

RESEARCH ARTICLE | JUNE 22 2023

A “lifting” method for exponential large deviation estimates and an application to certain non-stationary 1D lattice Anderson models

Special Collection: [New Directions in Disordered Systems: In Honor of Abel Klein](#)

Omar Hurtado 



J. Math. Phys. 64, 061902 (2023)

<https://doi.org/10.1063/5.0150430>



View
Online



Export
Citation

[CrossMark](#)

Articles You May Be Interested In

Singular-unbounded random Jacobi matrices

J. Math. Phys. (August 2019)

Homotopy arguments for quantized Hall conductivity

J. Math. Phys. (August 2001)

Power law logarithmic bounds of moments for long range operators in arbitrary dimension

J. Math. Phys. (March 2023)



Journal of Mathematical Physics

**Young Researcher Award:
Recognizing the Outstanding Work
of Early Career Researchers**

[Learn More!](#)



A “lifting” method for exponential large deviation estimates and an application to certain non-stationary 1D lattice Anderson models

Cite as: J. Math. Phys. 64, 061902 (2023); doi: 10.1063/5.0150430

Submitted: 14 March 2023 • Accepted: 30 May 2023 •

Published Online: 22 June 2023



Omar Hurtado^{a)}

AFFILIATIONS

University of California, Irvine, California 92697, USA

Note: Paper published as part of the Special Topic on New Directions in Disordered Systems: In Honor of Abel Klein.

^{a)} Author to whom correspondence should be addressed: ohurtad1@uci.edu

ABSTRACT

Proofs of localization for random Schrödinger operators with sufficiently regular distribution of the potential can take advantage of the fractional moment method introduced by Aizenman–Molchanov [Commun. Math. Phys. **157**(2), 245–278 (1993)] or use the classical Wegner estimate as part of another method, e.g., the multi-scale analysis introduced by Fröhlich–Spencer [Commun. Math. Phys. **88**, 151–184 (1983)] and significantly developed by Klein and his collaborators. When the potential distribution is singular, most proofs rely crucially on exponential estimates of events corresponding to finite truncations of the operator in question; these estimates in some sense substitute for the classical Wegner estimate. We introduce a method to “lift” such estimates, which have been obtained for many stationary models, to certain closely related non-stationary models. As an application, we use this method to derive Anderson localization on the 1D lattice for certain non-stationary potentials along the lines of the non-perturbative approach developed by Jitomirskaya–Zhu [Commun. Math. Physics **370**, 311–324 (2019)] in 2019.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0150430>

I. INTRODUCTION

The study of large deviations has featured prominently in many proofs of localization for ergodic Schrödinger operators, both in the random and quasi-periodic context. For the random case, the first proof of Anderson localization for the Anderson model with singular potentials on the one-dimensional lattice in Ref. 1 by Carmona–Klein–Martinelli used large deviation estimates coming from a study of the Lyapunov exponent; later proofs in, e.g., Refs. 2–5 either introduced new approaches or made simplifications but still relied on large deviation estimates in some way.

For the two-dimensional Anderson model with singular potentials, Ding–Smart in Ref. 6 recently obtained localization at the edges of the spectrum, and a key unique continuation result relied crucially on large deviation estimates in the form of the Azuma inequality. This unique continuation enabled them to use the strategy introduced by Bourgain–Kenig in Ref. 7 and elucidated by Germinet–Klein in Ref. 8. The result in Ref. 6 has since been improved upon in Ref. 9 and was combined with a $\mathbb{Z}^5$ unique continuation result to obtain the same localization result in the three-dimensional context in Ref. 10. In the quasi-periodic context, analysis of the large deviation sets is crucial for, e.g., the method introduced by Bourgain and his co-authors in Ref. 11, detailed in Ref. 12. Recent work in this vein includes Refs. 13 and 14.

To be more explicit, when we speak of discrete Schrödinger operators in dimension d , we mean operators on $\ell^2(\mathbb{Z}^d)$ of the form

$$[H\psi](n) = \sum_{|m-n|=1} \psi(m) + V_n\psi(n). \quad (1)$$

In both the random and quasi-periodic cases, when large deviation estimates are under consideration, they are often of the form

$$\mathbb{P}[\text{A square of length } \ell \text{ is "bad"}] \leq e^{-\eta \ell} \quad (2)$$

for some $\eta > 0$ and some contextual notion of “bad.” In general, most proofs of localization require statements like the above, but with $e^{-\eta \ell}$ possibly replaced by some other $f(\ell)$ vanishing as $\ell \rightarrow \infty$. For the special case in (2) of an exponential decay, we have produced a method by which to “lift” estimates of this form from some joint distribution $\mathbb{P}_1$ of the values $\{V_n(\omega)\}$ to another distribution $\mathbb{P}_0$, which is closely related. We note that our method admits some generalization to arbitrary $f(\ell)$, but both the constraints on $\mathbb{P}_0$ and the losses in the “lifted” rate can be severe for slower rates of decay, motivating our emphasis on exponential estimates.

In particular, the method is still somewhat robust for “almost exponential” decays such as $f(\ell) = e^{-\eta \ell^q}$ (where $0 < q < 1$) and not very robust for power law decays, i.e., $f(\ell) = \ell^{-q}$ with $q > 0$. The method requires estimates on how “close” $\mathbb{P}_0$ is to $\mathbb{P}_1$ when both are restricted to finitely many coordinates, so we expect applications to be mostly (or exclusively) in the random context, where the potential at any given site has no dependence on the potential at other sites; in such cases, said estimates can be obtained in fairly natural circumstances. We note, however, that the underlying probabilistic results apply in all contexts where the requisite estimates hold.

Specifically, we let μ be a non-trivial probability distribution on $\mathbb{R}$ satisfying an appropriate moment condition. (By non-trivial, we mean μ is supported on at least two points.) When $d = 1$ and V_n are identically distributed with law μ , the transfer matrices associated with the random operator,

$$[H\psi](n) = \psi(n+1) + \psi(n-1) + V_n\psi(n),$$

satisfy large deviation estimates as a consequence Ref. 15, which should be understood in the wider context of the Furstenberg theory of random matrix products. More explicitly, we let $\Omega = \mathbb{R}^{\mathbb{Z}}$, $\mathbb{P} = \mu^{\mathbb{Z}}$, and $V_n : \Omega \rightarrow \mathbb{R}$ be projection onto the n th coordinate. Then, the family of operators

$$[H_\omega\psi](n) = \psi(n+1) + \psi(n-1) + V_n(\omega)\psi(n) \quad (3)$$

are precisely distributed in this way. This explicit description is necessary both in comparing such families to families with slightly different distribution and in the method used in Ref. 5 specifically, which we hope to adapt to our context.

For a system where V_n are independently distributed with laws of the form $g_n\mu$, with $g_n > 0$, $\int g_n d\mu = 1$, and $g_n \in L^\infty(d\mu)$, we derive the existence of similar large deviation estimates for the transfer matrices associated with the said system so long as the following key condition holds:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=-N}^N \log \|g_n\|_\infty = 0, \quad (4)$$

i.e., $\sum_{n=-N}^N \log \|g_n\|_\infty = o(N)$. Furthermore, this system will show more or less the same Lyapunov behavior as if $V_n(\omega)$ were i.i.d. with law μ . By this, we mean that the “approximate” system with potential at site n distributed with law $g_n\mu$ has a Lyapunov exponent $\tilde{\gamma}(E)$, which describes the asymptotics of the associated transfer matrices, and at all energies, this coincides with the Lyapunov exponent $\gamma(E)$ describing the transfer matrices associated with the “exact” system where the potential at each site is distributed with law μ . This probabilistic result can loosely be understood as a multiplicative version of the standard result found in, e.g., Ref. 16 that large deviation principles for stochastic processes are preserved under certain forms of super-exponential approximation.

These estimates are then used to derive almost sure Anderson localization for the associated random family of Schrödinger operators using a strategy for proving localization along the lines of Refs. 5 and 17. In the case where μ has unbounded support, we will require further technical assumptions, the most salient of which is the strengthening of (4) to

$$\lim_{N \rightarrow \infty} \frac{1}{N} \left[\sup_{k \in \mathbb{Z}} \sum_{n=k-N}^{k+N} \log \|g_n\|_\infty \right] = 0. \quad (5)$$

To our knowledge, the only work that has obtained large deviation estimates for non-stationary matrix products is from Gorodetski–Kleptsyn in Ref. 18, and the author of Ref. 19 also analyzed the asymptotics of non-stationary random matrix products but does not produce large deviation estimates. We note that the large deviation estimates found in Ref. 18 hold in an overall more general setting than ours, forgoing any kind of convergence condition on the sequence of distributions and the requirement that they all be absolutely continuous with respect to some base distribution μ .

However, in this considerably more general setting, a Lyapunov exponent cannot be expected to exist for all the non-stationary matrix products satisfying the relevant assumptions. In particular, to our knowledge, the results in Ref. 18 cannot be combined with the Craig–Simon subharmonicity result in Ref. 20, which uses subharmonicity of the Lyapunov exponent $\gamma(z)$ to achieve a certain kind of uniformity in energy.

The general large deviations they obtain consequently do not seem to be compatible with the approach of Ref. 5, which we use to derive localization. We came to know while preparing this work that the forthcoming work of Gorodetski–Kleptsyn will prove Anderson localization in the context of distributions satisfying the requirements in Ref. 18 via completely different methods using a purely dynamical approach that builds upon the approach used in Ref. 4.

We emphasize that the results in Ref. 18 are on the whole considerably more general in the context of large deviation estimates corresponding to non-stationary matrix products. Similarly, the forthcoming localization result from the work of Gorodetski–Kleptsyn will be far more general than the localization results we obtain here. There are certain rare examples for which our results are applicable and those in Ref. 18 are not, but “most” examples satisfying the hypotheses of our work satisfy the hypotheses in Ref. 18.

At the same time, the method presented there is one-dimensional in an essential way, as the transfer matrix method that allows the study of generalized eigenfunctions in terms of matrix products is only available in one-dimensional and quasi-one-dimensional contexts. In contrast, our lifting method, while imposing much harsher restrictions on the distributions considered, is fundamentally an abstract probabilistic result about product measures and is easily adapted to the study of non-stationary potentials on $\mathbb{Z}^d$ for any d or any reasonably “tame” lattice.

Moreover, while the large deviation work in Ref. 18 does not require compact support of the distribution and makes morally the same mild moment assumptions as we do in this work and as are made in, e.g., Refs. 1 and 17, it is our understanding that the forthcoming localization proof from the work of Gorodetski–Kleptsyn will further the approach of Ref. 4 and, in particular, will require the potentials $V_n(\omega)$ to have a uniform bound. We point out that our localization result holds for unbounded distributions under our more stringent assumptions.

These results are of special interest in the case where μ is singular. With sufficient regularity, the study of the problem becomes amenable to the fractional moment method pioneered in Ref. 21; such methods are able to handle non identically distributed distributions in considerably higher generality than our methods under the requisite assumptions. The methods of Ref. 22, developed earlier, are also available for the one-dimensional problem specifically. In the continuum setting, Klein obtained localization for certain non-stationary Anderson models with no regularity assumptions made in Ref. 23. This was a consequence of Wegner estimates derived more generally using the quantitative unique continuation introduced by Bourgain–Kenig in Ref. 7 and further elaborated upon by Bourgain–Klein in Ref. 24.

In addition, the requirement that the laws of $V_n(\omega)$ are all absolutely continuous with respect to some base measure μ becomes considerably more natural in the context of finite valued potentials specifically. If, e.g., $\mu = \sum_{m=1}^M \alpha_m \delta_{x_m}$, then absolute continuity implies that $\mu_n = \sum_{m=1}^M \beta_{n,m} \delta_{x_m}$ and $\|g_n\|_\infty = \max \{ \frac{\beta_{n,m}}{\alpha_m} \}$. Physically, the situation corresponds to pockets of zero natural density where the probability differs from what it should be more than any given $\varepsilon > 0$; (4) is in this context precisely equivalent to $\beta_{n,m} \xrightarrow{d} \alpha_m$ for all m , where the convergence is in the sense of natural density or, more precisely, its obvious analog on $\mathbb{Z}$.

We also prove another result under even stronger assumptions than (5) that essentially makes all the statistics of a system with $V_n(\omega)$ distributed as $g_n \mu$ identical to those of i.i.d. $V_n(\omega)$ with law μ ; specifically, if

$$\sum_{n \in \mathbb{Z}} \log \|g_n\|_\infty < \infty, \quad (6)$$

then the joint distribution $\mathbb{P}_0$ corresponding to the non-stationary case is absolutely continuous with respect to the joint distribution $\mathbb{P}_1$ with $V_n(\omega)$ i.i.d. This argument, under more stringent conditions, essentially allows a great number of results, e.g., the dynamical localization results from Ref. 5, to be carried over wholesale from stationary contexts to non-stationary contexts.

We mention in brief that our model is just one of many where the potential incorporates randomness but is not just given by i.i.d. random variables. A comprehensive review is well beyond the scope of this work; we briefly mention the random polymer models studied in, e.g., Refs. 25 and 26, which have potential purely driven by randomness but allow some “local” dependence among the variables determining the potential, mixed models studied in, e.g., Refs. 27 and 28, which consider potentials given by random terms together with terms that are either quasi-periodic or periodic, and “trimmed” models considered in, e.g., Refs. 29 and 30.

The rest of this paper is organized as follows: in Sec. II, we introduce the necessary probabilistic definitions and our results, introducing our three main probabilistic results and their consequences for localization of non-stationary Anderson models. In Sec. III, we prove our probabilistic results. In Sec. IV, we recall basic facts about Schrödinger operators and results regarding stationary Anderson models. In Sec. V, we use our probabilistic results to derive large deviation estimates and prove important consequences thereof, namely, identical Lyapunov behavior for the non-stationary approximate system and applicability of the Craig–Simon subharmonicity result for the said system. In Sec. VI, we prove technical lemmas necessary to prove our main localization result concerning unbounded potentials, Theorem II.9, and comment on the small changes necessary to prove the similar Theorem II.5, which allows weaker hypotheses in the case that the potential is bounded. Finally, in Section VII, we prove Theorem II.9.

II. PRELIMINARIES AND STATEMENTS OF RESULTS

Throughout, $\mathbb{P}_0$ and $\mathbb{P}_1$ will denote distinct probability distributions on the same measure space $(X, \mathcal{F})$. When the discussion is specified to Schrödinger operators, $(X, \mathcal{F})$ will be $\mathbb{R}^{\mathbb{Z}}$ and the product Borel algebra. In this case, $\mathbb{P}_0$ and $\mathbb{P}_1$ can also be understood as joint distributions of variables $V_n(\omega)$ so that V_n are projection maps. The expectation with respect to $\mathbb{P}_0$ and $\mathbb{P}_1$ will be denoted $\mathbb{E}_0$ and $\mathbb{E}_1$, respectively. Whenever $\mathcal{G} \subset \mathcal{F}$ are σ -algebras and X is an $L^1(\mathbb{P}_i, \mathcal{F})$ random variable, we denote its conditional expectation with respect to $\mathcal{G}$ and $\mathbb{P}_i$ by $\mathbb{E}_i[X | \mathcal{G}]$. That is, $\mathbb{E}_i[X | \mathcal{G}]$ is the unique (up to $\mathbb{P}_i$ -a.e. equivalence) $\mathcal{G}$ measurable variable, which satisfies

$$\mathbb{E}_i[\chi_A \cdot X] = \mathbb{E}_i[\chi_A \cdot \mathbb{E}_i[X | \mathcal{G}]]$$

for all $A \in \mathcal{G}$. Throughout, $\mathbb{P}_1$ can be considered an “exact” distribution and $\mathbb{P}_0$ can be considered an “approximate” distribution.

For two measures ν and μ on a measure space, $\nu \ll \mu$ denotes the absolute continuity of ν with respect to μ . When we specify the conversation to product measures, we will fix a probability measure μ on $\mathbb{R}$ and consider non-negative $g_n \in L^\infty(d\mu)$ with $g \geq 0$ and $\int g d\mu = 1$. In a natural way, these functions correspond to probability measures $\nu \ll \mu$ with essentially bounded Radon–Nikodym derivatives. In some sense, these measures are the more relevant objects, but we identify them with their densities for notational simplicity so that we may write, e.g., g_n and $g_n \mu$ instead of $\frac{d\nu_n}{d\mu}$ and μ_n .

Given $(\Omega, \mathcal{F})$ some measure space, $\mathcal{F}_n$ a filtration, and $\mathbb{P}_0, \mathbb{P}_1$ two probability measures such that $\mathbb{P}_0|_{\mathcal{F}_n} \ll \mathbb{P}_1|_{\mathcal{F}_n}$, we will define

$$H_n := \frac{d(\mathbb{P}_0|_{\mathcal{F}_n})}{d(\mathbb{P}_1|_{\mathcal{F}_n})}.$$

Sometimes, we will consider families of filtrations, i.e., collections $(\mathcal{F}_n^k)$ where for any fixed $k \in \mathbb{Z}$, $(\mathcal{F}_n^{k_0})$ is a filtration. In this case, we set

$$H_n^k := \frac{d(\mathbb{P}_0|_{\mathcal{F}_n^k})}{d(\mathbb{P}_1|_{\mathcal{F}_n^k})}.$$

We introduce notions relevant only to the study of Schrödinger operators later in the paper.

Definition II.1. We let $(X, \mathcal{M})$ be a measurable space and $(A_n^E)_{(n,E) \in \mathbb{N} \times \mathbb{R}}$ be a collection of measurable sets. (This should be understood as a collection of sequences indexed by $E \in \mathbb{R}$.) We say that the collection decays exponentially uniformly in E with respect to a probability measure $\mathbb{P}$ if there exists $N \in \mathbb{N}$ and $\eta > 0$ such that for $n > N$ and all E , we have

$$\mathbb{P}[A_n^E] \leq e^{-\eta n}.$$

Definition II.2. We say a collection of sequences of events $(A_n^E)_{(n,E)}$ (with $n \in \mathbb{N}, E \in \mathbb{R}$) is adapted to a filtration of σ -algebras $(\mathcal{F}_n)_{n \in \mathbb{N}}$ if for all E and n , we have $A_n^E \in \mathcal{F}_n$.

Note that there is nothing special about the choice of $\mathbb{R}$ for indexing our collection of sequences, save that it is what we will use later in this work, with E representing energy.

Theorem II.3. Let $(X, \mathcal{M})$ be a measurable space. Furthermore, let $\Omega = X^{\mathbb{Z}}$ and $\mathcal{B}$ be the σ -algebra on $X^{\mathbb{Z}}$ generated by measurable cylinder sets. Let μ be a probability measure on $(X, \mathcal{M})$ and g_n be a sequence of non-negative functions with $g_n \in L^\infty(d\mu)$ and $\int g_n d\mu = 1$. We define on $(\Omega, \mathcal{B})$ the product measures,

$$\mathbb{P}_0 = \bigotimes_{n \in \mathbb{Z}} g_n \mu, \quad \mathbb{P}_1 = \mu^{\mathbb{Z}},$$

and the coordinate projections V_n by $V_n(\omega) = \omega_n$ for $\Omega \ni \omega = (\omega_n)_{n \in \mathbb{Z}}$. Finally, we define the σ -algebras as

$$\mathcal{F}_n = \begin{cases} \sigma(V_{-n}, V_{1-n}, \dots, V_{n-1}, V_n), & n > 0, \\ \{\Omega, \emptyset\}, & n = 0. \end{cases}$$

If g_n satisfy

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=-N}^N \log \|g_n\|_\infty = 0, \quad (4)$$

then any collection of events (A_n^E) , which is adapted to $(\mathcal{F}_n)_{n \in \mathbb{N}}$ and exponentially decaying uniformly in E in $\mathbb{P}_1$, is also exponentially decaying uniformly in E in $\mathbb{P}_0$.

A straightforward consequence of Theorem II.3 is that large deviation estimates are valid for the joint distribution $\mathbb{P}_0$ and, as a consequence, the existence of a Lyapunov exponent, the same one existing for the joint distribution $\mathbb{P}_1$.

Theorem II.4. Let $\Omega = \mathbb{R}^{\mathbb{Z}}$ and $\mathbb{P}_1 = \mu^{\mathbb{Z}}$ for some non-trivial (i.e., supported on at least two points) μ such that there is $\alpha > 0$ for which $\int |x|^\alpha d\mu(x) < \infty$ and define the family of operators H_ω by

$$H_\omega \psi(n) = \psi(n+1) + \psi(n-1) + V_n(\omega) \psi(n).$$

Moreover, let $S_{[1,n]}^z(\omega)$ be the $SL_2(\mathbb{C})$ matrices satisfying

$$S_{[1,n]}^z(\omega) \begin{pmatrix} \psi(1) \\ \psi(0) \end{pmatrix} = \begin{pmatrix} \psi(n+1) \\ \psi(n) \end{pmatrix}$$

for any solution $\psi \in \mathbb{C}^{\mathbb{Z}}$ to $H_\omega \psi = z\psi$ and $\gamma(z) = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}_1 [\log \|S_{[1,n]}^z(\omega)\|]$. Then, for any fixed $z \in \mathbb{C}$ and $\varepsilon > 0$, there exist $\eta = \eta(z, \varepsilon) > 0$ and $N = N(z, \varepsilon)$ such that

$$\mathbb{P}_0 \left[\left| \frac{1}{n} \log \|S_{[1,n]}^z(\omega)\| - \gamma(z) \right| > \varepsilon \right] \leq e^{-\eta n}$$

for $n > N$. In particular, $\frac{1}{n} \|S_{[1,n]}^z(\omega)\| \rightarrow \gamma(z)$ $\mathbb{P}_0$ -almost surely.

These results suffice to produce the necessary estimates to show Anderson localization for the bounded case:

Theorem II.5. Let $(X, \mathcal{M})$ be $\mathbb{R}$ equipped with its Borel σ -algebra, and further let $(\Omega, \mathcal{B})$, $\mathbb{P}_0$, $\mathbb{P}_1$, $\mathcal{F}_n$, and V_n be as in Theorem II.3, with Eq. (4) satisfied.

Assume moreover that μ has compact support. Then, there is a set $\Omega_0 \subset \Omega$ with $\mathbb{P}_0[\Omega_0] = 1$ such that for all $\omega \in \Omega_0$, the operators on $\ell^2(\mathbb{Z})$ defined by

$$H_\omega \psi(n) = \psi(n-1) + \psi(n+1) + V_n(\omega) \psi(n)$$

has pure point spectrum with all eigenfunctions exponentially decaying.

However, for distributions μ with unbounded support, we need for technical reasons to be able to discuss uniformity across different choices of filtrations $(\mathcal{F}_n)^k$, which physically corresponds to uniformity across different choices of centers for an interval. Concretely, in the stationary Anderson model context, the statistics for the transfer matrix $S_{[a,b],\omega}^E$ only depend on the length $b-a+1$ and the energy E . While we cannot recover this exact statement in the non-stationary case, we can under stronger assumptions lift *a priori* large deviation estimates, which are uniform across such choices from a stationary context to similar large deviation estimates in a non-stationary context. We introduce additional definitions specific to these considerations. Throughout, “uniform in filtration” can loosely be read as “depending only on the length of the associated interval.” This is in the full level of abstraction not quite correct but describes our specific application. We introduce the following necessary definitions:

Definition II.6. Given a measure space $(X, \mathcal{M})$ and a collection of filtrations $(\mathcal{F}_n)^k$, with $n \in \mathbb{N}, k \in \mathbb{Z}$, we say that a collection of events $(A_n^{k,E})$ decays exponentially uniformly in E and k with respect to $\mathbb{P}$ if there exists N and $\eta > 0$ such that for all E , all k , and all $n > N$, we have

$$\mathbb{P}[A_n^{k,E}] < e^{-\eta n}.$$

Definition II.7. We say that $(A_n^{k,E})$ is $(\mathcal{F}_n)^k$ adapted if for any fixed $k_0 \in \mathbb{Z}$, the collection of events $A_n^{k_0,E}$ is $\mathcal{F}_n^{k_0}$ adapted.

With these natural extensions of earlier definitions, we can state a version of Theorem II.3, which is uniform in filtration.

Theorem II.8. Let $(X, \mathcal{M})$ be a measurable space. Furthermore, let $\Omega = X^{\mathbb{Z}}$ and $\mathcal{B}$ be the σ -algebra on $X^{\mathbb{Z}}$ generated by measurable cylinder sets. Let μ be a probability measure on $(X, \mathcal{M})$ and g_n be a sequence of non-negative functions with $g_n \in L^\infty(d\mu)$ and $\int g_n d\mu = 1$. We define on $(\Omega, \mathcal{B})$ the product measures

$$\mathbb{P}_0 = \bigotimes_{n \in \mathbb{Z}} g_n \mu, \quad \mathbb{P}_1 = \mu^{\mathbb{Z}}$$

and the coordinate projections V_n by $V_n(\omega) = \omega_n$ for $\Omega \ni \omega = (\omega_n)_{n \in \mathbb{Z}}$. Finally, we define the σ -algebras

$$\mathcal{F}_n^k = \begin{cases} \sigma(V_{k-n}, V_{k+1-n}, \dots, V_{k+n-1}, V_{k+n}), & n > 0, \\ \{\Omega, \emptyset\}, & n = 0. \end{cases}$$

If g_n satisfy

$$\lim_{N \rightarrow \infty} \frac{1}{N} \left[\sup_{k \in \mathbb{Z}} \sum_{n=k-N}^{k+N} \log \|g_n\|_\infty \right] = 0,$$

then any collection of events $(A_n^{E,k})$, which is adapted to $(\mathcal{F}_n)^k$ and exponentially decaying uniformly in E and k with respect to $\mathbb{P}_1$, is also exponentially decaying uniformly in E and k with respect to $\mathbb{P}_0$.

This version, uniform in filtration (i.e., center), suffices to show Anderson localization for μ_n “converging” to μ , which have unbounded support but satisfy a mild moment condition, more or less the condition found in the original Carmona–Klein–Martinelli work.¹

Theorem II.9. Let $(X, \mathcal{M})$ be $\mathbb{R}$ equipped with its Borel σ -algebra, and furthermore, let $(\Omega, \mathcal{B})$, $\mathbb{P}_0$, $\mathbb{P}_1$, $\mathcal{F}_n$, and V_n be as in Theorem II.5 with Eq. (5) satisfied.

Assume also that there are $\alpha > 0$ and $M < \infty$ such that $\int |x|^\alpha d\mu(x) < K$ and $\int |x|^\alpha g_n(x) d\mu(x) < M$ for all $n \in \mathbb{Z}$. Set $\Omega_0 \subset \Omega$ with $\mathbb{P}_0[\Omega_0] = 1$ such that for all $\omega \in \Omega_0$, the operators on $\ell^2(\mathbb{Z})$ defined by

$$H_\omega \psi(n) = \psi(n-1) + \psi(n+1) + V_n(\omega) \psi(n)$$

has the pure point spectrum with all eigenfunctions exponentially decaying, i.e., H_ω exhibits Anderson localization.

Remark II.10. Our additional conditions in the unbounded case amount to imposing uniformity in various ways; this in some sense is required to make up for the loss of uniformity, which came from the existence of a bound on V_n . In particular, (5) allows the extraction of large deviation estimates uniform in “center” of the corresponding interval or square by prohibiting arbitrarily long stretches of abnormally high $\|g_n\|$. We will explicitly go through the unbounded case through the rest of this paper and point out when they arise, the places where boundedness allows one to discard assumptions.

Finally, a condition stronger than even (5) forces essentially all the relevant statistics of $\mathbb{P}_0$ to coincide with those of $\mathbb{P}_1$, i.e., absolute continuity with an essential bound on $\frac{d\mathbb{P}_0}{d\mathbb{P}_1}$.

Theorem II.11. Let $(\Omega, \mathcal{B})$, $\mathbb{P}_0$, and $\mathbb{P}_1$ be as in Theorem II.3. If the Radon–Nikodym derivatives satisfy the stronger condition

$$\sum_{n \in \mathbb{Z}} \log \|g_n\|_\infty < \infty, \quad (6)$$

then $\mathbb{P}_0 \ll \mathbb{P}_1$, and moreover, for any $A \in \mathcal{B}$, we have $\mathbb{P}_0[A] \leq C \mathbb{P}_1[A]$, where $C := \prod_{n \in \mathbb{Z}} \|g_n\|_\infty$.

A direct application of this result to the results in Ref. 5 gives the following result:

Theorem II.12. Let everything be as in Theorem II.5; assume further that μ has compact support, and the Radon–Nikodym derivatives obey the stronger condition,

$$\sum_{n \in \mathbb{Z}} \log \|g_n\|_\infty < \infty. \quad (7)$$

Then, H_ω is almost surely exponentially dynamically localized, in the sense of Ref. 5.

III. LIFTING METHOD, PROBABILISTIC RESULTS

We prove a simple lemma, and a useful corollary, before proving Theorem II.3, which allows us to lift large deviation estimates under certain conditions.

Lemma III.1. Let $(X, \mathcal{M})$ be a measurable space and $\mathbb{P}_0, \mathbb{P}_1$ be two probability measures on it. Let $(\mathcal{F}_n)_{n \in \mathbb{N}}$ be a sequence of σ -subalgebras of $\mathcal{M}$, and further assume that $\mathbb{P}_0|_{\mathcal{F}_n} \ll \mathbb{P}_1|_{\mathcal{F}_n}$ for all $n \in \mathbb{N}$. Then, for any collection of events (A_n^E) , which decays uniformly exponentially with respect to $\mathbb{P}_1$ at rate η , we will also have uniform exponential decay with respect to $\mathbb{P}_0$ if

$$\eta_0 := \limsup_{n \rightarrow \infty} \left(\sup_E \frac{1}{n} \log \left(\|\chi_{A_n^E} H_n\|_\infty \right) \right) < \eta, \quad (8)$$

where

$$H_n := \frac{d(\mathbb{P}_0|_{\mathcal{F}_n})}{d(\mathbb{P}_1|_{\mathcal{F}_n})}$$

and $\|\cdot\|_\infty$ denotes the $L^\infty(d\mathbb{P}_1)$ norm.

Proof. Under the assumption of (8), we have as before for any $\varepsilon > 0$ some $N_0 = N_0(\varepsilon)$ such that

$$e^{-(\eta_0 + \varepsilon)n} \|\chi_{A_n^E} H_n\|_\infty \leq 1$$

for all $n > N_0$ and $\alpha \in A$. By our assumption of uniform exponential decay with respect to $\mathbb{P}_1$, there is $N_1 \in \mathbb{N}$ such that

$$\mathbb{P}_1[A_n^E] \leq e^{-\eta n}$$

for all α and $n > N_1$. Fixing $\varepsilon < \eta - \eta_0$ and $N = \max(N_0, N_1)$, we get

$$\mathbb{P}_0[A_n^E] \leq e^{-(\eta - \eta_0 - \varepsilon)n}$$

for $n > N$ and all E , establishing uniform exponential decay of the family with respect to $\mathbb{P}_0$. $\square$

Control of the Radon–Nikodym derivatives on the whole space gives a general result.

Theorem III.2. *If $(\Omega, \mathcal{M})$, $\mathcal{F}_n$, $\mathbb{P}_0$, $\mathbb{P}_1$, and H_n are as above and, moreover, we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|H_n\| = 0, \quad (9)$$

then any family adapted to $\mathcal{F}_n$ exponentially decaying uniformly in E with respect to $\mathbb{P}_1$ is exponentially decaying uniformly in E with respect to $\mathbb{P}_0$.

Proof. For any family of events, we have

$$\sup_E \|\chi_{A_n^E} H_n\|_\infty \leq \|H_n\|_\infty$$

for all n . By monotonicity of log, the required bound (8) holds as before. $\square$

Along essentially the same lines, we also have a uniformized version.

Theorem III.3. *If $(\Omega, \mathcal{M})$, $(\mathcal{F}_n)^k$, $\mathbb{P}_0$, $\mathbb{P}_1$, and H_n^k are as above and, moreover, we have*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left[\sup_{k \in \mathbb{Z}} \log \|H_n^k\|_\infty \right] = 0, \quad (10)$$

then any family adapted to $\mathcal{F}_n$ exponentially decaying uniformly in E and k with respect to $\mathbb{P}_1$ is exponentially decaying uniformly in E and k with respect to $\mathbb{P}_0$.

Remark III.4. *These results have a natural analog for other large deviation estimates. In particular, if $r(n) \geq 0$ is monotone increasing and $r(n) \rightarrow \infty$, then for $\mathbb{P}_0$ and $\mathbb{P}_1$ satisfying the obvious analog of (9)*

$$\lim_{n \rightarrow \infty} \frac{1}{r(n)} \log \|H_n\|_\infty = 0, \quad (11)$$

any adapted sequence A_n eventually satisfying $\mathbb{P}_1[A_n] \leq e^{-r(n)}$ also satisfies $\mathbb{P}_0[A_n] \leq e^{-(1-\varepsilon)r(n)}$ eventually for any $\varepsilon > 0$, and this can be made uniform over a parameter or over filtrations with the appropriate uniformity assumptions imposed on (11).

For our proof of localization, absent the uniformity in filtration coming from Theorem III.3, we can get a weaker form of uniformity by examining arithmetic progressions within a filtration.

Corollary III.5. *If $(\Omega, \mathcal{M})$, $\mathcal{F}_n$, $\mathbb{P}_0$, and $\mathbb{P}_1$ are as above and (9) holds, any family of events $\{A_n^E\}$, which is adapted with respect to a subfiltration $\{\mathcal{G}_n\}$ of the form $\mathcal{G}_n = \mathcal{F}_{kn+l}$, uniformly exponentially decays with respect to $\mathbb{P}_0$ if it does the same for $\mathbb{P}_1$.*

Proof. It suffices to show that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|H_{kn+l}\| = 0.$$

For large enough n , we have

$$\frac{1}{n} \log \|H_{kn+l}\| \leq \frac{2k}{kn+l} \log \|H_{kn+l}\|$$

with the right-hand side term going to zero as a consequence of (9). $\square$

We will later extract uniformity over finitely many (sub-)filtrations in the context of Schrödinger operators. Our work thus far now suffices to prove Theorem II.3, after recalling some definitions and a fundamental probabilistic result. Its uniformized (in filtration) variant Theorem II.8 will follow along the same lines.

Definition III.6. *If Ω is a set, we call $\mathcal{A} \subset \mathcal{P}(\Omega)$ a π -system over Ω if*

1. $\mathcal{A}$ is non-empty,
2. $\mathcal{A}$ is closed under finite intersections, i.e., for $A_1, A_2 \in \mathcal{A}$, we have $A_1 \cap A_2 \in \mathcal{A}$.

Definition III.7. *If Ω is a set, we call $\mathcal{X} \subset \mathcal{P}(\Omega)$ a λ -system if*

1. $\Omega \in \mathcal{L}$,
2. $\mathcal{L}$ is closed under complementation; for any $A \in \mathcal{L}$, we have $A^C \in \mathcal{L}$,
3. $\mathcal{L}$ is closed under disjoint countable unions; if $(A_n)_{n \in \mathbb{N}}$ are pairwise disjoint and all in $\mathcal{L}$, then $\cup A_n \in \mathcal{L}$.

A result of Dynkin, found in, e.g., Ref. 31, allows us to prove equality of measures in terms of these systems.

Theorem III.8 (Dynkin). *If $\mathcal{A}$ is a π -system contained in some λ -system $\mathcal{L}$, then $\sigma(\mathcal{A})$ is also contained in $\mathcal{L}$, where $\sigma(\mathcal{A})$ is the σ -algebra generated by $\mathcal{A}$.*

We prove one more lemma, which, combined with Theorem III.2, implies Theorem II.3.

Lemma III.9. *In the setting of Theorem II.3, we have for $N > 0$,*

$$\|H_N\|_\infty = \prod_{n=-N}^N \|g_n\|_\infty.$$

Proof. We prove this by showing that, in fact, we have the pointwise (almost surely with respect to $\mathbb{P}_1$) equality,

$$H_N(\omega) = \prod_{n=-N}^N g_n(V_n(\omega)).$$

The right-hand side of this clearly has norm less than or equal to $\prod_{n=-N}^N \|g_n\|_\infty$ by submultiplicativity of $\|\cdot\|_\infty$ and can be shown to have norm at least that more or less as a direct consequence of the product measure structure. This equality of functions (pointwise a.s.) will follow from showing equality of their integrals on any $\mathcal{F}_n$ measurable set. We define

$$\tilde{H}_N := H_N(\omega) - \prod_{n=-N}^N g_n(V_n(\omega))$$

so that the equality of their integrals is the same as vanishing of the integral of $\tilde{H}_N$.

We let $\mathcal{Z}_N$ be the family of sets $A \in \mathcal{F}_N$ such that

$$\int_A \tilde{H}_N(\omega) d\mathbb{P}_1 = 0.$$

These families are closed under countable disjoint union as a consequence of the dominated convergence theorem. (Note that $\mathbb{E}_1[|H_N|] \leq 2$.) If we assume that $\Omega \in \mathcal{Z}_N$, we obtain

$$\int_A \tilde{H}_N(\omega) d\mathbb{P}_1 = - \int_{A^C} \tilde{H}_N(\omega) d\mathbb{P}_1.$$

If we can show $\Omega \in \mathcal{Z}_N$, we will also have closure under complementation and, hence, that $\mathcal{Z}_N$ are λ -systems.

We let $\mathcal{A}_N$ denote the family of $\mathcal{F}_N$ measurable cylinder sets, i.e., $A = \prod_{n \in \mathbb{Z}} A_n$ with $A_n \in \mathcal{M}$. It is clear that $\mathcal{F}_N$ measurability requires that $A_n \in \{\emptyset, X\}$ for $n > N$ or $n < -N$. We will show $\mathcal{A}_N \subset \mathcal{Z}_N$. The result is trivial if $A_n = \emptyset$ for any n , then the set has measure zero; we thus assume that $A_n = X$ for $n > N$ or $n < -N$. The result then follows by a computation using the fact that $\mathbb{P}_1$ is a product measure,

$$\begin{aligned} \mathbb{E}_1[\chi_A H_N(\omega)] &= \mathbb{P}_0[A] \\ &= \prod_{n=-N}^N \mathbb{E}_1[\chi_{A_n} g_n(V_n(\omega))]. \end{aligned}$$

We note that $\Omega \in \mathcal{A}_N$ for all N , as clearly $\Omega = \prod_{n \in \mathbb{Z}} X$; hence, $\Omega \in \mathcal{Z}_N$, and $\mathcal{Z}_N$ are λ -systems. Moreover, $\mathcal{A}_N$ are π -systems; clearly,

$$\prod_{m \in \mathbb{Z}} A_m \cap \prod_{m \in \mathbb{Z}} B_m = \prod_{m \in \mathbb{Z}} (A_m \cap B_m)$$

and measurability is preserved. Hence, $\sigma(\mathcal{A}_N) \subset \mathcal{Z}_N$ by the π - λ theorem. However, $\sigma(\mathcal{A}_N)$ is precisely $\mathcal{F}_N$; the pre-images of the projections are cylinder sets, giving us the desired equality. $\square$

Using a similar argument as in the previous lemma, we can prove an abstract result, which has Theorem II.11 as an immediate consequence. First, we recall a probabilistic result of Doob's, whence we will obtain the existence of $H_\infty = \lim_{N \rightarrow \infty} H_N$. The result can be found in, e.g., Ref. 31.

Theorem III.10 (Doob). *Let Y_n be a martingale on a probability space $(X, \mathcal{F}, P)$ with respect to a filtration $\mathcal{F}_n$ such that $Y_n \geq 0$, and also let $\mathcal{F}_\infty = \sigma(\mathcal{F}_1, \dots)$. Then, there is Y_∞ , which is $\mathcal{F}_\infty$ measurable such that $Y_n \rightarrow Y_\infty$ almost surely.*

Using this result, we can prove the following lemma:

Lemma III.11. *Let $(\Omega, \mathcal{F})$, $\mathbb{P}_0$, $\mathbb{P}_1$, and $(\mathcal{F}_n)$ be as in the previous lemma and corollary, and let $\mathcal{F}_\infty = \sigma(\mathcal{F}_1, \mathcal{F}_2, \dots)$. If*

$$C := \sup_n \log \|H_n\| < \infty, \quad (12)$$

then $\mathbb{P}_0|_{\mathcal{F}_\infty} \ll \mathbb{P}_1|_{\mathcal{F}_\infty}$ and $\mathbb{P}_0[A] \leq e^C \cdot \mathbb{P}_1[A]$ for all $A \in \mathcal{F}$.

Proof. First, we note that even in the absence of the moment condition, the stochastic process H_n is, in fact, a martingale with respect to $(\mathcal{F}_n)$ and $\mathbb{P}_1$. Indeed, the martingale condition requires that $\mathbb{E}_1[H_{n+1} | \mathcal{F}_n] = H_n$. It follows from the definition that H_n is $\mathcal{F}_n$ measurable, and because $(\mathcal{F}_n)$ is a filtration, $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ so that for $A \in \mathcal{F}_n$, we have

$$\begin{aligned} \mathbb{E}_1[\chi_A H_n] &= \mathbb{P}_0[A] \\ &= \mathbb{E}_1[\chi_A H_{n+1}] \\ &= \mathbb{E}_1[\mathbb{E}_1[\chi_A \cdot H_{n+1} | \mathcal{F}_n]] \\ &= \mathbb{E}_1[\chi_A \cdot \mathbb{E}_1[H_{n+1} | \mathcal{F}_n]], \end{aligned}$$

with the very last equality a consequence of the $\mathcal{F}_n$ measurability of A . Hence, H_n is, in fact, $\mathbb{E}_1[H_{n+1} | \mathcal{F}_n]$.

Non-negativity of the Radon–Nikodym derivatives is obvious, so there exists (up to almost everywhere equivalence) a $\mathbb{P}_1$ almost sure limit,

$$H_\infty(\omega) := \lim_{n \rightarrow \infty} H_n(\omega),$$

which is $\mathcal{F}_\infty$ measurable; moreover, under our condition (12), satisfies $H_\infty \leq e^C$ almost surely. As a consequence of the dominated convergence theorem, $\mathbb{E}_1[H_\infty] = 1$. The theorem will follow immediately once we show that $\mathbb{P}_0[A] = \mathbb{E}_1[\chi_A \cdot H_\infty]$ for all $A \in \mathcal{F}_\infty$, i.e., H_∞ is the Radon–Nikodym derivative $\frac{d(\mathbb{P}_0|_{\mathcal{F}_\infty})}{d(\mathbb{P}_1|_{\mathcal{F}_\infty})}$. Toward this end, we use another π – λ argument. We let $\mathcal{A} = \cup_{n \in \mathbb{N}} \mathcal{F}_n$. Non-emptiness and closure under finite intersections are both obvious. We let $\mathcal{Z}$ be the collection of $A \in \mathcal{F}$ such that $\mathbb{P}_0[A] = \mathbb{E}_1[\chi_A \cdot H_\infty]$. As noted earlier, $\mathbb{E}_1[H_\infty] = 1$ by dominated convergence, and so $\Omega \in \mathcal{Z}$. Naturally then, $\mathbb{E}_1[\chi_{A^c} \cdot H_\infty] = 1 - \mathbb{E}_1[\chi_A \cdot H_\infty]$ so that $\mathcal{Z}$ is closed under complementation. Finally, using monotone convergence twice, we obtain for any countable collection of disjoint $(A_n)_{n \in \mathbb{N}} \subset \mathcal{Z}$,

$$\begin{aligned} \mathbb{P}_0[\cup A_n] &= \sum_n \mathbb{P}_0[A_n] \\ &= \sum_n \mathbb{E}_1[\chi_{A_n} \cdot H_\infty] \\ &= \mathbb{E}_1[\chi_{\cup A_n} \cdot H_\infty] \end{aligned}$$

so that $\mathcal{Z}$ is a λ -system. Consequently, $\sigma(\mathcal{A}) = \mathcal{F}_\infty$ is contained in $\mathcal{Z}$; this completes the argument. $\square$

Remark III.12. Condition (12) is a strict strengthening of (9) and relies only on the coarse data coming from the norms $\|H_n\|_\infty$. However, strictly speaking, the weaker condition

$$\mathbb{E}_1 \left[\sup_{n \in \mathbb{N}} H_n(\omega) \right] < \infty$$

suffices to give $\mathbb{P}_0 \ll \mathbb{P}_1$, as the bound (12) is only used to justify application of the dominated convergence theorem. In this more general setting, a bound of the form $\mathbb{P}_0[A] \leq C \cdot \mathbb{P}_1[A]$ for $A \in \mathcal{F}_\infty$ will hold if and only if the limit of these Radon–Nikodym derivatives H_∞ is essentially bounded. Furthermore, we note that (12) is sharp in the sense that no weaker condition formulated solely in terms of the asymptotics of $\|H_n\|_\infty$ can serve as a sufficient condition for absolute continuity.

We note that Theorem II.11 is an immediate consequence of this result together with Lemma III.9. Theorem II.12 is then immediate consequence of Ref. 5 together with Theorem II.11. While we omit the details, Ref. 17 establishing Anderson localization for Jacobi operators can also be extended to non-i.i.d. potentials in the strongly converging regime using Theorem II.11.

IV. SCHRÖDINGER OPERATOR PRELIMINARIES

Having shown some general results, we can now introduce notions relevant to the analysis of random Schrödinger operators. We also give some remarks on the general strategy of the non-perturbative approach. More or less, the entire section follows Refs. 5 and 17 closely. For an introduction to fundamental results in the theory of random Schrödinger operators, we recommend Sec. 9 of Ref. 32. Throughout this

section, we let $\Omega = \mathbb{R}^{\mathbb{Z}}$, $\mathcal{B}$ be the σ -algebra generated by cylinder sets (with respect to the Borel σ -algebra on $\mathbb{R}$), and V_n be the coordinate projections $V_n(\omega) := \omega_n$. Then, to any $\omega \in \Omega$, there is a Schrödinger operator defined on $\ell^2(\mathbb{Z})$ by

$$[H_\omega \psi](n) = \psi(n+1) + \psi(n-1) + V_n(\omega)\psi(n). \quad (3)$$

Definition IV.1. A probabilistic family of Schrödinger operators is a Borel probability measure $\mathbb{P}$ on $(\Omega, \mathcal{B})$.

Remark IV.2. In the case where $\mathbb{P}$ has support contained in $[-M, M]^{\mathbb{Z}}$ for some M , we can consider H_ω as a random variable valued in $B(\ell^2(\mathbb{Z}))$, which is the pushforward of $\mathbb{P}$ under $\omega \mapsto H_\omega$; doing this explicitly in the general (unbounded) case is unwieldy, hence our identification with the probability on the space of potentials $\mathbb{R}^{\mathbb{Z}}$. In either setting, H_ω is weakly measurable in an appropriate sense.

Throughout the rest of this paper, we freely identify a probabilistic family of Schrödinger operators with the corresponding probability distribution $\mathbb{P}$ on $(\Omega, \mathcal{B})$. Moreover, while much of what we discuss in this section holds in considerable generality, we restrict ourselves to considering two types of $\mathbb{P}$. We will consider distributions $\mathbb{P}_1$ of the form $\mathbb{P}_1 = \mu^{\mathbb{Z}}$ for some Borel measure μ on $\mathbb{R}$ and $\mathbb{P}_0$ of the form $\mathbb{P}_0 = \bigotimes_{n \in \mathbb{Z}} g_n \mu$ for some μ . $\mathbb{P}_0$ under consideration will always satisfy at least (4), if not one of the stronger conditions (5) or (6), so that we can understand $\mathbb{P}_0$ as, in fact, “close” in some sense to the $\mathbb{P}_1$ corresponding to the base distribution μ . In relation to each other, we will call $\mathbb{P}_0$ approximate and $\mathbb{P}_1$ exact.

Given fairly mild assumptions on μ , there are many results regarding localization for the exact system $\mathbb{P}_1$. In Ref. 22, it was shown that if μ was absolutely continuous with bounded density, then Anderson localization and a form of dynamical localization hold. This result was extended to hold for singular measures in Ref. 1 by Carmona–Klein–Martinelli, who found Anderson localization to hold for arbitrary non-trivial μ satisfying the moment condition

$$\int |x|^\alpha d\mu(x) < \infty \quad (13)$$

for some $\alpha > 0$. This paper used results regarding large deviations for the Lyapunov exponent, together with the multi-scale analysis developed in Ref. 33.

At least in the context of operators studied in this work, two properties are of interest, both corresponding to localization in some sense of the mass of $e^{-itH_\omega} \psi$ as t ranges over $\mathbb{R}$.

Definition IV.3. We say an operator H_ω is Anderson localized if the spectrum is entirely pure point, and its eigenfunctions are exponentially decaying.

Definition IV.4. An operator H_ω is dynamically localized if there is some A and $\mu > 0$ such that

$$\sup_{t \in \mathbb{R}} |\langle \delta_x, e^{-itH_\omega} \delta_y \rangle| \leq A e^{-\mu|x-y|}.$$

We introduce the latter definition for the sake of completeness; as we have mentioned earlier, dynamical localization can be obtained in the setting of Theorem II.12 via our probabilistic method without any further work. The rest of our paper focuses on the proof of Anderson localization in the settings of Theorems II.5 and II.9.

Definition IV.5. $E \in \mathbb{R}$ is called a generalized eigenvalue of H_ω if there exists some non-zero $\psi \in \mathbb{C}^{\mathbb{Z}}$ with $|\psi(n)|$ polynomially bounded as $|n| \rightarrow \infty$ satisfying

$$H_\omega \psi = E\psi. \quad (14)$$

Such ψ is then called a generalized eigenfunction.

The study of these suffices more or less entirely to show Anderson localization. This is a consequence of Sch’nol’s theorem, which can be found in, e.g., Ref. 32.

Theorem IV.6 (Sch’nol). If all the generalized eigenfunctions of H_ω are exponentially decaying (i.e., $|\psi(n)|$ decays exponentially as $|n| \rightarrow \infty$), then H_ω has only pure point spectrum.

We analyze the asymptotics of $|\psi(n)|$ using by working over truncations to finite boxes. We thus define for $a \leq b$ the operator $H_{[a,b],\omega}$ as the restriction of H_ω to $[a, b]$. This is $PH_\omega P$ for an appropriate choice of projection P , and we identify it with a $(b-a+1) \times (b-a+1)$ matrix.

In relation to these truncations, we can define additional quantities,

$$P_{[a,b],\omega}^E = \det(H_{[a,b],\omega} - E)$$

and

$$G_{[a,b],\omega}^E(x,y) = \langle \delta_x, (H_\omega - E)^{-1} \delta_y \rangle$$

calling the latter quantity the Green's function. Note that the Green's function is only defined for $E \notin \sigma(H_{[a,b],\omega})$ and $x, y \in [a, b]$ and that, moreover, the inverse $(H - E)^{-1}$ is the matrix inverse, not the $B(\ell^2(\mathbb{Z}))$ inverse. The importance of these quantities comes from the well-known formulas

$$\psi(x) = -G_{[a,b],\omega}^E(x,a)\psi(a-1) - G_{[a,b],\omega}^E(x,b)\psi(b+1), \quad x \in [a, b], \quad (15)$$

and

$$|G_{[a,b],\omega}^E(x,y)| = \frac{|P_{[a,x-1],\omega}^E| \cdot |P_{[y+1,b],\omega}^E|}{|P_{[a,b],\omega}^E|}, \quad x \leq y. \quad (16)$$

The first formula reduces exponential decay of generalized eigenfunctions to exponential decay of the truncated Green's functions; the second allows us to study the asymptotics of these in terms of the asymptotics of determinants. Finally, we will study these using the transfer matrices. Given any solution $\psi \in \mathbb{C}^{\mathbb{Z}}$ to (14), it satisfies for all $n \in \mathbb{Z}$,

$$\begin{pmatrix} \psi(n+1) \\ \psi(n) \end{pmatrix} = \begin{pmatrix} E - V_n(\omega) & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \psi(n) \\ \psi(n-1) \end{pmatrix} \quad (17)$$

The 2×2 matrix in (16) is called the one-step transfer matrix, and we denote it by $S_n^E(\omega)$. For $a \leq b$, we can define $S_{[a,b]}^E(\omega)$ as the unique matrix such that

$$\begin{pmatrix} \psi(a) \\ \psi(a-1) \end{pmatrix} = S_{[a,b]}^E(\omega) \begin{pmatrix} \psi(b+1) \\ \psi(b) \end{pmatrix} \quad (18)$$

The asymptotics of these matrices encode the asymptotics of generic solutions $\psi \in \mathbb{C}^{\mathbb{Z}}$ to $H_\omega \psi = E\psi$; the formula below allows us to control the asymptotics of the truncated determinants and, through those, the truncated Green's functions. Finally, we make use of the formula

$$S_{[a,b],\omega}^E = \begin{pmatrix} P_{[a,b],\omega}^E & -P_{[a+1,b],\omega}^E \\ P_{[a,b-1],\omega}^E & -P_{[a+1,b-1],\omega}^E \end{pmatrix} \quad (19)$$

to estimate the determinants $P_{[a,b],\omega}^E$ by writing them as matrix elements,

$$P_{[a,b],\omega}^E = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, S_{[a,b],\omega}^E \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle, \quad (20)$$

whence, at least in the stationary case $\mathbb{P}_1 = \mu^{\mathbb{Z}}$, the Furstenberg theory gives us information regarding the asymptotics. If

$$\int |x|^\alpha d\mu(x) < \infty \quad (13)$$

for some $\alpha > 0$, then Furstenberg's theorem and extensions thereof are applicable. In particular, the work of Furstenberg–Kesten shows that under weaker conditions than those proposed, the quantity

$$\gamma(E) := \lim_{n \rightarrow \infty} \frac{1}{n} E[\log \|S_{[1,n]}^E\|]$$

is defined, and almost surely, we have

$$\frac{1}{n} \log \|S_{[1,n]}^E(\omega)\| \rightarrow \gamma(E),$$

and the Furstenberg Theorem implies that as long as μ is non-trivial, $\gamma(E) > 0$ for all E . Le Page showed in Ref. 15 that under our conditions, we have exponential decay in the probability of large deviations, both for these quantities and for the magnitude of the corresponding matrix elements. In Ref. 34, it was found that this could be made uniform over a parameter varying over a compact set; in particular, it is a straightforward application of Ref. 34, also proved in Ref. 3.

Theorem IV.7 (Tsai, Bucaj et al.). *Fixing $I \subset \mathbb{R}$ compact and $\varepsilon > 0$, there exist $\eta > 0$ and N such that for any $u, v \in \mathbb{R}^2$ with $\|u\| = \|v\| = 1$ and any $E \in \mathbb{R}$, we have*

$$\mathbb{P}_1 \left[\left| \frac{1}{n} \log |\langle u, S_{[1,n],\omega}^E v \rangle| - \gamma(E) \right| > \varepsilon \right] < e^{-\eta n}$$

for $n > N$.

Because $\mathbb{P}_1 = \mu^{\mathbb{Z}}$, in particular, the map T defined by $(T\omega)_n = \omega_{n+1}$ is measure preserving so that the statistics of $S_{[a,b]}^E$ and those of $S_{[1,b-a+1]}^E$ are identical. Hence, exponential large deviation estimates for $P_{[a,b],\omega}^E$, uniform over a compact interval I , are a corollary of Tsay's theorem together with (20).

Corollary IV.8. Fix $I \subset \mathbb{R}$ be compact and $\varepsilon > 0$. Then, there are $\eta > 0$ and N such that

$$\mathbb{P}_1 \left[\left| \frac{1}{b-a+1} \log |P_{[a,b],\omega}^E| - \gamma(E) \right| > \varepsilon \right] < e^{-\eta(b-a+1)} \quad (21)$$

for $b-a+1 > N$.

In particular, this result gives in a crude sense,

$$|G_{[a,b],\omega}^E(x, y)| \sim e^{-\gamma(E)|x-y|}.$$

A stronger, quantitative version of this result would imply Anderson localization as a consequence of (15). This motivates the following definition:

Definition IV.9. We say $x \in \mathbb{Z}$ is (C, n, E, ω) regular if

$$|G_{[x-n, x+n], \omega}^E(x, x \pm n)| \leq e^{-Cn}$$

and x is (C, n, E, ω) -singular if it is not regular for the same set of parameters.

In particular, there is a reformulation of Anderson localization in terms of this notion. Using formula (15), Theorems II.5 and II.9 are straightforward consequences of the following theorem:

Theorem IV.10. Under the assumptions of either Theorems II.5 or II.9, there is $\Omega_0 \subset \Omega$ with $\mathbb{P}_0[\Omega_0] = 1$ such that for every $\omega \in \Omega_0$ and $E \in \mathbb{R}$, there exist $N = N(E, \omega)$ and $C = C(E)$ such that for every $n > N$, both $2n$ and $2n+1$ are (C, n, E, ω) -regular.

All the relevant estimates in this section are established for a stationary $\mathbb{P}_1$, and so Sec. V derives the analogous large deviation estimate results for an appropriate $\mathbb{P}_0$, along with important consequences.

V. ADAPTATIONS FOR NON-STATIONARITY

Throughout this section, $\mathbb{P}_0$ is of the form specified in either Theorems II.5 or II.9; we explicitly indicate when a result only applies in one context. By applying Theorem II.3 to Corollary IV.8, we can produce a $\mathbb{P}_0$ analog of (21).

Theorem V.1. Fixing $I \subset \mathbb{R}$ compact, for any $\varepsilon > 0$ and $K \in \mathbb{N}$, there are $\eta = \eta(\varepsilon, K) > 0$ and $N = N(\varepsilon, K)$ such that

$$\mathbb{P}_0 \left[\left| \frac{1}{b-a+n(j_1+j_2)+1} \log |P_{[a+j_1n, b+j_2n], \omega}^E| - \gamma(E) \right| > \varepsilon \right] < e^{-\eta(b+j_2n-a-j_1n+1)}$$

for $-K \leq a \leq b \leq K$, $-K \leq j_1 < j_2 \leq K$, $n > N$, and $E \in I$.

Proof. Fix $a, b \in \mathbb{Z}$ and $j_1, j_2 \in \mathbb{N}_0$. Let $l = \max\{|a|, |b|\}$ and $j = \max\{|j_1|, |j_2|\}$. Clearly, the events

$$\left\{ \omega : \left| \frac{1}{b-a+j_1n+j_2n+1} \log |P_{[a-j_1n, b+j_2n], \omega}^E| - \gamma(E) \right| > \varepsilon \right\}$$

are $\mathcal{F}_{jn+l}$ measurable. Moreover, they have $\mathbb{P}_1$ exponential large deviation estimates by Corollary IV.8. By Corollary II.3, there are $\mathbb{P}_0$ exponential large deviation estimates, i.e., there are $\tilde{N}(\varepsilon, a, b, j_1, j_2)$ and $\tilde{\eta}(\varepsilon, a, b, j_1, j_2)$ such that

$$\mathbb{P}_0 \left[\left| \frac{1}{b-a+(j_1+j_2)n+1} \log |P_{[a-j_1n, b+j_2n], \omega}^E| - \gamma(E) \right| > \varepsilon \right] < e^{\tilde{\eta}(b-a+j_1n+j_2n+1)}$$

for $n > \tilde{N}$. Taking η to be the minimum of $\tilde{\eta}$ for $|a|, |b|, |j_1|, |j_2|$ all smaller than K and N to be the maximum of $\tilde{N}$ ranging over the same parameters, we obtain the desired η and N . $\square$

We proved this theorem that works even without the assumption of condition (9) to illustrate what is used in the bounded case if said condition fails; however, we have something stronger when (9) holds, allowing the use of Theorem II.8.

Theorem V.2. *In the context of Theorem II.9, but not necessarily in the context of Theorem II.5, fixing $I \subset \mathbb{R}$ compact, for any $\varepsilon > 0$, there are $\eta = \eta(\varepsilon) > 0$ and $N = N(\varepsilon)$ such that for $b - a > N$,*

$$\mathbb{P}_0 \left[\left| \frac{1}{b-a+1} \log |P_{[a,b],\omega}^E| - \gamma(E) \right| \right] < e^{-\eta(b-a+1)}.$$

Proof. By stationarity of the shift map $(T\omega)_n = \omega_{n+1}$ in the $\mathbb{P}_1$ context, the large deviation estimates furnished by Ref. 34 are, in fact, uniform in the endpoints of the interval, i.e., uniform in filtration. The result is then immediate as a consequence of Theorem II.8. $\square$

For a certain application of Ref. 20, we need to consider complexified energy, i.e., transfer matrices $S_{[a,b]}^z$, which characterize the solutions to $H_\omega \psi = z\psi$ in the sense of (18). For $z \in \mathbb{C}$ generally, the non-uniform version of Theorem V.1 (i.e., for a single fixed energy $z \in \mathbb{C}$) also holds by applying Theorem II.3 to the work of Le Page in Ref. 15. (We believe that the uniform result in Ref. 34 still applies in the context of complexified energy, but we are not sure of the details and do not need it.). There is analogously $\gamma(z) = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}_1 [\log \|S_{[1,n]}^z\|]$ and

$$\frac{1}{n} \log \|S_{[1,n],\omega}^z\| \rightarrow \gamma(z)$$

$\mathbb{P}_1$ almost surely. The large deviation results imply that these $\mathbb{P}_1$ almost sure limits are also $\mathbb{P}_0$ almost sure limits.

Theorem V.3. *For $\mathbb{P}_0$ and $\mathbb{P}_1$ as above and any $z \in \mathbb{C}$, there is Ω_z such that $\mathbb{P}_0[\Omega_z] = 1$ and*

$$\frac{1}{n} \log \|S_{[1,n],\omega}^z\| \rightarrow \gamma(z).$$

Proof. We note first that as a consequence of Theorem V.1 or its non-uniform complex analog, there are $\eta > 0$ and $N \in \mathbb{N}$ such that for $n > N$, we have

$$\mathbb{P}_0 \left[\left| \frac{1}{n} \log \|S_{[1,n],\omega}^z\| - \gamma(z) \right| > \varepsilon \right] < e^{-\eta n}.$$

The eventual exponential decay implies summability so that by Borel–Cantelli, there exist for all $\varepsilon > 0$ subsets $\Omega_{z,\varepsilon} \subset \Omega$ with $\mathbb{P}_0[\Omega_{z,\varepsilon}] = 1$, and for all $\omega \in \Omega_{z,\varepsilon}$, we have some $N = N(\omega)$ such that $n > N$ implies

$$\left| \frac{1}{n} \log \|S_{[1,n],\omega}^z\| - \gamma(z) \right| < \varepsilon.$$

We note that $\frac{1}{n} \log \|S_{[1,n],\omega}^z\| \rightarrow \gamma(z)$ precisely if $\omega \in \bigcap_{m \in \mathbb{N}} \Omega_{z,1/m}$. This intersection has probability 1 and so is the requisite Ω_z . $\square$

This last result is of some interest in its own right, demonstrating an ability to “lift” Lyapunov behavior across contexts; it is also necessary for the proof of localization in making it possible to apply certain results exploiting subharmonicity of $\gamma(z)$ in Ref. 20.

We go through the details here, although the argument is fundamentally the same as in the original paper. For any fixed $z \in \mathbb{C}$, there is $\Omega_z \subset \Omega$ with $\mathbb{P}_0[\Omega_z] = 1$ such that for $\omega \in \Omega_z$, the quantity

$$\gamma^+(\omega, z) := \limsup_{n \rightarrow +\infty} \frac{\log \|S_{[1,n],\omega}^z\|}{n}$$

coincides with $\gamma(z)$. A result of Craig and Simon makes this in some sense uniform in the context of the exact Anderson model and, moreover, says the same for the quantity

$$\gamma^-(\omega, z) := \limsup_{n \rightarrow +\infty} \frac{\log \|S_{[-n,-1],\omega}^z\|}{n}.$$

(In fact, the quantities $\gamma^+(\omega, z)$ and $\gamma^-(\omega, z)$ always coincide, and so we denote this quantity going forward by $\bar{\gamma}(\omega, z)$.)

Theorem V.4 (Ref. 20). *Let μ be a distribution on $\mathbb{R}$ satisfying the condition*

$$\int \max\{0, \log |x|\} d\mu(x) < \infty$$

and $\mathbb{P}_1 = \mu^{\mathbb{Z}}$. Then, there exists $\Omega_1 \subset \Omega = \mathbb{R}^{\mathbb{Z}}$ with $\mathbb{P}_1[\Omega_1] = 1$ such that $\bar{\gamma}(\omega, E) \leq \gamma(E)$ for all $\omega \in \Omega_1$ and $E \in \mathbb{R}$.

Remark V.5. Craig and Simon originally proved the result for bounded Schrödinger operators, which in our context implies an absolute bound on $V_n(\omega)$. However, the proof straightforwardly generalizes to any family H_ω satisfying this mild moment condition, without which $\gamma(E)$ is not even guaranteed to exist.

The key step to proving this result concerning $E \in \mathbb{R}$ was a theorem regarding the Lyapunov exponent in complexified energy.

Theorem V.6 (Ref. 20). *For an exact system whose parameters satisfy the assumptions in Theorem V.4, $\gamma(z)$ is subharmonic, and for all $\omega \in \Omega$, the function $\bar{\gamma}(\omega, z)$ is submean.*

That $\bar{\gamma}(\omega, E)$ is still submean in the context of the approximate system is obvious; the change from the exact to approximate system amounts only to a change in probability measure and $\bar{\gamma}(\omega, E)$ is not an averaged quantity. On the other hand, subharmonicity of the $\mathbb{P}_0$ Lyapunov exponent $\bar{\gamma}(z)$ is only obtained by showing its equality with the $\mathbb{P}_1$ Lyapunov exponent $\gamma(z)$; we recall some basic facts from the theory of subharmonic and submean functions.

Proposition V.7. *If f is submean and E_0 is fixed, then*

$$f(E_0) \leq \lim_{r \rightarrow 0} \frac{1}{\pi r^2} \int_{|E-E_0| < r} f(E) d^2 E,$$

and if f is subharmonic and E_0 is fixed, then

$$f(E_0) = \lim_{r \rightarrow 0} \frac{1}{\pi r^2} \int_{|E-E_0| < r} f(E) d^2 E.$$

Using this, we can now prove, more or less along the lines of the original argument in the original paper of Craig and Simon, the following:

Theorem V.8. *There exists a subset $\Omega_0 \subset \Omega$ with $\mathbb{P}_0[\Omega_0] = 1$ such that for all $E \in \mathbb{R}$,*

$$\bar{\gamma}(\omega, E) \leq \gamma(E).$$

Proof. Recall that $\bar{\gamma}(\omega, z)$ is submean and $\gamma(z)$ is subharmonic. Moreover, we have shown that for any fixed $z \in \mathbb{C}$, there is a $\mathbb{P}_0$ full measure subset $\Omega_z \subset \Omega$ such that $\gamma(\omega, z) = \gamma(z)$. By Fubini, there is a $\mathbb{P}_0$ probability 1 subset Ω_0 such that for $\omega \in \Omega_0$, we have $\gamma(\omega, z) = \gamma(z)$ for a Lebesgue almost every $z \in \mathbb{C}$.

Hence, it suffices to show that for any ω such that $\bar{\gamma}(\omega, z) \leq \gamma(z)$ for Lebesgue almost all z , we, in fact, have it for all z . For such ω , we have necessarily for any fixed $E \in \mathbb{R}$ and $r > 0$,

$$\int_{|z-E| < r} \bar{\gamma}(\omega, z) d^2 z = \int_{|z-E| < r} \gamma(z) d^2 z.$$

It follows immediately that

$$\begin{aligned} \bar{\gamma}(\omega, E) &\leq \lim_{r \rightarrow 0} \frac{1}{\pi r^2} \int_{|z-E| < r} \bar{\gamma}(\omega, z) d^2 z \\ &= \lim_{r \rightarrow 0} \frac{1}{\pi r^2} \int_{|z-E| < r} \gamma(z) d^2 z \\ &= \gamma(E). \end{aligned}$$

□

We reformulate this result quantitatively and in terms of the transfer matrices.

Corollary V.9. *For $\mathbb{P}_0$ -a.s. ω and any $\varepsilon > 0$, there exists $N = N(\omega, \varepsilon)$ such that for $n > N$, we have*

$$\max \left\{ \|S_{[1,n],E,\omega}\|, \|S_{[-n,-1],E,\omega}^{-1}\| \right\} \leq e^{(\gamma(E)+\varepsilon)n}$$

and

$$\max \left\{ \|S_{[n+1,2n],E,\omega}\|, \|S_{[2n+2,3n],E,\omega}^{-1}\| \right\} \leq e^{(\gamma(E)+\varepsilon)n}.$$

VI. MAIN LEMMAS

We use now the large deviation results we have obtained to prove several technical lemmas, generally following Ref. 17, making some simplifications due to our exclusive consideration of the Schrödinger case, rather than Jacobi operators as a whole. Throughout this section,

we fix a compact interval of energies $[s, t] =: I \subset \mathbb{R}$. Of course, if we have localization almost surely with respect to any given compact interval, then (taking a countable intersection) we have almost sure localization for all energies. Continuity of $\gamma(E)$ was shown in Ref. 17, using ideas from work in Ref. 35, so that, in particular, $\inf_{E \in I} \gamma(E) > 0$. The proof proceeds by analyzing the sets where large deviations occur, so we define subsets of $I \times \Omega$,

$$B_{[a,b],\varepsilon}^+ = \left\{ (E, \omega) : |P_{[a,b],\omega}^E| \geq e^{(\gamma(E)+\varepsilon)(b-a+1)} \right\}$$

and

$$B_{[a,b],\varepsilon}^- = \left\{ (E, \omega) : |P_{[a,b],\omega}^E| \leq e^{(\gamma(E)-\varepsilon)(b-a+1)} \right\},$$

and the sections

$$B_{[a,b],E,\varepsilon}^\pm = \{ \omega : (E, \omega) \in B_{[a,b],\varepsilon}^\pm \}$$

and

$$B_{[a,b],\omega,\varepsilon}^\pm = \{ E : (E, \omega) \in B_{[a,b],\varepsilon}^\pm \}.$$

Moreover, we let $E_{j,[a,b],\omega}$ denote the $b - a + 1$ eigenvalues (with multiplicity) of $H_{[a,b],\omega}$. An immediate consequence of (16) is as follows:

Lemma VI.1. *If $n \geq 2$ and $0 < \varepsilon < v_I/8$ and x is $(\gamma(E) - 8\varepsilon, E, n, \omega)$ -singular, then*

$$E \in B_{[x-n,x+n],\varepsilon}^- \cup B_{[x-n,x],\varepsilon}^+ \cup B_{[x,x+n],\varepsilon}^+.$$

Our work in lifting LDEs and the existence of these for the exact case will allow us to prove technical lemmas like those in Refs. 5 and 17. In particular, Theorem V.1 suffices for most of these. We take $\eta_0 = \eta(\varepsilon_0, 4)$, where η is the large deviation parameter from Theorem V.1. We now proceed through some technical lemmas, proved in either Refs. 5 or 17 for $\mathbb{P}_1$, and comment on any differences arising in the proof in a non-stationary context.

Lemma VI.2. *Let $0 < \delta_0 < \eta_0$. There is $\Omega_1 \subset \Omega$ such that $\mathbb{P}_0[\Omega_1] = 1$, and for all $\omega \in \Omega_1$, there is $N_1 = N_1(\omega)$ such that for all $n > N_1$, we have*

$$\max \{ |B_{[n+1,3n+1],\varepsilon_0,\omega}^-|, |B_{[-n,n],\varepsilon_0,\omega}^-| \} \leq e^{-(\eta_0-\delta_0)(2n+1)}.$$

This is done by a Borel–Cantelli argument in Ref. 17 for $\mathbb{P}_1$, which uses large deviation estimates and carries over with no modification to our $\mathbb{P}_0$ context as a consequence of Theorem V.1.

Lemma VI.3. *For any $\varepsilon > 0$ and $p > 4/\eta_\varepsilon$, where η_ε is the large deviation parameter furnished by Theorem V.2, there is $\Omega_2 = \Omega_2(\varepsilon, p) \subset \Omega$ with full probability such that for every $\omega \in \Omega_2$, there is $N_2 = N_2(\omega)$ so that for $n > N_2$, any y_1, y_2 satisfying $-n \leq y_1 \leq y_2 \leq n$, and $| -n - y_1 | > p \log n$ and $| n - y_2 | > p \log n$, we have*

$$E_{j,[n+1,3n+1],\omega} \notin B_{[-n,y_1],\varepsilon,\omega} \cup B_{[y_2,n],\varepsilon,\omega} \cup B_{[-n,n],\varepsilon,\omega}$$

for all $j \in [1, b - a + 1]$.

Proof. Following Ref. 5, we analyze the events

$$A_n = \{ \omega : \exists j \in [1, 2n + 1]; E_{j,[n+1,3n+1],\omega} \in B_{[-n,y_1],\varepsilon,\omega} \cup B_{[y_2,n],\varepsilon,\omega} \}$$

and, leveraging independence and union bounds together with large deviation estimates coming from Theorem V.2, obtain

$$\mathbb{P}_0[A_n] \leq 2(2n + 1)^3 e^{-\eta_\varepsilon p \log n + 2} \quad (22)$$

for sufficiently large n . Because $p\eta_\varepsilon > 4$, this is summable, whence the result follows. $\square$

This result has an analog in the context of Theorem II.5, which can be proven using only Theorem V.1 rather than requiring Theorem V.2.

Lemma VI.4. *For any $\varepsilon > 0$ and $L > 1$, there is $\Omega_2 = \Omega_2(\varepsilon, L) \subset \Omega$ with full probability such that for every $\omega \in \Omega_2$, there is $N_2 = N_2(\omega)$ so that for $n > N_2$, any y_1, y_2 satisfying $-n \leq y_1 \leq y_2 \leq n$, and $| -n - y_1 | > \frac{n}{L}$ and $| n - y_2 | > \frac{n}{L}$, we have*

$$E_{j,[n+1,3n+1],\omega} \notin B_{[-n,y_1],\varepsilon,\omega} \cup B_{[y_2,n],\varepsilon,\omega} \cup B_{[-n,n],\varepsilon,\omega}$$

for all $j \in [1, b - a + 1]$.

Its proof is more or less the same as that of Lemma VI.3, but replacing $p \log n$ with $\frac{n}{L}$ gives (in the original stationary context) a stronger estimate than (22),

$$\mathbb{P}_1[A_n] \leq 2(2n + 1)^3 e^{-\tilde{\eta} \frac{n}{L} + 2}, \quad (23)$$

where $\bar{\eta}$ is the $\mathbb{P}_1$ large deviation parameter. For any $\varepsilon > 0$ and sufficiently large n , we then obtain

$$\mathbb{P}_1[A_n] \leq e^{-\left(\frac{\bar{\eta}}{L} - \varepsilon\right)n}.$$

In particular, then, instead of carrying over exponential estimates beforehand to ultimately produce the summable but subexponential estimates of (22), we derive exponential estimates of (23) in the stationary context and then carry them over using Theorem V.1, obtaining

$$\mathbb{P}_0[A_n] \leq e^{-\eta n}. \quad (24)$$

(Even Theorem V.1 is in some sense more than we need; another direct application of Theorem II.3 suffices.) This result is also true in the unbounded context of Theorem II.9, but in carrying out the localization proof for the unbounded case, Lemma VI.3 turns out to be necessary. This is because the next pair of results concerning determinants corresponding to the edge of a box is weaker than its bounded context analog.

Lemma VI.5. *For fixed $r > 1$ and $p > 0$, for almost all ω , there exists $N = N(\omega)$ such that for $n > N$ and $m \in [-n, n]$ with $| -n - m| \leq p \log n$ or $|n - m| \leq p \log n$,*

$$|V_m(\omega)| \leq n^{r/\alpha},$$

where $\alpha > 0$ is such that $\sup_n \int |x|^\alpha g_n(x) d\mu(x) < \infty$.

Proof. We set $C := \sup_n \int |x|^\alpha g_n(x) d\mu(x)$. By Chebyshev,

$$\mathbb{P}_0[V_m(\omega) \geq n^{r/\alpha}] \leq \frac{C}{n^r}$$

for any $m \in [-n, n]$. Hence, the probability that there exists some m with $|V_m(\omega)|$ exceeding $n^{r/\alpha}$ and also $| -n - m| \leq p \log n$ or $|n - m| \leq p \log n$ is bounded by (using a union bound)

$$\frac{2C}{n^r} (1 + 2p \log n),$$

which is summable, whence the result follows from the work of Borel–Cantelli. $\square$

A crucial corollary is the following:

Corollary VI.6. *Fixing $I \subset \mathbb{R}$ compact, for $p > 0$ and $r > 1$, there is a probability 1 subset $\Omega_3 = \Omega_3(p, r) \subset \Omega$ such that for $\omega \in \Omega_3$, there is $N = N(\omega) \geq 3$ such that if $n > N$ and $| -n - y| > p \log n$, then*

$$|P_{[-n, y], \omega}^E| \leq e^{4pra^{-1}(\log n)^2}. \quad (25)$$

The same result also holds for $|n - y| < p \log n$ and $|P_{[y, n], \omega}^E|$ substituted in appropriately.

Proof. The Proof of Lemma VI.5 only relied on the existence of a uniform (in m) bound on the α th moments of $|V_m(\omega)|$. Clearly, if E varies over compact $I \subset \mathbb{R}$, there is similarly a uniform (in m) bound on the α th moments of $\sup_{E \in I} |V_m(\omega) - E|$. Hence, replacing C in our proof with larger $\tilde{C}$ if necessary, we obtain the necessary result. Because of the tri-diagonal nature of $H_{[-n, y], \omega}$, we have

$$P_{[-n, y], \omega}^E = \prod_{m=-n}^y (V_m(\omega) - E),$$

whence we obtain

$$|P_{[-n, y], \omega}^E| \leq n^{\frac{2\alpha}{r}(p \log n + 1)(\log n)}$$

for n large. Because $1 < p \log n$ for large n , the result follows immediately. $\square$

In the bounded context, there is an obvious analog of this result; there exists M such that $|V_n(\omega)| \leq M$ for all n , and so (taking $\tilde{M}$ slightly larger if necessary to account for varying E over I) $|P_{[-n, y], \omega}^E| \leq \tilde{M}^{b-a+1}$, without even making any restrictions to edges. Hence, in particular, we have the following proposition:

Proposition VI.7. *In the setting of Theorem II.5 with $\text{supp } \mathbb{P}_0 \subset [-M, M]^\mathbb{Z}$, if we fix $I \subset \mathbb{R}$ compact and $L > 1$, then for all $\omega \in \text{supp } \mathbb{P}_0$, there is $N = N(\omega)$ such that if $n > N$ and $| -n - y| > \frac{n}{L}$, then*

$$|P_{[-n, y], \omega}^E| \leq \tilde{M}^{\frac{n}{L}}, \quad (26)$$

where $\tilde{M} = \sup_{E \in I} |M - E|$. The same result also holds for $|n - y| < \frac{n}{L}$ and $|P_{[y,n],\omega}^E|$ substituted in appropriately.

The availability of such a bound is what allows the use of Lemma VI.4 instead of VI.3. Ultimately, the problem of using Lemma VI.4 with our weaker control of $|V_m(\omega)|$ comes down to the superexponential growth of $n^{\frac{r}{k}}$, whereas $n^{pr \log n} = e^{pr(\log n)^2}$ grows subexponentially and $\tilde{M}^{n/L}$ grows exponentially. In the following Proof of Theorem IV.10 under the assumptions of II.9, we note that replacing $\eta_0 := \eta(\varepsilon_0)$ and $\eta_\varepsilon := \eta(\varepsilon)$ from Theorem V.2 with $\tilde{\eta}_0 := \eta(\varepsilon_0, 4)$ and $\tilde{\eta}_\varepsilon := \eta(\varepsilon, 4)$ from Theorem V.1, Lemma VI.3 with Lemma VI.4, and $e^{pr(\log n)^2}$ with $\tilde{M}^{n/L}$, where L is chosen to be sufficiently large (specifically $L > \frac{3 \log \tilde{M}}{\varepsilon_0 - \delta_0}$ in terms of parameters introduced in Sec. VII), we obtain a Proof of Theorem IV.10 under the assumptions of Theorem II.5.

VII. PROOF OF THEOREM IV.10

Having established these technical results in our non-stationary case, we are ready to prove $\mathbb{P}_0$ almost sure eventual $(\gamma(E) - 8\varepsilon, n, E, \omega)$ -regularity of $2n + 1$, where $\varepsilon < v_I/8$. The proof for $2n$, $-2n$, and $-2n - 1$ follows by a nearly identical argument (in the i.i.d. potential setting, symmetry considerations obviate any need to consider the negative cases even in passing; the absence of stationarity here technically forces one to run the argument in the negative direction as well, though no unique technical difficulties arise).

This follows Ref. 17, which made some necessary adjustments to account for unboundedness. We note that many technical details there are absent here because they arise from considering the Jacobi case rather than only the Schrödinger case; more or less of all of these reduce to showing bounds on the growth of hopping terms, which are uniformly 1 in the Schrödinger case. We believe that these additional considerations do not prevent the argument from going through in the more general Jacobi case but have not gone through the details.

Although the argument more or less follows Refs. 5 and 17, the former discussing bounded Schrödinger operators and the latter unbounded Jacobi operators, both of which were inspired by Ref. 36, we say a few words about the general strategy. A particular site n being $(\gamma(E) - \varepsilon, n, E, \omega)$ -singular forces a “resonance” of a sort; for sufficiently large n , Corollary V.9 is used to establish the existence of eigenvalues E_i and E_j for truncations to $[-n, n]$ and $[n + 1, 3n + 1]$, which give “sub”-deviations from the expected Lyapunov behavior, either on the whole intervals $[-n, n]$ and $[n + 1, 3n + 1]$ or on subintervals. Using Lemmas VI.3 and VI.5, we derive three inequalities for different cases, at least one holding for each instance of singularity. All of these inequalities fail for large enough n ; it follows that there cannot be infinitely many singular points with respect to any set of parameters satisfying our assumptions, giving localization with respect to the compact interval I .

Proof of Theorem IV.10. We fix $0 < \varepsilon_0 < v_I/8$. We then choose η_0 a large deviation parameter satisfying the conclusion of Theorem V.2 for ε_0 . Then, given these parameters, we fix $0 < \delta_0 < \eta_0$ and $0 < \varepsilon < \min\{(\eta_0 - \delta_0)/3, \varepsilon_0\}$. Given these parameters, we take $\tilde{\Omega}$ with probability 1 such that the conclusions of Corollary V.9 and Lemmas VI.3 and VI.5 hold.

We let $\omega \in \tilde{\Omega}$ and $E \in I$ be a generalized eigenvalue of H_ω . We further let ψ be the associated generalized eigenfunction. At least one of $\psi(0), \psi(1)$ is non-zero; we assume without loss of generality that $\psi(0) \neq 0$. There is $N(\omega)$ satisfying the conclusions of Corollary V.9 and Lemmas VI.3 and VI.5 and so that, furthermore, for $n > N$, we have that 0 is $(\gamma(E) - 8\varepsilon_0, n, E, \omega)$ -singular [this singularity is a consequence of $\psi(0) \neq 0$ together with (15) and the polynomial bound on growth of $|\psi(n)|$].

We suppose (toward a contradiction) that for infinitely many $n > N$, $2n + 1$ is also $(\gamma(E) - 8\varepsilon_0, n, E, \omega)$ -singular. By Lemma VI.1 and Corollary V.9, we have that $E \in B_{[n+1, 3n+1], \varepsilon_0, \omega}^-$. There is then for any fixed $n > N$ some E_j , an eigenvalue of $H_{\omega, [n+1, 3n+1]}$ such that E_j lies in a slightly larger band of energies $\tilde{I} := [s - 1, t + 1]$ (recall that $I = [s, t]$) and $|E - E_j| \leq e^{-(\eta_0 - \delta_0)(2n+1)}$. Were this not the case, then either all eigenvalues would lie to one side of $\tilde{I}$ or some would lie on each side. In full detail, we show that the latter case is impossible, with a proof that generalizes straightforwardly to the former. We let E_{j-} be the largest eigenvalue to the left of I and E_{j+} be the smallest to the right.

All E_j are the real zeroes of $P_{[n+1, 3n+1], E, \omega}$, which is a polynomial in E of degree $2n + 1$. Then, $P_{[a, b], \omega}^E$ is monotone on one of $[E_{j-}, E]$ or $[E, E_{j+}]$. Then, in particular, we have

$$1 \leq \min\{|E_{j-} - E|, |E_{j+} - E|\} \leq m(B_{[a, b], \omega}^-) \leq e^{-(\eta_0 - \delta_0)(2n+1)} < 1,$$

whence we conclude it is impossible that there are no eigenvalues in $\tilde{I}$, and so there is some $E_j \in \tilde{I}$. Our proof, in particular, showed that one such E_j satisfies $|E - E_j| \leq e^{-(\eta_0 - \delta_0)(2n+1)}$. We can repeat this argument to produce an eigenvalue E_i of $H_{\omega, [-n, n]} \in B_{[-n, n], \varepsilon, \omega}^-$ such that $|E_i - E| \leq e^{-(\eta_0 - \delta_0)(2n+1)}$. Then, $|E_i - E_j| \leq 2e^{-(\eta_0 - \delta_0)(2n+1)}$. By Lemma VI.3, we have, in particular, that $E_j \notin B_{[-n, n], \omega, \varepsilon}$. However, because E_i is an eigenvalue of $H_{[-n, n], \omega}$, we obtain

$$\|G_{[-n, n], E_j, \omega}\| \geq \frac{1}{2} e^{(\eta_0 - \delta_0)(2n+1)},$$

and by equivalence of norms, there are $y_1, y_2 \in [-n, n]$ such that $y_1 \leq y_2$ and

$$|G_{[-n, n], E_j, \omega}(y_1, y_2)| \geq \frac{1}{2\sqrt{2n+1}} e^{(\eta_0 - \delta_0)(2n+1)}. \quad (27)$$

These two facts together will yield precisely the sought contradiction. Because $E_j \notin B_{[-n,n],\varepsilon,\omega}$, we obtain

$$|P_{[-n,n],\omega}| \geq e^{(\gamma(E_j)-\varepsilon)(2n+1)},$$

which we can combine with (16) to obtain

$$|P_{[-n,y_1-1],\omega}^{E_j}| \cdot |P_{[y_2+1,n],E_j,\omega}^{E_j}| \geq \frac{e^{(\eta_0-\delta_0+\gamma(E_j)-\varepsilon)(2n+1)}}{2\sqrt{2n+1}}. \quad (28)$$

Because $y_1 < y_2$, our analysis can be split into (essentially) three cases: $|-n-y_1| \geq p \log n$ and $|n-y_2| \geq p \log n$, $|-n-y_1| \geq \log n$ but $|n-y_2| < \log n$, and $|-n-y_1| < \log n$ and $|n-y_2| < p \log n$. (A fourth case mirroring the second also exists; we omit any explicit consideration as the argument is the same.) In the first case, it is an immediate consequence of Lemma VI.3 that (28) yields

$$e^{(\gamma(E_j)+\varepsilon)(2n+1)} \geq \frac{1}{2\sqrt{2n+1}} e^{(\eta_0-\delta_0+\gamma(E_j)-\varepsilon)(2n+1)}. \quad (29)$$

By our choice of $\varepsilon < (\delta_0 - \eta_0)/3$, this cannot hold for arbitrarily large n .

For the second and third cases, we use Lemma VI.5 to bound the term “close” to the edge. Hence, for the second case, (28) yields

$$e^{3pr\alpha^{-1}(\log n)^2} e^{(\gamma(E_j)+\varepsilon)(2n+1)} \geq \frac{1}{2\sqrt{2n+1}} e^{(\eta_0-\delta_0+\gamma(E_j)-\varepsilon)(2n+1)},$$

which also cannot hold for arbitrarily large n . In the third case, (28) yields

$$e^{6pr\alpha^{-1}(\log n)^2} \geq \frac{1}{2\sqrt{2n+1}} e^{(\eta_0-\delta_0+\gamma(E_j)-\varepsilon)(2n+1)},$$

which again cannot hold for arbitrarily large n . Hence, it is impossible that there are infinitely many n such that $2n+1$ is $(\gamma(E)-\varepsilon, n, E, \omega)$ -singular. $\square$

ACKNOWLEDGMENTS

We would like to thank Lana Jitomirskaya for suggesting the problem to us and for many helpful discussions, Rui Han for reading early drafts closely and providing helpful comments, and Anton Gorodetski and Victor Kleptsyn also for reading an early draft and making suggestions. This work was partially supported by Grant Nos. NSF DMS-2052899, DMS-2155211, and Simons 681675.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Omar Hurtado: Writing – original draft (equal).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

- ¹R. Carmona, A. Klein, and F. Martinelli, “Anderson localization for Bernoulli and other singular potentials,” *Commun. Math. Phys.* **108**(1), 41–66 (1987).
- ²C. Shubin, R. Vakilian, and T. Wolff, “Some harmonic analysis questions suggested by Anderson-Bernoulli models,” *Geom. Funct. Anal.* **8**(5), 932–964 (1998).
- ³V. Bucaj, D. Damanik, J. Fillman, V. Gerbuz, T. Vandenboom, F. Wang, and Z. Zhang, “Localization for the one-dimensional Anderson model via positivity and large deviations for the Lyapunov exponent,” *Trans. Am. Math. Soc.* **372**, 3619 (2019).
- ⁴A. Gorodetski and V. Kleptsyn, “Parametric Furstenberg theorem on random products of $SL(2, \mathbb{R})$ matrices,” *Adv. Math.* **378**, 107522 (2021).
- ⁵S. Jitomirskaya and X. Zhu, “Large deviations of the Lyapunov exponent and localization for the 1D Anderson model,” *Commun. Math. Phys.* **370**, 311–324 (2019).
- ⁶J. Ding and C. K. Smart, “Localization near the edge for the Anderson Bernoulli model on the two dimensional lattice,” *Invent. Math.* **219**(2), 467–506 (2020).
- ⁷J. Bourgain and C. E. Kenig, “On localization in the continuous Anderson-Bernoulli model in higher dimension,” *Invent. Math.* **161**(2), 389–426 (2005).
- ⁸F. Germinet and A. Klein, “Bootstrap multiscale analysis and localization in random media,” *Commun. Math. Phys.* **222**(2), 415–448 (2001).

- ⁹L. Li, “Anderson-Bernoulli localization at large disorder on the 2D lattice,” *Commun. Math. Phys.* **393**(1), 151–214 (2020).
- ¹⁰L. Li and L. Zhang, “Anderson-Bernoulli Localization on the 3D lattice and discrete unique continuation principle,” *Duke Math. J.* **171**(2), 327–415 (2019).
- ¹¹J. Bourgain and M. Goldstein, “On nonperturbative localization with quasi-periodic potential,” *Ann. Math.* **152**(3), 835–879 (2000).
- ¹²J. Bourgain, *Green’s Function Estimates for Lattice Schrödinger Operators and Applications*, *Annals of Mathematics Studies* (Princeton University Press, Princeton, NJ, 2005), No. 158. Includes bibliographical references.
- ¹³S. Jitomirskaya, W. Liu, and Y. Shi, “Anderson localization for multi-frequency Schrödinger operators on $\mathbb{Z}^d$,” *Geom. Funct. Anal.* **30**(2), 457–481 (2020).
- ¹⁴W. Liu, “Quantitative inductive estimates for Green’s functions of non-self-adjoint matrices,” *Anal. PDE* **15**(8), 2061–2108 (2023).
- ¹⁵E. Le Page, “Theoremes limites pour les produits de matrices aleatoires,” in *Probability Measures on Groups*, edited by H. Heyer (Springer, Berlin, Heidelberg, 1982), pp. 258–303.
- ¹⁶P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*. *Wiley Series in Probability and Statistics* (Wiley-Interscience, 1997).
- ¹⁷N. Rangamani, “Singular-unbounded random Jacobi matrices,” *J. Math. Phys.* **60**(8), 081904 (2019).
- ¹⁸A. Gorodetski and V. Kleptsyn, “Non-stationary version of Furstenberg theorem on random matrix products,” *Adv. Math.* **378**, 107522 (2022).
- ¹⁹I. Goldsheid, “Exponential growth of products of non-stationary Markov-dependent matrices,” *Int. Math. Res. Not.* **2022**(8), 6310–6346.
- ²⁰W. Craig and B. Simon, “Subharmonicity of the Lyapunov index,” *Duke Math. J.* **50**(2), 551–560 (1983).
- ²¹M. Aizenman and S. Molchanov, “Localization at large disorder and at extreme energies: An elementary derivations,” *Commun. Math. Phys.* **157**(2), 245–278 (1993).
- ²²H. Kunz and B. Souillard, “Sur le spectre des opérateurs aux différences finies aléatoires,” *Commun. Math. Phys.* **78**(2), 201–246 (1980).
- ²³A. Klein, “Unique continuation principle for spectral projections of Schrödinger operators and optimal Wegner estimates for non-ergodic random Schrödinger operators,” *Commun. Math. Phys.* **323**(3), 1229–1246 (2013).
- ²⁴J. Bourgain and A. Klein, “Bounds on the density of states for Schrödinger operators,” *Invent. Math.* **194**(1), 41–72 (2013).
- ²⁵S. Jitomirskaya, H. Schulz-Baldes, and G. Stolz, “Delocalization in random polymer models,” *Commun. Math. Phys.* **233**(1), 27–48 (2003).
- ²⁶N. Rangamani, “Exponential dynamical localization for random word models,” *Ann. Henri Poincaré* **23**(12), 4171–4193 (2022).
- ²⁷A. Cai, P. Duarte, and S. Klein, “Furstenberg theory of mixed random-quasiperiodic cocycles,” [arXiv:2201.04745](https://arxiv.org/abs/2201.04745) (2022).
- ²⁸D. Damanik, J. Fillman, and P. Gohlke, “Spectral characteristics of Schrödinger operators generated by product systems,” *J. Spectr. Theory* **12**(4), 1659–1718 (2022).
- ²⁹A. Elgart and A. Klein, “Ground state energy of trimmed discrete Schrödinger operators and localization for trimmed Anderson models,” *J. Spectral Theory* **4**(2), 391–419 (2014).
- ³⁰A. Elgart and S. Sodin, “The trimmed Anderson model at strong disorder: Localisation and its breakup,” *J. Spectral Theory* **7**(1), 87–110 (2017).
- ³¹R. Durrett, *Probability: Theory and Examples*, 5th ed., *Cambridge Series in Statistical and Probabilistic Mathematics* (Cambridge University Press, 2019).
- ³²H. L. Cycon, R. G. Froese, B. Simon, and W. Kirsch, *Schrödinger Operators: With Applications to Quantum Mechanics and Global Geometry*, Springer Study Edition (Springer, 1987).
- ³³J. Fröhlich and T. Spencer, “Absence of diffusion in the Anderson tight binding model for large disorder or low energy,” *Commun. Math. Phys.* **88**, 151–184 (1983).
- ³⁴J. Tsay, “Some uniform estimates in products of random matrices,” *Taiwan. J. Math.* **3**(3), 291–302 (1999).
- ³⁵H. Furstenberg and Y. Kifer, “Random matrix products and measures on projective spaces,” *Isr. J. Math.* **46**(1), 12–32 (1983).
- ³⁶S. Y. Jitomirskaya, “Metal-insulator transition for the almost Mathieu operator,” *Ann. Math.* **150**(3), 1159–1175 (1999).