (viewing $\mathcal{H}_{\mathrm{QW}} \cong \bigoplus_{n\in\mathbb{Z}} \mathbb{C}^2$), that is,

$$\mathbf{Q}\delta_n^+ = q_n^{11}\delta_n^+ + q_n^{21}\delta_n^-, \qquad \mathbf{Q}\delta_n^- = q_n^{12}\delta_n^+ + q_n^{22}\delta_n^-.$$

If every Q_n lies in the set $i\,\mathrm{SU}(1, 1) = \{Q \in \mathrm{U}(1, 1) : \det(Q) = -1\}$, then $\mathbf{U}$ is a CMV matrix. Concretely, if $\Lambda : \mathcal{H}_{\mathrm{QW}} \to \ell^2(\mathbb{Z})$ is the unitary operator defined

by $\Lambda\delta_n^+ = \delta_{2n-1}$ and $\Lambda\delta_n^- = \delta_{2n}$, then $\mathbf{\Lambda U \Lambda}^*$ is a CMV matrix with Verblunsky coefficients

$$\alpha_{2j} = 0, \qquad \alpha_{2j-1} = \overline{q_n^{21}}.$$

More generally, any quantum walk operator is unitarily equivalent to a CMV operator after a suitable gauge transformation, but we do not need this more general fact [11].

Thus, the work of Ahlbrecht, Scholz, and Werner [1] gives localization for CMV matrices for which $\alpha_{2j} \equiv 0$ and $\{\alpha_{2j-1}\}_{j\in\mathbb{Z}}$ is a sequence of i.i.d. random variables in $\mathbb{D}$. It remains to be seen how one can then deduce localization for random CMV matrices without zeros interlaced. This is a standard construction (known as *sieving*), which we now describe.

Given a CMV matrix $\mathcal{E}$ with Verblunsky coefficients $\{\alpha_n\}_{n\in\mathbb{Z}}$, consider $\widetilde{\mathcal{E}}$, the CMV matrix with coefficients

$$\widetilde{\alpha}_{2j} = 0, \quad \widetilde{\alpha}_{2j-1} = \alpha_j, \quad j \in \mathbb{Z}.$$

Then one can verify by hand that $\widetilde{\mathcal{E}}^2 \cong \mathcal{E} \oplus \mathcal{E}^\top$. The calculation is known to experts but may not be obvious to the uninitiated, so we will sketch the outline for the reader's convenience. First, let $\widetilde{\mathcal{E}} = \widetilde{\mathcal{L}}\widetilde{\mathcal{M}}$ denote the factorization of $\widetilde{\mathcal{E}}$ as in (7.5). Then, straightforward calculations using the definitions yield

$$\widetilde{\mathcal{L}}\delta_{2j} = \delta_{2j+1}, \qquad \widetilde{\mathcal{L}}\delta_{2j+1} = \delta_{2j}$$

and

$$\widetilde{\mathcal{M}}\delta_{2j-1} = \overline{\alpha_j}\delta_{2j-1} + \rho_j\delta_{2j}, \qquad \widetilde{\mathcal{M}}\delta_{2j} = \rho_j\delta_{2j-1} - \alpha_j\delta_{2j}.$$

Therefore, one can verify that

(A.1)
$$\widetilde{\mathcal{E}}^2\delta_{4n-1} = \overline{\alpha_{2n}}\rho_{2n-1}\delta_{4n-4} - \overline{\alpha_{2n}}\alpha_{2n-1}\delta_{4n-1} + \overline{\alpha_{2n+1}}\rho_{2n}\delta_{4n} + \rho_{2n+1}\rho_{2n}\delta_{4n+3},$$

(A.2)
$$\widetilde{\mathcal{E}}^2\delta_{4n} = \rho_{2n}\rho_{2n-1}\delta_{4n-4} - \rho_{2n}\alpha_{2n-1}\delta_{4n-1} - \overline{\alpha_{2n+1}}\alpha_{2n}\delta_{4n} - \rho_{2n+1}\alpha_{2n}\delta_{4n+3},$$

(A.3)
$$\widetilde{\mathcal{E}}^2\delta_{4n+1} = \overline{\alpha_{2n+1}}\rho_{2n}\delta_{4n-2} - \overline{\alpha_{2n+1}}\alpha_{2n}\delta_{4n+1} + \overline{\alpha_{2n+2}}\rho_{2n+1}\delta_{4n+2} + \rho_{2n+2}\rho_{2n+1}\delta_{4n+5},$$

(A.4)
$$\widetilde{\mathcal{E}}^2\delta_{4n+2} = \rho_{2n+1}\rho_{2n}\delta_{4n-2} - \rho_{2n+1}\alpha_{2n}\delta_{4n+1} - \overline{\alpha_{2n+2}}\alpha_{2n+1}\delta_{4n+2} - \rho_{2n+2}\alpha_{2n+1}\delta_{4n+5}.$$

Defining subspaces

$$\mathcal{X} = \ell^2(\{k \in \mathbb{Z} : k \equiv 0 \mod 4 \text{ or } k \equiv 3 \mod 4\}),$$
$$\mathcal{Y} = \ell^2(\{k \in \mathbb{Z} : k \equiv 1 \mod 4 \text{ or } k \equiv 2 \mod 4\}),$$

the calculations in equations (A.1)–(A.4) show that $\widetilde{\mathcal{E}}^2$ leaves $\mathcal{X}$ and $\mathcal{Y}$ invariant, and that $\widetilde{\mathcal{E}}^2|_{\mathcal{X}} \cong \mathcal{E}^\top$ and $\widetilde{\mathcal{E}}^2|_{\mathcal{Y}} \cong \mathcal{E}$. In particular, spectral localization for $\widetilde{\mathcal{E}}^2$ implies spectral localization for $\mathcal{E}$.

Acknowledgments

We are grateful to Anton Gorodetski and Günter Stolz for their useful input.

References

[1] A. Ahlbrecht, V. B. Scholz, and A. H. Werner, *Disordered quantum walks in one lattice dimension*, J. Math. Phys. **52** (2011), no. 10, 102201, 48, DOI 10.1063/1.3643768. MR2894584

[2] M. Aizenman, *Localization at weak disorder: Some elementary bounds*, Rev. Math. Phys. **6** (1994), no. 5A, 1163–1182, DOI 10.1142/S0129055X94000419. MR1301371

[3] M. Aizenman and S. Molchanov, *Localization at large disorder and at extreme energies: An elementary derivation*, Comm. Math. Phys. **157** (1993), no. 2, 245–278. MR1244867

[4] M. Aizenman, J. H. Schenker, R. M. Friedrich, and D. Hundertmark, *Finite-volume fractional-moment criteria for Anderson localization*, Comm. Math. Phys. **224** (2001), no. 1, 219–253, DOI 10.1007/s002200100441. MR1868998

[5] J. Avron and B. Simon, *Almost periodic Schrödinger operators. II. The integrated density of states*, Duke Math. J. **50** (1983), no. 1, 369–391, DOI 10.1215/S0012-7094-83-05016-0. MR700145

[6] I. Binder, M. Goldstein, and M. Voda, *On fluctuations and localization length for the Anderson model on a strip*, J. Spectr. Theory **5** (2015), no. 1, 193–225, DOI 10.4171/JST/94. MR3340179

[7] I. Binder, M. Goldstein, and M. Voda, *On the sum of the non-negative Lyapunov exponents for some cocycles related to the Anderson model*, Ergodic Theory Dynam. Systems **37** (2017), no. 2, 369–388, DOI 10.1017/etds.2015.59. MR3614029

[8] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, Progress in Probability and Statistics, vol. 8, Birkhäuser Boston, Inc., Boston, MA, 1985. MR886674

[9] J. Bourgain and M. Goldstein, *On nonperturbative localization with quasi-periodic potential*, Ann. of Math. (2) **152** (2000), no. 3, 835–879, DOI 10.2307/2661356. MR1815703

[10] J. Bourgain and W. Schlag, *Anderson localization for Schrödinger operators on* $\mathbf{Z}$ *with strongly mixing potentials*, Comm. Math. Phys. **215** (2000), no. 1, 143–175, DOI 10.1007/PL00005538. MR1800921

[11] M.-J. Cantero, F. A. Grünbaum, L. Moral, and L. Velázquez, *Matrix-valued Szegő polynomials and quantum random walks*, Comm. Pure Appl. Math. **63** (2010), no. 4, 464–507. MR2604869

[12] R. Carmona, A. Klein, and F. Martinelli, *Anderson localization for Bernoulli and other singular potentials*, Comm. Math. Phys. **108** (1987), no. 1, 41–66. MR872140

[13] W. Craig and B. Simon, *Subharmonicity of the Lyaponov index*, Duke Math. J. **50** (1983), no. 2, 551–560, DOI 10.1215/S0012-7094-83-05025-1. MR705040

[14] D. Damanik, *A short course on one-dimensional random Schrödinger operators*, arXiv:1107.1094 (2011).

[15] D. Damanik and J. Fillman, *Spectral theory of discrete one-dimensional ergodic Schrödinger operators* (in preparation).

[16] D. Damanik, J. Fillman, M. Lukic, and W. Yessen, *Characterizations of uniform hyperbolicity and spectra of CMV matrices*, Discrete Contin. Dyn. Syst. Ser. S **9** (2016), no. 4, 1009–1023, DOI 10.3934/dcdss.2016039. MR3543643

[17] D. Damanik and A. Gorodetski, *An extension of the Kunz-Souillard approach to localization in one dimension and applications to almost-periodic Schrödinger operators*, Adv. Math. **297** (2016), 149–173, DOI 10.1016/j.aim.2016.04.006. MR3498796

[18] D. Damanik and P. Stollmann, *Multi-scale analysis implies strong dynamical localization*, Geom. Funct. Anal. **11** (2001), no. 1, 11–29, DOI 10.1007/PL00001666. MR1829640

[19] E. B. Davies and B. Simon, *Eigenvalue estimates for non-normal matrices and the zeros of random orthogonal polynomials on the unit circle*, J. Approx. Theory **141** (2006), no. 2, 189–213, DOI 10.1016/j.jat.2006.03.006. MR2252099

[20] R. del Rio, S. Jitomirskaya, Y. Last, and B. Simon, *Operators with singular continuous spectrum. IV. Hausdorff dimensions, rank one perturbations, and localization*, J. Anal. Math. **69** (1996), 153–200, DOI 10.1007/BF02787106. MR1428099

[21] F. Delyon, H. Kunz, and B. Souillard, *One-dimensional wave equations in disordered media*, J. Phys. A **16** (1983), no. 1, 25–42. MR700179

[22] J. Ding and C. K. Smart, *Localization near the edge for the Anderson Bernoulli model on the two dimensional lattice*, arXiv:1809.09041 (2018).

[23] H. von Dreifus and A. Klein, *A new proof of localization in the Anderson tight binding model*, Comm. Math. Phys. **124** (1989), no. 2, 285–299. MR1012868

[24] A. Elgart and A. Klein, *An eigensystem approach to Anderson localization*, J. Funct. Anal. **271** (2016), no. 12, 3465–3512, DOI 10.1016/j.jfa.2016.09.008. MR3558248
[25] J. Fröhlich, F. Martinelli, E. Scoppola, and T. Spencer, *Constructive proof of localization in the Anderson tight binding model*, Comm. Math. Phys. **101** (1985), no. 1, 21–46. MR814541
[26] J. Fröhlich and T. Spencer, *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*, Comm. Math. Phys. **88** (1983), no. 2, 151–184. MR696803
[27] H. Furstenberg, *Noncommuting random products*, Trans. Amer. Math. Soc. **108** (1963), 377–428, DOI 10.2307/1993589. MR0163345
[28] H. Furstenberg and Y. Kifer, *Random matrix products and measures on projective spaces*, Israel J. Math. **46** (1983), no. 1-2, 12–32, DOI 10.1007/BF02760620. MR727020
[29] F. Germinet and S. De Bièvre, *Dynamical localization for discrete and continuous random Schrödinger operators*, Comm. Math. Phys. **194** (1998), no. 2, 323–341, DOI 10.1007/s002200050360. MR1627657
[30] F. Germinet and A. Klein, *Bootstrap multiscale analysis and localization in random media*, Comm. Math. Phys. **222** (2001), no. 2, 415–448, DOI 10.1007/s002200100518. MR1859605
[31] F. Germinet and A. Klein, *A characterization of the Anderson metal-insulator transport transition*, Duke Math. J. **124** (2004), no. 2, 309–350, DOI 10.1215/S0012-7094-04-12423-6. MR2078370
[32] F. Germinet and A. Klein, *New characterizations of the region of complete localization for random Schrödinger operators*, J. Stat. Phys. **122** (2006), no. 1, 73–94, DOI 10.1007/s10955-005-8068-9. MR2203782
[33] F. Gesztesy and M. Zinchenko, *Weyl-Titchmarsh theory for CMV operators associated with orthogonal polynomials on the unit circle*, J. Approx. Theory **139** (2006), no. 1-2, 172–213, DOI 10.1016/j.jat.2005.08.002. MR2220038
[34] M. Goldstein and W. Schlag, *Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions*, Ann. of Math. (2) **154** (2001), no. 1, 155–203, DOI 10.2307/3062114. MR1847592
[35] A. Gorodetski and V. Kleptsyn, *Parametric Fürstenberg theorem on random products of* $\mathrm{SL}(2,\mathbb{R})$ *matrices*, arXiv:1809.00416 (2018).
[36] E. Hamza, A. Joye, and G. Stolz, *Localization for random unitary operators*, Lett. Math. Phys. **75** (2006), no. 3, 255–272, DOI 10.1007/s11005-005-0044-4. MR2211031
[37] S. Jitomirskaya and X. Zhu, *Large deviations of the Lyapunov exponent and localization for the 1D Anderson model*, arxiv:1803.10697 (2018); Comm. Math. Phys. (to appear).
[38] A. Joye, *Fractional moment estimates for random unitary operators*, Lett. Math. Phys. **72** (2005), no. 1, 51–64, DOI 10.1007/s11005-005-3256-8. MR2150261
[39] W. Kirsch, *An invitation to random Schrödinger operators* (English, with English and French summaries), Random Schrödinger operators, Panor. Synthèses, vol. 25, Soc. Math. France, Paris, 2008, pp. 1–119. With an appendix by Frédéric Klopp. MR2509110
[40] H. Krüger, *Orthogonal polynomials on the unit circle with Verblunsky coefficients defined by the skew-shift*, Int. Math. Res. Not. IMRN **18** (2013), 4135–4169, DOI 10.1093/imrn/rns173. MR3106885
[41] H. Kunz and B. Souillard, *Sur le spectre des opérateurs aux différences finies aléatoires* (French, with English summary), Comm. Math. Phys. **78** (1980/81), no. 2, 201–246. MR597748
[42] É. Le Page, *Théorèmes limites pour les produits de matrices aléatoires* (French), Probability measures on groups (Oberwolfach, 1981), Lecture Notes in Math., vol. 928, Springer, Berlin-New York, 1982, pp. 258–303. MR669072
[43] É. Le Page, *Régularité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes et applications* (French, with English summary), Ann. Inst. H. Poincaré Probab. Statist. **25** (1989), no. 2, 109–142. MR1001021
[44] L. A. Pastur, *Spectral properties of disordered systems in the one-body approximation*, Comm. Math. Phys. **75** (1980), no. 2, 179–196. MR582507
[45] D. Ruelle, *Ergodic theory of differentiable dynamical systems*, Inst. Hautes Études Sci. Publ. Math. **50** (1979), 27–58. MR556581
[46] I. É. Shnol, *On the behavior of eigenfunctions* (Russian), Doklady Akad. Nauk SSSR (N.S.) **94** (1954), 389–392. MR0062902

[47] C. Shubin, R. Vakilian, and T. Wolff, *Some harmonic analysis questions suggested by Anderson-Bernoulli models*, Geom. Funct. Anal. **8** (1998), no. 5, 932–964, DOI 10.1007/s000390050078. MR1650106

[48] B. Simon, *Spectrum and continuum eigenfunctions of Schrödinger operators*, J. Funct. Anal. **42** (1981), no. 3, 347–355, DOI 10.1016/0022-1236(81)90094-X. MR626449

[49] B. Simon, *Localization in general one-dimensional random systems. I. Jacobi matrices*, Comm. Math. Phys. **102** (1985), no. 2, 327–336. MR820578

[50] B. Simon, *Orthogonal polynomials on the unit circle. Part* 1*: Classical theory*, American Mathematical Society Colloquium Publications, vol. 54, American Mathematical Society, Providence, RI, 2005. MR2105088

[51] B. Simon, *Orthogonal polynomials on the unit circle. Part* 2*: Spectral theory*, American Mathematical Society Colloquium Publications, vol. 54, American Mathematical Society, Providence, RI, 2005. MR2105089

[52] 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, DOI 10.1002/cpa.3160390105. MR820340

[53] A. V. Teplyaev, *The pure point spectrum of random orthogonal polynomials on the circle* (Russian), Dokl. Akad. Nauk SSSR **320** (1991), no. 1, 49–53; English transl., Soviet Math. Dokl. **44** (1992), no. 2, 407–411. MR1151514

[54] D. Thouless, *A relation between the density of states and range of localization for one-dimensional systems*, J. Phys. C **5** (1972), 77–81.

[55] J. Tsay, *Some uniform estimates in products of random matrices*, Taiwanese J. Math. **3** (1999), no. 3, 291–302, DOI 10.11650/twjm/1500407129. MR1705982

[56] F. Wang and D. Damanik, *Anderson localization for quasi-periodic CMV matrices and quantum walks*, J. Funct. Anal. **276** (2019), no. 6, 1978–2006, DOI 10.1016/j.jfa.2018.10.016. MR3912798

[57] Z. Zhang, *Uniform positivity of the Lyapunov exponent for monotone potentials generated by the doubling map*, arXiv:1610.02137 (2016).

Department of Mathematics, United States Military Academy, West Point, New York 10996
Email address: valmir.bucaj@westpoint.edu

Department of Mathematics, Rice University, Houston, Texas 77005
Email address: damanik@rice.edu

Department of Mathematics, Virginia Tech, 225 Stanger Street—0123, Blacksburg, Virginia 24061
Email address: fillman@vt.edu

Department of Mathematics, Rice University, Houston, Texas 77005
Email address: vitaly.gerbuz@rice.edu

Department of Mathematics, Yale University, New Haven, Connecticut 06511
Email address: thomas.vandenboom@yale.edu

School of Mathematics (Zhuhai), Sun Yat-sen University, Zhuhai, Guangdong 519082, People's Republic of China
Email address: wfpouc@163.com

Department of Mathematics, University of California, Riverside, California 92521
Email address: zhenghe.zhang@ucr.edu