Anderson localisation for quasi-one-dimensional random operators

Davide Macera[1] and Sasha Sodin[2]

Abstract

In 1990, Klein, Lacroix, and Speis proved (spectral) Anderson localisation for the Anderson model on the strip of width $W \geqslant 1$, allowing for singular distribution of the potential. Their proof employs multi-scale analysis, in addition to arguments from the theory of random matrix products (the case of regular distributions was handled earlier in the works of Goldsheid and Lacroix by other means). We give a proof of their result avoiding multi-scale analysis, and also extend it to the general quasi-one-dimensional model, allowing, in particular, random hopping. Furthermore, we prove a sharp bound on the eigenfunction correlator of the model, which implies exponential dynamical localisation and exponential decay of the Fermi projection.

The method is also applicable to operatos on the half-line with arbitrary (deterministic, self-adjoint) boundary condition.

Our work generalises and complements the single-scale proofs of localisation in pure one dimension ($W = 1$), recently found by Bucaj–Damanik–Fillman–Gerbuz–VandenBoom–Wang–Zhang, Jitomirskaya–Zhu, Gorodetski–Kleptsyn, and Rangamani.

1 Introduction

1.1 The operator, transfer matrices and Lyapunov exponents

Let $W \geqslant 1$. Let $\{L_x\}_{x \in \mathbb{Z}}$ be a sequence of identically distributed $W \times W$ random matrices in $\mathrm{GL}(W, \mathbb{R})$, and let $\{V_x\}_{x \in \mathbb{Z}}$ be a sequence of identically distributed $W \times W$ real symmetric matrices, so that $\{L_x\}_{x \in \mathbb{Z}}, \{V_x\}_{x \in \mathbb{Z}}$ are jointly independent. Denote by $\mathcal{L}$ the support of L_0 and by $\mathcal{V}$ – the support of V_0. Throughout this paper we assume that

(A) there exists $\eta > 0$ such that

$$\mathbb{E}(\|V_0\|^\eta + \|L_0\|^\eta + \|L_0^{-1}\|^\eta) < \infty \; ;$$

(B) the Zariski closure of the group generated by $\mathcal{L}\mathcal{L}^{-1}$ in $\mathrm{GL}(W, \mathbb{R})$ intersects $\mathcal{L}$ (this holds for example when $\mathbb{1} \in \mathcal{L}$);

---

[1] Department of Mathematics and Physics, Roma Tre University, Largo San Murialdo 1, 00146 Roma, Italy. Email: davide.macera@uniroma3.it

[2] School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom. Email: a.sodin@qmul.ac.uk.

(C) $\mathcal{V}$ is irreducible (i.e. has no common invariant subspaces except for $\{0\}$ and $\mathbb{R}^W$), and $\mathcal{V} - \mathcal{V}$ contains a matrix of rank one.

We are concerned with the spectral properties of the random operator H acting on (a dense subspace of) $\ell_2(\mathbb{Z} \to \mathbb{C}^W)$ via

$$(H\psi)(x) = L_x \psi(x+1) + V_x \psi(x) + L_{x-1}^{\mathsf{T}} \psi(x-1) \, , \quad x \in \mathbb{Z} \, . \tag{1}$$

This model, often referred to as a quasi-one-dimensional random operator, is the general Hamiltonian describing a quantum particle with W internal degrees of freedom in random potential and with nearest-neighbour random hopping. The special case $L_x \equiv \mathbb{1}$ is known as the block Anderson model; it is in turn a generalisation of the Anderson model on the strip $\mathbb{Z} \times \{1, \cdots, W\}$, and, more generally, on $\mathbb{Z} \times \Gamma$, where Γ is any connected finite graph (the assumption that Γ is connected ensures that $\mathcal{V}$ is irreducible). Another known special case of (1) is the Wegner orbital model.

Fix $E \in \mathbb{R}$. If $\psi : \mathbb{Z} \to \mathbb{C}^W$ is a formal solution of the equation

$$L_x \psi(x+1) + V_x \psi(x) + L_{x-1}^{\mathsf{T}} \psi(x-1) = E\psi(x) \, , \quad x \geqslant 1 \, ,$$

then

$$\begin{pmatrix} \psi(x+1) \\ \psi(x) \end{pmatrix} = T_x \begin{pmatrix} \psi(x) \\ \psi(x-1) \end{pmatrix} , \tag{2}$$

where the one-step transfer matrix $T_x \in \mathrm{GL}(2W, \mathbb{R})$ is given by

$$T_x = \begin{pmatrix} L_x^{-1}(E\mathbb{1} - V_x) & -L_x^{-1}L_{x-1}^{\mathsf{T}} \\ \mathbb{1} & 0 \end{pmatrix} . \tag{3}$$

The multi-step transfer matrices $\Phi_{x,y} \in \mathrm{GL}(2W, \mathbb{R})$, $x, y \in \mathbb{Z}$, are defined by

$$\Phi_{x,y} = \begin{cases} T_{x-1} \cdots T_y \, , & x > y \\ \mathbb{1} \, , & x = y \\ T_x^{-1} \cdots T_{y-1}^{-1} \, , & x < y \, , \end{cases} \tag{4}$$

so that

$$\Phi_{x,y} \begin{pmatrix} \psi(y) \\ \psi(y-1) \end{pmatrix} = \begin{pmatrix} \psi(x) \\ \psi(x-1) \end{pmatrix} . \tag{5}$$

In particular, $T_x = \Phi_{x+1,x}$. We abbreviate $\Phi_N = \Phi_{N,0}$. The Lyapunov exponents $\gamma_j(E)$, $1 \leqslant j \leqslant 2W$, are defined as

$$\gamma_j(E) = \lim_{N \to \infty} \frac{1}{N} \mathbb{E} \log s_j(\Phi_N(E)) \, ,$$

where s_j stands for the j-th singular value. It is known [11] that (for fixed E) this limit in expectation is also an almost sure limit. The cocycle $\{\Phi_{x,y}\}$ is conjugate to a symplectic one (see Section 3.1), and hence

$$\gamma_j(E) = -\gamma_{2W+1-j}(E) \, , \quad j = 1, \cdots, W \, .$$

Further, as we shall see in Section 3.2, using the work of Goldsheid [15] to verify the conditions of the Goldsheid–Margulis theorem [16] on the simplicity of the Lyapunov spectrum, that

$$\gamma_1(E) > \gamma_2(E) > \cdots > \gamma_W(E) > 0 \ .$$

We also mention that the Lyapunov exponents $\gamma_j(E)$ are continuous functions of E. This was proved by Furstenberg and Kifer in [12]; it can also be deduced from the large deviation estimate (27) – see Duarte and Klein [8].

1.2   The main results

Theorem 1. *Assume (A)–(C). Then the spectrum of H is almost surely pure point. Moreover, if*

$$\mathcal{E}[H] = \left\{ (E, \psi) \in \mathbb{R} \times \ell_2(\mathbb{Z} \to \mathbb{C}^W) \ : \ \|\psi\| = 1 \ , \ H\psi = E\psi \right\}$$

is the collection of eigenpairs of H, then

$$\mathbb{P}\left\{ \forall (E, \psi) \in \mathcal{E}[H] \ \limsup_{x \to \pm\infty} \frac{1}{|x|} \log \|\psi(x)\| \leqslant -\gamma_W(E) \right\} = 1 \ , \qquad (6)$$

i.e. each eigenfunction decays exponentially, with the rate lower-bounded by the slowest Lyapunov exponent.

Remark 1.1. *It is believed that the lower bound is sharp, i.e. the rate of decay can not be faster than the slowest Lyapunov exponent:*

$$\mathbb{P}\left\{ \forall (E, \psi) \in \mathcal{E}[H] \ \liminf_{x \to \pm\infty} \frac{1}{|x|} \log \|\psi(x)\| \geqslant -\gamma_W(E) \right\} = 1. \qquad (7)$$

We refer to [18] for a discussion and partial results in this direction. For $W = 1$, (7) was proved by Craig and Simon in [7].

The property of having pure point spectrum with exponentially decaying eigenfunctions is a manifestation of Anderson localisation of the random operator H. The mathematical work on Anderson localisation in one dimension was initiated by Goldsheid, Molchanov and Pastur [17], who considered the case $W = 1$, $L_x \equiv 1$ and established the pure point nature of the spectrum under the assumption that the distribution of V_x is regular enough (absolutely continuous with bounded density). A different proof of the result of [17] was found by Kunz and Souillard [25]. Under the same assumptions, the exponential decay of the eigenfunctions was established by Molchanov [29]. The case of singular distributions was treated by Carmona, Klein, and Martinelli [6].

The case $W > 1$ was first considered by Goldsheid [14], who established the pure point nature of the spectrum for the case of the Schrödinger operator on the strip, i.e. when $L_x \equiv \mathbb{1}$, V_x is tridiagonal with the off-diagonal entries equal to 1 and the diagonal ones independent and identically distributed, under the assumption that the distribution of the diagonal etries of V_x is regular. In the same setting, Lacroix [26, 27, 28] proved that the eigenfunctions decay exponentially. The case of the Anderson model on a strip

with general (possibly, singular) distributions was settled by Klein–Lacroix–Speis [23], who established localisation in the strong form (6).

Unlike the earlier, more direct arguments treating regular distributions, the works [6, 23] allowing singular distributions involve a multi-scale argument (as developed in the work of Fröhlich–Spencer [10] on localisation in higher dimension); the theory of random matrix products is used to verify the initial hypothesis of multi-scale analysis. Recently, proofs of the result of [6] avoiding multi-scale analysis were found by Bucaj et al. [5], Jitomirskaya and Zhu [22], and Gorodetski and Kleptsyn [20]; the general one-dimensional case (allowing for random hopping) was settled by Rangamani [30]. Our Theorem 1 can be seen as a generalisation of these works, and especially of [22, 30], to which our arguments are closest in spirit: we give a relatively short and single-scale proof of localisation which applies to arbitrary $W \geqslant 1$, and allows for rather general distributions of V_0 and L_0 (under no regularity assumptions on the distribution of the potential). In particular, we recover and generalise the result of [23].

In fact, we prove a stronger result pertaining to the eigenfunction correlators, introduced by Aizenman [1] (see further the monograph of Aizenman–Warzel [3]). If $\Lambda \subset \mathbb{Z}$ is a finite set, denote by H_Λ the restriction of H to $\ell_2(\Lambda \to \mathbb{C}^W)$, i.e.

$$H_\Lambda = P_\Lambda H P_\Lambda^* ,$$

where $P_\Lambda : \ell_2(\mathbb{Z} \to \mathbb{C}^W) \to \ell_2(\Lambda \to \mathbb{C}^W)$ is the coordinate projection. If $I \subset \mathbb{R}$ is a compact interval, denote

$$Q_I^\Lambda(x, y) = \sup \left\{ \| f(H_\Lambda)_{x,y} \| \ : \ \mathrm{supp}\, f \subset I \, , \, |f| \leqslant 1 \right\} \, , \quad Q_I(x, y) = \sup_{a \leqslant x, y \leqslant b} Q_I^{[a,b]}(x, y) \, .$$

Here $\| f(H_\Lambda)_{x,y} \|$ is the operator norm of the (x, y) block of $f(H_\Lambda)$, and the functions f in the supremum are assumed to be, say, Borel measurable.

Theorem 2. *Assume (A)–(C). For any compact interval $I \subset \mathbb{R}$,*

$$\mathbb{P} \left\{ \limsup_{x \to \pm\infty} \frac{1}{|x|} \log Q_I(x, y) \leqslant - \inf_{E \in I} \gamma_W(E) \right\} = 1 \, . \tag{8}$$

It is known (see [3]) that Theorem 2 implies Theorem 1. By plugging in various choices of f, it also implies (almost sure) dynamical localisation with the sharp rate of exponential decay, the exponential decay of the Fermi projection, et cet. (see e.g. [2] and [3]). We chose to state Theorem 1 as a separate result rather than a corollary of Theorem 2 since its direct proof is somewhat shorter than that of the latter.

We refer to Bucaj et al. [5], Jitomirskaya–Zhu [22], and Ge-Zhao [13] for earlier results on dynamical localisation for $W = 1$.

1.3   Main ingredients of the proof

Similarly to many of the previous works, including [6, 23] and also the recent works [5, 22, 20], the two main ingredients of the proof of localisation are a large deviation estimate and a Wegner-type estimate. We state these in the generality required here. Let $I \subset \mathbb{R}$ be a compact interval, and let $F \subset \mathbb{R}^{2W}$ be a Lagrangian subspace (see Section 3). Denote by $\pi_F : \mathbb{R}^{2W} \to F$ the orthogonal projection onto F.

Proposition 1.2. *Assume (A)–(C). For any $\epsilon > 0$ there exist $C, c > 0$ such that for any $E \in I$ and any Lagrangian subspace $F \subset \mathbb{R}^{2W}$*

$$\mathbb{P}\left\{\left|\frac{1}{N}\log s_W(\Phi_N(E)\pi_F^*) - \gamma_W(E)\right| \geqslant \epsilon\right\} \leqslant Ce^{-cN} . \tag{9}$$

The proof is essentially given in [23]; we outline the necessary reductions in Section 3.1. The second proposition could be also proved along the lines of the special case considered in [23]; we present an alternative (arguably, simpler) argument in Section 3.3.

For an operator H and E in the resolvent set of H, we denote by $G_E[H] = (H - E)^{-1}$ the resolvent of H and by $G_E[H](\cdot, \cdot)$ its matrix elements. If E lies in the spectrum of H, we set $G_E[H](\cdot, \cdot) \equiv \infty$.

Proposition 1.3. *Assume (A)–(C). For any $\epsilon > 0$ there exist $C, c > 0$ such that for any $E \in I$ and $N \geqslant 1$*

$$\begin{aligned}
\mathbb{P}\left\{\|G_E[H_{[-N,N]}](i,i)\| \leqslant e^{-\epsilon N}\right\} &\leqslant Ce^{-cN} \quad (i \in [-N, N]) \\
\mathbb{P}\left\{\|G_E[H_{[-N,N]}](i,i \pm 1)\| \leqslant e^{-\epsilon N}\right\} &\leqslant Ce^{-cN} \quad (i, i \pm 1 \in [-N, N])
\end{aligned} \tag{10}$$

Remark 1.4. *The arguments which we present can be applied to deduce the following strengthening of (10):*

$$\mathbb{P}\left\{\mathrm{dist}(E, \sigma(H_{[-N,N]})) \leqslant e^{-\epsilon N}\right\} \leqslant Ce^{-cN} .$$

We content ourselves with (10) which suffices for the proof of the main theorems.

Klein, Lacroix and Speis [23] use (special cases of) Propositions 1.2 and 1.3 to verify the assumptions required for multi-scale analysis. We deduce Theorems 1 and 2 directly from these propositions. In this aspect, our general strategy is similar to the cited works [5, 22, 20]. However, several of the arguments employed in these works rely on the special features of the model for $W = 1$; therefore our implementation of the strategy differs in several crucial aspects.

2   Proof of the main theorems

2.1   Resonant sites; the main technical proposition

Let $\tau > 0$ be a (small) number. We say that $x \in \mathbb{Z}$ is (τ, E, N)-non-resonant ($x \notin \mathrm{Res}(\tau, E, N)$) if

$$\begin{cases}
\|L_x\| \leqslant e^{\tau N} , \\
\|G_E[H_{[x-N,x+N]}](x, x \pm N)\| \leqslant e^{-(\gamma_W(E)-\tau)N} ,
\end{cases} \tag{11}$$

and (τ, E, N)-resonant ($x \in \mathrm{Res}(\tau, E, N)$) otherwise. The following proposition is the key step towards the proof of Theorems 1 and 2.

Proposition 2.1. *Assume (A)–(C). Let $I \subset \mathbb{R}$ be a compact interval, and let $\tau > 0$. There exist $C, c > 0$ such that for any $N \geqslant 1$*

$$\mathbb{P}\left\{ \max_{E \in I} \mathrm{diam}(\mathrm{Res}(\tau, E, N) \cap [-N^2, N^2]) > 2N \right\} \leqslant C e^{-cN} \ .$$

The remainder of this section is organised as follows. In Section 2.2, we express the Green function in terms of the transfer matrices. Using this expression and Propositions 1.2 and 1.3, we show that the probability that $x \in \mathrm{Res}(\tau, E, N)$ (for a fixed $E \in \mathbb{R}$) is exponentially small. In Section 2.3, we rely on this estimate to prove Proposition 2.1. Then we use this proposition to prove Theorem 1 (Section 2.4) and Theorem 2 (Section 2.5).

2.2   Reduction to transfer matrices

Fix $N \geqslant 1$. Consider the $W \times W$ matrices

$$\begin{aligned}
\Psi_i^+ &= (\mathbb{1} \ 0)\, \Phi_{i,N+1} \begin{pmatrix} 0 \\ \mathbb{1} \end{pmatrix} = (0 \ \mathbb{1}) \Phi_{i+1,N+1} \begin{pmatrix} 0 \\ \mathbb{1} \end{pmatrix} \ , \\
\Psi_i^- &= (\mathbb{1} \ 0)\, \Phi_{i,-N} \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix} = (0 \ \mathbb{1}) \Phi_{i+1,-N} \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix} \ .
\end{aligned} \tag{12}$$

The Green function of $H_{[-N,N]}$ can be expressed in terms of these matrices using the following claim, which holds deterministically for any H of the form (1). A similar expression has been employed already in [14].

Claim 2.2. *If $E \notin \sigma(H_{[-N,N]})$, then:*

1.

$$\begin{pmatrix} \Psi_{\pm 1}^\pm \\ \Psi_0^\pm \end{pmatrix} G_E[H_{[-N,N]}](0, \pm N) = \begin{pmatrix} G_E[H_{[-N,N]}](0, \pm 1) \\ G_E[H_{[-N,N]}](0, 0) \end{pmatrix};$$

2. for any $i, j \in [-N, N]$,

$$G_E[H_{[-N,N]}](i, i) = \begin{cases} \Psi_j^- (\Psi_i^-)^{-1} \left(\Psi_{i+1}^+ (\Psi_i^+)^{-1} - \Psi_{i+1}^- (\Psi_i^-)^{-1} \right)^{-1} L_i^{-1} \ , & i \geq j \\ \Psi_j^+ (\Psi_i^+)^{-1} \left(\Psi_{i+1}^+ (\Psi_i^+)^{-1} - \Psi_{i+1}^- (\Psi_i^-)^{-1} \right)^{-1} L_i^{-1} \ , & i \leq j \ . \end{cases}$$

Proof. Abbreviate $G_E = G_E[H_{[-N,N]}]$, and set $G_E(i, j) = 0$ for $j \notin [-N, N]$. The matrices $G_E(i, j)$, $-N \leqslant j \leqslant N$, are uniquely determined by the system of equations

$$L_j G_E(i, j+1) + (V_j - E\mathbb{1}) G_E(i, j) + L_{j-1}^\mathsf{T} G_E(i, j-1) = \delta_{j,i} \mathbb{1} \ , \quad -N \leqslant j \leqslant N \ . \tag{13}$$

We look for a solution of the form

$$G_E(i, j) = \begin{cases} \Psi_j^- \alpha_i^- \ , j \leqslant i \\ \Psi_j^+ \alpha_i^+ \ , j \geqslant i \ , \end{cases} \tag{14}$$

where

$$\Psi_i^- \alpha_i^- - \Psi_i^+ \alpha_i^+ \;=\; 0 \tag{15}$$

$$\Psi_{i+1}^- \alpha_i^- - \Psi_{i+1}^+ \alpha_i^+ \;=\; -L_i^{-1} \; . \tag{16}$$

The first equation ensures that (14) defines $G_E(i,i)$ consistently, while the second one guarantees that (13) holds for $j = i$. For the other values of j, (13) follows from the construction of the matrices $\Psi_j^\pm$.

The solution to (15)–(16) is explicitly found by elimination:

$$\alpha_i^- = (\Psi_i^-)^{-1}\Psi_i^+ \alpha_i^+ \;, \quad \alpha_i^+ = -(\Psi_{i+1}^-(\Psi_i^-)^{-1}\Psi_i^+ - \Psi_{i+1}^+)^{-1}L_i^{-1} \; .$$

This implies the second part of the claim. For the first part, note that for $j \geqslant i$

$$G_E(0,j) = \Psi_j^+ \alpha_0^+ = \Psi_j^+ (\Psi_0^+)^{-1} G_E(0,0) = \Psi_j^+ (\Psi_1^+)^{-1} G_E(0,1) \; .$$

Observing that $\Psi_N^+ = \mathbb{1}$, we conclude that

$$G_E(0,N) = (\Psi_0^+)^{-1} G_E(0,0) = (\Psi_1^+)^{-1} G_E(0,1) \; ,$$

as claimed. Similarly,

$$G_E(0,-N) = (\Psi_0^-)^{-1} G_E(0,0) = (\Psi_{-1}^-)^{-1} G_E(0,-1) \; . \qquad \square$$

2.3   Proof of Proposition 2.1

Fix a small $\tau > 0$. Without loss of generality I is short enough to ensure that

$$\max_{E \in I} \gamma_W(E) - \min_{E \in I} \gamma_W(E) \leqslant \frac{\tau}{2}$$

(this property is valid for short intervals due to the continuity of γ_W; the statement for larger intervals I follows by compactness). Fix such I (which will be suppressed from the notation), and let

$$\gamma = \frac{1}{2}(\max_{E \in I} \gamma_W(E) + \min_{E \in I} \gamma_W(E)) \;, \quad \text{so that} \quad \sup_{E \in I} |\gamma_W(E) - \gamma| \leqslant \frac{\tau}{4} \; .$$

For $x \in \mathbb{Z}$, let

$$\mathrm{Res}^*(\tau,x,N) = \left\{ E \in I \;:\; \max_\pm \|G_E[H_{[x-N,x+N]}](x,x\pm N)\|_{1,\infty} \geqslant e^{-(\gamma(E)-\frac{\tau}{2})N} \right\} \;,$$

where $\|A\|_{1,\infty} = \max_{1 \leqslant \alpha,\beta \leqslant W} |A_{\alpha,\beta}|$. For N large enough ($N \geqslant N_0(\tau)$),

$$\left(\|L_x\| \leqslant e^{\tau N} \right) \text{ and } (E \notin \mathrm{Res}^*(\tau,x,N)) \Longrightarrow x \notin \mathrm{Res}(\tau,E,N) \; .$$

By (A) and the Chebyshev inequality

$$\mathbb{P}\left\{ \exists x \in [-N^2,N^2] \;:\; \|L_x\| \geqslant e^{\tau N} \right\} \leqslant (2N^2+1)\frac{\mathbb{E}\|L_0\|^\eta}{e^{\tau\eta N}} \leqslant C_1 e^{-c_1 N} \; .$$

Hence the proposition boils down to the following statement:

$$|x-y| > 2N \Longrightarrow \mathbb{P}\left\{ \mathrm{Res}^*(\tau,x,N) \cap \mathrm{Res}^*(\tau,y,N) \neq \varnothing \right\} \leqslant C e^{-cN} \; . \tag{17}$$

The proof of (17) rests on two claims. The first one is deterministic:

Claim 2.3. $\mathrm{Res}^*(\tau, x, N)$ *is the union of at most $C_W N$ disjoint closed intervals.*

Proof. By Cramer's rule, for each $\alpha, \beta \in \{1, \cdots, W\}$ and $\pm$ the function

$$g_{\alpha,\beta}^{\pm} : E \mapsto (G_E[H_{[x-N,x+N]}](x, x \pm N))_{\alpha,\beta}$$

is the ratio of two polynomials of degree $\leqslant W(2N + 1)$. Hence the level set

$$\left\{ E \; : \; |g_{\alpha,\beta}^{\pm}(E)| = e^{-(\gamma - \frac{\tau}{2})N} \right\}$$

is of cardinality $\leqslant W(2N + 1)$ (note that the $\leqslant W(2N + 1)$ discontinuity points of $g_{\alpha,\beta}^{\pm}$ are poles, hence they can not serve as the endpoints of the superlevel sets of this function). Hence our set

$$\left\{ E \; : \; |g_{\alpha,\beta}^{\pm}(E)| \geqslant e^{-(\gamma - \frac{\tau}{2})N} \right\}$$

is the union of at most $\leqslant W(2N + 1)/2$ closed intervals, and $\mathrm{Res}^*(\tau, x, N)$ is the union of at most

$$2 \frac{W(W + 1)}{2} \frac{W(2N + 1)}{2} \leqslant C_W N$$

closed intervals. $\square$

Claim 2.4. *Assume (A)–(C). For any compact interval $I \subset \mathbb{R}$ there exist $C, c > 0$ such that for any $N \geqslant 1$ and any $E \in I$,*

$$\mathbb{P}\left\{ E \in \mathrm{Res}^*(\tau, x, N) \right\} \leqslant C e^{-cN} .$$

Proof. According to Claim 2.2,

$$\|G_E[H_{[-N,N]}](0, \pm N)\| \leqslant \left\{ s_W \begin{pmatrix} \Psi_{\pm 1}^{\pm} \\ \Psi_0^{\pm} \end{pmatrix} \right\}^{-1} \left\| \begin{pmatrix} G_E[H_{[-N,N]}](0, \pm 1) \\ \|G_E[H_{[-N,N]}](0, 0) \end{pmatrix} \right\| ;$$

hence

$$\mathbb{P}\left\{ \|G_E[H_{[-N,N]}](0, \pm N)\| \geqslant e^{-(\gamma_W(E) - \frac{\tau}{4})N} \right\}$$

$$\leqslant \mathbb{P}\left\{ s_W \begin{pmatrix} \Psi_{\pm 1}^{\pm} \\ \Psi_0^{\pm} \end{pmatrix} \leqslant e^{(\gamma_W(E) - \frac{\tau}{8})N} \right\} + \mathbb{P}\left\{ \left\| \begin{pmatrix} G_E[H_{[-N,N]}](0, \pm 1) \\ \|G_E[H_{[-N,N]}](0, 0) \end{pmatrix} \right\| \geqslant e^{\frac{\tau}{8}N} \right\} .$$

By Propositions 1.2 and 1.3, both terms decay exponentially in N, locally uniformly in E. $\square$

Now we can prove (17). By Claim 2.3 both $\mathrm{Res}^*(\tau, x, N)$ and $\mathrm{Res}^*(\tau, y, N)$ are unions of at most $C_W N$ closed intervals. If these two sets intersect, then either one of the edges of the intervals composing the first one lies in the second one, or vice versa. The operators $H_{[x-N,x+N]}$ and $H_{[y-N,y+N]}$ are independent due to the assumption $|x - y| > 2N$, hence by Claim 2.4

$$\mathbb{P}\left\{ \mathrm{Res}^*(\tau, x, N) \cap \mathrm{Res}^*(\tau, y, N) \neq \varnothing \right\} \leqslant 4 C_W N \times C e^{-cN} \leqslant C_1 e^{-c_1 N} .$$

This concludes the proof of (17) and of Proposition 2.1. $\square$

2.4 Spectral localisation: proof of Theorem 1

The proof of localisation is based on Schnol's lemma, which we now recall (see [21] for a version applicable in the current setting). A function $\psi : \mathbb{Z} \to \mathbb{C}^W$ is called a generalised eigenfunction corresponding to a generalised eigenvalue $E \in \mathbb{R}$ if

$$L_x \psi(x+1) + V_x \psi(x) + L_{x-1}^\mathsf{T} \psi(x-1) = E\psi(x) , \quad x \geqslant 0 \tag{18}$$

$$\limsup_{|x| \to \infty} \tfrac{1}{|x|} \log \|\psi(x)\| = 0 . \tag{19}$$

Schnol's lemma asserts that any spectral measure of H is supported on the set of generalised eigenvalues. Thus we need to show that (with full probability) any generalised eigenpair (E, ψ) satisfies

$$\limsup_{|x| \to \infty} \frac{1}{|x|} \log \|\psi(x)\| \leqslant -\gamma_W(E) . \tag{20}$$

Fix a compact interval $I \subset \mathbb{R}$, and $\tau > 0$. Consider the events

$$\mathcal{G}_M(I, \tau) = \left\{ \forall E \in I \ \forall N \geqslant M \ \operatorname{diam}(\operatorname{Res}(\tau, E, N) \cap [-N^2, N^2]) \leqslant 2N \right\} .$$

By Proposition 2.1 and the Borel–Cantelli lemma,

$$\mathbb{P}\left(\bigcup_{M \geqslant 1} \mathcal{G}_M(I, \tau) \right) = 1 .$$

We shall prove that on any $\mathcal{G}_M(I, \tau)$ every generalised eigenpair (E, ψ) with $E \in I$ satisfies

$$\limsup_{|x| \to \infty} \frac{1}{|x|} \log \|\psi(x)\| \leqslant -\gamma_W(E) + 3\tau . \tag{21}$$

From (18), we have for any x

$$\psi(x) = -G_E[H_{[x-N,x+N]}](x, x-N) L_{-N-1}^\mathsf{T} \psi(x-N-1) - G_E[H_{[x-N,x+N]}](x, x+N) L_N \psi(x+N+1) .$$

If $x \notin \operatorname{Res}(\tau, E, N)$, this implies

$$\|\psi(x)\| \leqslant e^{-(\gamma_W(E)-2\tau)N}(\|\psi(x-N-1)\| + \|\psi(x+N+1)\|)$$
$$\leqslant 2e^{-(\gamma_W(E)-2\tau)N} \max(\|\psi(x-N-1)\|, \|\psi(x+N+1)\|) ,$$

whence $f_\tau(x) \stackrel{\text{def}}{=} e^{-\tau|x|} \|\psi(x)\|$ satisfies

$$f_\tau(x) \leqslant 2e^{-(\gamma_W(E)-3\tau)N} \max(f_\tau(x-N-1), f_\tau(x+N+1))) . \tag{22}$$

The function f_τ is bounded due to (19), hence it achieves a maximum at some $x_\psi \in \mathbb{Z}$. For

$$N > \log 2/(\gamma_W(E) - 3\tau) ,$$

(22) can not hold at $x = x_\psi$, thus on $\mathcal{G}_M(I, \tau)$ x for all

$$N \geqslant N_0 \stackrel{\text{def}}{=} \max(M, \log 2/(\gamma_W(E) - 3\tau), |x_\psi|)$$

we have:

$$\mathrm{Res}(\tau, E, N) \cap [-N^2, N^2] \subset [x_\psi - 2N, x_\psi + 2N] \subset [-3N, 3N] \ .$$

Thus (22) holds whenever x, N are such that $3N < |x| \leqslant N^2$ and $N \geqslant N_0$.

For each $x \in \mathbb{Z}$, let $N(x)$ be such that $N^2/10 \leqslant |x| \leqslant N^2/5$. If $|x|$ is large enough, $N(x) \geqslant N_0$. Applying (22) $\lfloor |x|/(N+1) \rfloor - 4$ times, we obtain

$$f_\tau(x) \leqslant (2e^{-(\gamma_W(E) - 3\tau)N})^{\lfloor x/(N+1) \rfloor - 4} \times \max f_\tau \leqslant e^{-(\gamma_W(E) - 3\tau)|x| + C(\sqrt{|x|} + 1)} \times \max f_\tau \ ,$$

which implies (21). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

2.5   Eigenfunction correlator: proof of Theorem 2

Fix a compact interval $I \subset \mathbb{R}$, and let $\gamma = \min_{E \in I} \gamma_W(E)$. The proof of (8) relies on the following fact from [9, Lemma 4.1], based on an idea from [3]:

$$Q_I^\Lambda(x, y) \leqslant \lim_{\epsilon \to +0} \frac{\epsilon}{2} \int_I \|G_E[H_\Lambda](x, y)\|^{1-\epsilon} dE \leqslant W \ . \tag{23}$$

Our goal is to bound on this quantity uniformly in the interval $\Lambda \supset \{x, y\}$. Without loss of generality we can assume that $x = 0$. Choose N such that $N^2/10 \leqslant |y| \leqslant N^2/5$. By Proposition 2.1, for any $\tau \in (0, \gamma)$

$$\mathbb{P}\left\{ \forall E \in I \ \mathrm{diam}(\mathrm{Res}(\tau, E, N) \cap [-N^2, N^2]) \leqslant 2N \right\} \geqslant 1 - Ce^{-cN} \ .$$

We show that on the event

$$\left\{ \forall E \in I \ \mathrm{diam}(\mathrm{Res}(\tau, E, N) \cap [-N^2, N^2]) \leqslant 2N \right\} \tag{24}$$

we have

$$Q_I^\Lambda(0, y) \leqslant e^{-(\gamma - 2\tau)|y|} \ , \quad |y| > C_0(\gamma - \tau) \ . \tag{25}$$

Expand the Green function $G_E[H_\Lambda](0, y)$ as follows. First, iterate the resolvent identity

$$G_E[H_\Lambda](x, y) = G_E[H_{[x-N, x+N]}](x, x - N)L_{x-N-1}^\mathsf{T} G_E[H_\Lambda](x - N - 1, y)$$
$$+ G_E[H_{[x-N, x+N]}](x, x + N)L_{x+N} \quad G_E[H_\Lambda](x + N + 1, y)$$

starting from $x = 0$ at most $|y|/N$ times, or until the first argument of $G_E[H_\Lambda]$ reaches the set $\mathrm{Res}(\tau, E, N)$. Then apply the identity

$$G_E[H_\Lambda](x, u) = G_E[H_\Lambda](x, u - N - 1)L_{u-N-1} G_E[H_{[u-N, u+N]}](u - N, u)$$
$$+ G_E[H_\Lambda](x, u + N + 1)L_{u+N}^\mathsf{T} \quad G_E[H_{[u-N, u+N]}](u + N, u)$$

starting from $u = y$ at most $|y|/N$ times, or until the second argument of $G_E[H_\Lambda]$ reaches the set $\mathrm{Res}(\tau, E, N)$. The resulting expansion has $\leqslant 2^{2|y|/N}$ addends, each of which has the form

$$G_E[H_{[x_0-N, x_0+N]}](x_0, x_1) \cdots G_E[H_{[x_{k-1}-N, x_{k-1}+N]}](x_{k-1}, x_k)$$
$$G_E[H_\Lambda](x_k, y_\ell) \tag{26}$$
$$G_E[H_{[y_{\ell-1}-N, y_{\ell-1}+N]}](y_\ell, y_{\ell-1}) \cdots G_E[H_{[y_0-N, y_0+N]}](y_1, y_0) \ ,$$

where $x_0 = 0$, $x_{j+1} = x_j \pm N$, $y_0 = y$, $y_{j+1} = y_j \pm N$, and (by the construction of the event (24)) $k + \ell \geqslant |y|/N - 4$. All the terms in the first and third line of (26) are bounded in norm by $e^{-(\gamma - \tau)N}$, hence

$$\|G_E[H_\Lambda](0, y)\| \leqslant 64 \left(4e^{-(\gamma - \tau)N} \right)^{|y|/N - 4} \sum_{u,v \leqslant 2|y|} \|G_E[H_\Lambda](u, v)\| \ .$$

Now we raise this estimate to the power $1 - \epsilon$ and integrate over $E \in I$:

$$\frac{\epsilon}{2} \int_I \|G_E[H_\Lambda](0, y)\|^{1-\epsilon} dE \leqslant 64^{1-\epsilon} \left(4e^{-(\gamma - \tau)N} \right)^{(1-\epsilon)(|y|/N - 4)} \sum_{u,v \leqslant 2|y|} \frac{\epsilon}{2} \int_I \|G_E[H_\Lambda](u, v)\|^{1-\epsilon} dE \ .$$

It remains to let $\epsilon \to +0$ while making use of the two inequalities in (23). $\qquad\square$

3   Properties of transfer matrices

3.1   Preliminaries

Denote

$$J = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \in \mathrm{GL}(2W, \mathbb{R}) \ .$$

A matrix $Q \in \mathrm{GL}(2W, \mathbb{R})$ is called symplectic, $Q \in \mathrm{Sp}(2W, \mathbb{R})$, if $Q^\intercal J Q = J$.

The matrices T_x are, generally speaking, not symplectic. However, the cocycle $\{\Phi_{x,y}\}_{x,y,\in\mathbb{Z}}$ is conjugate to a symplectic one. Indeed, observe that

Claim 3.1. *If $L \in \mathrm{GL}(W, \mathbb{R})$ and Z is $W \times W$ real symmetric, then $Q(L, Z) = \begin{pmatrix} L^{-1}Z & -L^{-1} \\ L^\intercal & 0 \end{pmatrix}$ is symplectic.*

Denote $D_x = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & L_x^\intercal \end{pmatrix}$, then

$$\widetilde{T}_x(E) \overset{\text{def}}{=} D_x T_x(E) D_{x-1}^{-1} = Q(L_x, E\mathbb{1} - V_x) \in \mathrm{Sp}(2W, \mathbb{R}) \ .$$

Thus also

$$\widetilde{\Phi}_{x,y}(E) = D_{x-1} \Phi_{x,y}(E) D_{y-1}^{-1} = \begin{cases} \widetilde{T}_{x-1}(E) \cdots \widetilde{T}_y(E) \ , & x > y \\ \mathbb{1} \ , & x = y \\ \widetilde{T}_x^{-1}(E) \cdots T_{y-1}^{-1}(E) \ , & x < y \end{cases} \in \mathrm{Sp}(2W, \mathbb{R}) \ .$$

3.2   Simplicity of the Lyapunov spectrum and large deviations

Goldsheid and Margulis showed [16] that if g_j are independent, identically distributed random matrices in $\mathrm{Sp}(2W, \mathbb{R})$, and the group generated by the support of g_1 is Zariski dense in $\mathrm{Sp}(2W, \mathbb{R})$, then the Lyapunov spectrum of a random matrix product $\{g_N \cdots g_1\}$ is simple, i.e.

$$\gamma_1 > \cdots > \gamma_W > 0 \ .$$

Goldsheid showed [15] that if $\mathcal{V}$ is irreducible and $\mathcal{V} - \mathcal{V}$ contains a rank-one matrix, then for any $E \in \mathbb{R}$ the group generated by $Q(\mathbb{1}, E\mathbb{1} - V)$, $V \in \mathcal{V}$, is Zariski dense in $\mathrm{Sp}(2W, \mathbb{R})$.

Corollary 3.2. *Assume (A)–(C). Then for any $E \in \mathbb{R}$*

$$\gamma_1(E) > \cdots > \gamma_W(E) > 0 \ .$$

Proof. Observe that

$$Q(L, E\mathbb{1} - V) = \begin{pmatrix} L^{-1} & 0 \\ 0 & L^{\intercal} \end{pmatrix} Q(\mathbb{1}, E\mathbb{1} - V) \ ,$$

whence

$$Q(\widehat{L}, E\mathbb{1} - V)^{-1} Q(L, E\mathbb{1} - V) = \begin{pmatrix} \widehat{L}L^{-1} & 0 \\ 0 & \widehat{L}^{-\intercal}L^{\intercal} \end{pmatrix} \ .$$

If the Zariski closure of the group generated by $\mathcal{L}\mathcal{L}^{-1}$ intersects $\mathcal{L}$, then the Zariski closure of the group generated by $\{Q(L, E\mathbb{1} - V)\}_{L \in \mathcal{L}, V \in \mathcal{V}}$ contains that of the group generated by $\{Q(\mathbb{1}, E\mathbb{1} - V)\}_{V \in \mathcal{V}}$. $\qquad\square$

Having the corollary at hand, we deduce from [23, Proposition 2.7] applied to the matrices $\widetilde{\Phi}_N(E)$:

Proposition 3.3. *Assume (A)–(C). For any $\epsilon > 0$ there exist $C, c > 0$ such that for any $E \in I$ and $1 \leqslant j \leqslant W$*

$$\mathbb{P}\left\{ \left| \frac{1}{N} \log s_j(\widetilde{\Phi}_N(E)) - \gamma_j(E) \right| \geqslant \epsilon \right\} \leqslant Ce^{-cN} \ . \tag{27}$$

Further, for any Lagrangian subspace $F \subset \mathbb{R}^{2W}$

$$\mathbb{P}\left\{ \left| \frac{1}{N} \log s_j(\widetilde{\Phi}_N(E)\pi_F^*) - \gamma_j(E) \right| \geqslant \epsilon \right\} \leqslant Ce^{-cN} \ . \tag{28}$$

Proof. The estimate (28) is a restatement of [23, Proposition 2.7], whereas (27) follows from (28) applied to a δ-net on the manifold of Lagrangian subspaces of $\mathbb{R}^{2W}$ (the Lagrangian Grassmannian). We note that (27) is also proved directly in [8]. $\qquad\square$

Note that Proposition 1.2 follows from (28).

Now fix ϵ and a Lagrangian subspace F, and let

$$\Omega_\epsilon^F[\widetilde{\Phi}_N] = \left\{ \max_{j=1}^W \left[|\frac{1}{N} \log s_j(\widetilde{\Phi}_N) - \gamma_j| + |\frac{1}{N} \log s_j(\widetilde{\Phi}_N \pi_F^*) - \gamma_j| \right] \leqslant \frac{\epsilon}{100W} \right\} \ . \tag{29}$$

According to Proposition 3.3,

$$\mathbb{P}(\Omega_\epsilon^F[\widetilde{\Phi}_N(E)]) \geqslant 1 - C(\epsilon, E)e^{-c(\epsilon, E)N} \ ,$$

where the constants are locally uniform in E. Let

$$\widetilde{\Phi}_N(E) = U_N(E)\Sigma_N(E)V_N(E)^{\mathsf{T}}$$

be the singular value decomposition of $\widetilde{\Phi}_N(E)$. Assume that the singular values on the diagonal of $\Sigma_N(E)$ are arranged in non-increasing order; the choice of the additional degrees of freedom is not essential for the current discussion. Denote

$$F_+ = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R}^W \right\} \subset \mathbb{R}^{2W} , \quad F_- = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} : y \in \mathbb{R}^W \right\} \subset \mathbb{R}^{2W} . \qquad (30)$$

Claim 3.4. *Let $F \subset \mathbb{R}^{2W}$ be a Lagrangian subspace. For N large enough (depending on ϵ), one has (deterministically) on the event $\Omega_\epsilon^F[\widetilde{\Phi}_N(E)]$ defined in (29)*

$$s_W(\pi_{F_+} V_N(E)^{\mathsf{T}} \pi_F^*) \geqslant e^{-\frac{\epsilon}{25}N} .$$

Remark 3.5. *For future reference, we also record the dual version of the claim: on $\Omega_\epsilon^F[\widetilde{\Phi}_N(E)^{\mathsf{T}}]$*

$$s_W(\pi_F^* U_N(E)\pi_{F_+}) \geqslant e^{-\frac{\epsilon}{25}N} .$$

Proof. We abbreviate $\Sigma = \Sigma_N(E)$, $V = V_N(E)$, and $\gamma_j = \gamma_j(E)$. On the other hand, the constants with ϵ not explicitly present in the notation will be uniform in $\epsilon \to +0$.

Clearly, $s_j(\pi_{F_+} V^{\mathsf{T}} \pi_F^*) \leqslant \|\pi_{F_+} V^{\mathsf{T}} \pi_F^*\| \leqslant 1$. Hence it will suffice to show that on $\Omega_\epsilon^F[\widetilde{\Phi}_N(E)]$

$$\prod_{k=1}^{W} s_k(\pi_{F_+} V^{\mathsf{T}} \pi_F^*) \geqslant e^{-\epsilon N} . \qquad (31)$$

Let Σ^+ be the diagonal matrix obtained by setting the (k,k) matrix entries of Σ to zero for $k > W$. Then on $\Omega_\epsilon^F[\widetilde{\Phi}_N(E)]$ we have

$$\|\Sigma - \Sigma^+\| \leqslant \exp(-cN)$$

(with $c > 0$ uniform in $\epsilon \to +0$). Thus $s_j(\widetilde{\Phi}_N \pi_F^*) = s_j(\Sigma V^{\mathsf{T}} \pi_F^*)$ satisfies

$$|s_j(\widetilde{\Phi}_N \pi_F^*) - s_j(\Sigma^+ V^{\mathsf{T}} \pi_F^*)| \leqslant e^{-cN} .$$

Observing that $s_j(\Sigma^+ V^{\mathsf{T}} \pi_F^*) = s_j(\widehat{\Sigma}^+ \pi_{F_+} V^{\mathsf{T}} \pi_F^*)$, where $\widehat{\Sigma}^+ = \pi_{F_+} \Sigma^+ \pi_{F_+}^*$, and that

$$s_j(\widetilde{\Phi}_N \pi_F^*) \geqslant e^{(\gamma_j - \frac{\epsilon}{100W})N}$$

on Ω_ϵ^F, we get (for sufficiently large N):

$$s_j(\widehat{\Sigma}^+ \pi_{F_+} V^{\mathsf{T}} \pi_F^*) \geqslant e^{(\gamma_j - \frac{\epsilon}{50W})N} , \quad \prod_{k=1}^{j} s_k(\widehat{\Sigma}^+ \pi_{F_+} V^{\mathsf{T}} \pi_F^*) \geqslant e^{(\gamma_1 + \cdots + \gamma_j - \frac{\epsilon}{50})N} .$$

On the other hand, using the submultiplicativity of the operator norm and the equalitiy between the norm of the j-th wedge power of a matrix and the product of its j top singular values, we have

$$\prod_{k=1}^{j} s_k(\widehat{\Sigma}^+ \pi_{F_+} V^\mathsf{T} \pi_F^*) \leqslant \prod_{k=1}^{j} s_k(\widehat{\Sigma}^+) \times \prod_{k=1}^{j} s_k(\pi_{F_+} V^\mathsf{T} \pi_F^*)$$

$$\leqslant e^{(\gamma_1 + \cdots + \gamma_j + \frac{\epsilon}{100})N} \prod_{k=1}^{j} s_k(\pi_{F_+} V^\mathsf{T} \pi_F^*) \,,$$

whence

$$\prod_{k=1}^{j} s_k(\pi_{F_+} V^\mathsf{T} \pi_F^*) \geqslant e^{-\frac{\epsilon}{25} N} \,, \quad 1 \leqslant j \leqslant W \,,$$

thus concluding the proof of (31) and of the claim. □

3.3   Wegner-type estimate: proof of Proposition 1.3

Let us first show that for any $i \in [-N, N]$

$$\mathbb{P}\left\{ \|G_E[H_{-N,N}](i,i)\| \geqslant e^{\epsilon N} \right\} \leqslant C_\epsilon e^{-c_\epsilon N} \,. \tag{32}$$

By Claim 2.2,

$$G_E[H_{[-N,N]}](i,i) = \left(\Psi_{i+1}^+ (\Psi_i^+)^{-1} - \Psi_{i+1}^- (\Psi_i^-)^{-1} \right)^{-1} L_i^{-1} \,,$$

where

$$\begin{pmatrix} \Psi_{i+1}^+ \\ \Psi_i^+ \end{pmatrix} = \Phi_{i+1,N+1} \begin{pmatrix} 0 \\ \mathbb{1} \end{pmatrix} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & L_i^{-\mathsf{T}} \end{pmatrix} \widetilde{\Phi}_{i+1,N+1} \begin{pmatrix} 0 \\ L_N^\mathsf{T} \end{pmatrix} \,,$$

$$\begin{pmatrix} \Psi_{i+1}^- \\ \Psi_i^- \end{pmatrix} = \Phi_{i+1,-N} \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & L_i^{-\mathsf{T}} \end{pmatrix} \widetilde{\Phi}_{i+1,-N} \begin{pmatrix} \mathbb{1} \\ 0 \end{pmatrix} \,.$$

Hence

$$G_E[H_{[-N,N]}](i,i) = L_i^{-\mathsf{T}} \left(X^+ - X^- \right)^{-1} L_i^{-1} \,,$$

where

$$X^+ = (\widetilde{\Phi}_{i+1,N+1})_{12} (\widetilde{\Phi}_{i+1,N+1})_{22}^{-1} \,, \quad X^- = (\widetilde{\Phi}_{i+1,-N})_{11} (\widetilde{\Phi}_{i+1,-N})_{21}^{-1} \,,$$

and the subscripts 11 and 21 represent extracting the corresponding $W \times W$ blocks from a $2W \times 2W$ matrix (i.e. $Y_{11} = \pi_{F_+} Y \pi_{F_+}^*$, $Y_{21} = \pi_{F_-} Y \pi_{F_+}^*$, in the notation of (30)). Both matrices $X^\pm$ are Hermitian, as follows from the symplectic property of the transfer matrices.

Without loss of generality we can assume that $i \geqslant 0$. We shall prove that

$$\mathbb{P}\left\{ s_W(X^+ - X^-) \leqslant e^{-\epsilon N} \mid X^+ \right\} \leqslant C_\epsilon e^{-c_\epsilon N} \,.$$

To this end, denote

$$F = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2W} \; : \; y = -X^+ x \right\} \,.$$

In the notation of Claim 3.4, consider the transfer matrix $\widetilde{\Phi}_{i+1,-N}$, and let

$$\Omega_\epsilon = \Omega_\epsilon^F[\widetilde{\Phi}^*] \cap \Omega_\epsilon^{F_+}[\widetilde{\Phi}^*] \cap \Omega_\epsilon^{F_-}[\widetilde{\Phi}^*] \cap \Omega_\epsilon^{F_+}[\widetilde{\Phi}]$$

(note that $\widetilde{\Phi}_{i+1,-N}$ is independent of X^+ and thus also of F). It suffices to show that on Ω_ϵ

$$s_W(X^+ - X^-) \geqslant e^{-\frac{\epsilon}{2}N} . \tag{33}$$

Let us write the singular value decomposition of $\widetilde{\Phi} = \widetilde{\Phi}_{i+1,-N}$ in block form:

$$\begin{pmatrix} \widetilde{\Phi}_{11} & \widetilde{\Phi}_{12} \\ \widetilde{\Phi}_{21} & \widetilde{\Phi}_{22} \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} \widehat{\Sigma}^+ & \\ & \widehat{\Sigma}^- \end{pmatrix} \begin{pmatrix} V_{11}^\mathsf{T} & V_{21}^\mathsf{T} \\ V_{12}^\mathsf{T} & V_{22}^\mathsf{T} \end{pmatrix}$$

whence on Ω_ϵ

$$\|\widetilde{\Phi}_{11} - U_{11}\widehat{\Sigma}^+ V_{11}^\mathsf{T}\| \, , \|\widetilde{\Phi}_{21} - U_{21}\widehat{\Sigma}^+ V_{11}^\mathsf{T}\| \leqslant e^{-cN} .$$

Further, by Claim 3.4 we have on Ω_ϵ:

$$s_W(U_{11}) \, , s_W(U_{21}) \, , s_W(V_{22}) \geqslant e^{-\frac{\epsilon}{25}N} . \tag{34}$$

Let us show that

$$\|X^- - U_{11}U_{21}^{-1}\| \leqslant e^{-c'N} . \tag{35}$$

To this end, start with the relation

$$X^- = (U_{11}\widehat{\Sigma}^+ V_{11}^\mathsf{T} + E_1)(U_{21}\widehat{\Sigma}^+ V_{11}^\mathsf{T} + E_2)^{-1}, \quad \|E_1\|, \|E_2\| \leqslant e^{-cN} .$$

In view of the bound

$$s_W(U_{21}\widehat{\Sigma}^+ V_{11}^\mathsf{T}) \geqslant s_W(U_{21})s_W(\widehat{\Sigma}^+)s_W(V_{11}^\mathsf{T}) \geqslant e^{+c_1 N} ,$$

we can set $E_2' = E_2(U_{21}\widehat{\Sigma}^+ V_{11}^\mathsf{T})^{-1}$ and rewrite

$$(U_{21}\widehat{\Sigma}^+ V_{11}^\mathsf{T} + E_2)^{-1} = (U_{21}\widehat{\Sigma}^+ V_{11}^\mathsf{T})^{-1}(1 + E_2') , \quad \|E_2'\| \leqslant e^{-c_2 N} ,$$

which implies (35).

Now, the matrix X^+ is symmetric, therefore $x - X^+ y = 0$ for $\binom{x}{y} \in F^\perp$, whence for any $\binom{x}{y} \in \mathbb{R}^{2W}$

$$x - X^+ y = (\mathbb{1} \mid -X^+)\pi_F^* \pi_F \begin{pmatrix} x \\ y \end{pmatrix}$$

(where the first term is a 1×2 block matrix). Therefore we have, by another application of Claim 3.4:

$$s_W(U_{11} - X^+ U_{21}) = s_W((\mathbb{1} \mid -X^+)\pi_F^* \pi_F U \pi_{F_+}^*)$$
$$\geqslant s_W((\mathbb{1} \mid -X^+)\pi_F^*)s_W(\pi_F U \pi_{F_+}^*) \geqslant s_W(\pi_F U \pi_{F_+}^*) \geqslant e^{-\frac{\epsilon}{25}N} .$$

This, together with (35) and (34), concludes the proof of (33), and of (32).

Now we consider the elements $G_E[H_{[-N,N]}](i, i \pm 1)$. We have:

$$G_E[H_{[-N,N]}](i, i \pm 1) = \Psi_{i+1}^\pm (\Psi_i^\pm)^{-1} G_E[H_{[-N,N]}](i, i) .$$

The norm of $G_E[H_{[-N,N]}](i, i)$ is controlled by (32), whereas $\Psi_{i+1}^\pm (\Psi_i^\pm)^{-1} = L_i^{-1} X^\pm$ are controled using (35) and Claim 3.4. $\qquad \square$

4   On generalisations

Other distributions   The assumptions (A)–(C) in Theorems 1 and 2 can probably be relaxed. Instead of a finite fractional moment in (A), it should be sufficient to assume the existence of a sufficiently high logarithmic moment:

$$\mathbb{E}(\log_+^A \|V_0\| + \log_+^A \|L_0\| + \log_+^A \|L_0^{-1}\|) < \infty$$

for a sufficiently large $A > 1$. To carry out the proof under this assumption in place of (A), one would need appropriate versions of large deviation estimates for random matrix products.

As we saw in the previous section, the rôle of the assumptions (B)–(C) is to ensure that the conditions of the Goldsheid–Margulis theorem [16] are satisfied. That is, our argument yields the following:

Theorem 3. *Let $I \subset \mathbb{R}$ be a compact interval. Assume (A) and that for any $E \in I$ the group generated by*

$$\{Q(L, E\mathbb{1} - V)\}_{L \in \mathcal{L}, V \in \mathcal{V}}$$

is Zariski-dense in $\mathrm{Sp}(2W, \mathbb{R})$. Then:

1. The spectrum of H in I is almost surely pure point, and

$$\mathbb{P}\left\{\forall(E, \psi) \in \mathcal{E}[H] \ \ E \in I \implies \limsup_{x \to \pm\infty} \frac{1}{|x|} \log \|\psi(x)\| \leqslant -\gamma_W(E)\right\} = 1 \ ; \quad (36)$$

2. for any compact subinterval $I' \subset I$ (possibly equal to I) one has:

$$\mathbb{P}\left\{\limsup_{x \to \pm\infty} \frac{1}{|x|} \log Q_I(x, y) \leqslant -\inf_{E \in I} \gamma_W(E)\right\} = 1 \ . \quad (37)$$

As we saw in the previous section, the second condition of this theorem is implied by our assumptions (B)–(C). Most probably, weaker assumptions should suffice, and, in fact, we believe that the conclusions of Theorems 1 and 2 hold as stated without the assumption (B). A proof would require an appropriate generalisation of the results of Goldsheid [15].

Another interesting class of models appears when $V_x \equiv 0$. The complex counterpart of this class, along with a generalisation in which the distribution of L_x depends on the parity of x, has recently been considered by Shapiro [31], in view of applications to topological insulators. An interesting feature of such models is that the slowest Lyapunov exponent $\gamma_W(E)$ may vanish at $E = 0$. This circle of questions (in partiular, the positivity of the smallest Lyapunov exponent and Anderson localisation) is studied in [31] under the assumption that the distribution of L_0 in $\mathrm{GL}(W, \mathbb{C})$ is regular. In order to extend the results of [31] (for matrices complex entries) to singular distributions, one would first need an extension of [16] to the Hermitian symplectic group.

Returning to the (real) setting of the current paper, assume that (B)–(C) are replaced with

(B$'$) the group generated by $\mathcal{L}$ is Zariski-dense in $\mathbf{GL}(W, \mathbb{R})$;

(C$'$) $V_x \equiv 0$.

Along the arguments of [31], one can check that the conditions of [16] hold for any $E \neq 0$. From Theorem 3, one deduces that the conclusion of Theorem 1 holds under the assumptions (A), (B$'$), (C$'$), whereas the conclusion (37) of Theorem 2 holds for compact intervals I not containing 0. If $\gamma_W(0) = 0$, (37) is vacuous for $I \ni 0$. If $\gamma_W(0) > 0$, (37) is meaningful and probably true for such intervals, however, additional arguments are required to establish the large deviation estimates required for the proof.

Finally, we note that Theorem 3 remains valid if the independence assumption is relaxed as follows: $\{(V_x, L_x)\}_{x \in \mathbb{Z}}$ are jointly independent (i.e. we can allow dependence between V_x and the corresponding L_x).

The half-line Similar results can be established for random operators on the half-line. For simplicity, we focus on the case $L_x \equiv \mathbb{1}$. Fix a Lagrangian subspace $F \subset \mathbb{R}^{2W}$, and consider the space $\mathcal{H}_F$ of square-summable sequences $\psi : \mathbb{Z}_+ \to \mathbb{C}^W$ such that $\left(\begin{smallmatrix} \psi(1) \\ \psi(0) \end{smallmatrix} \right) \in F$. Define an operator H_F acting on $\mathcal{H}_F$ so that

$$(H_F \psi)(x) = L_x \psi(x+1) + V_x \psi(x) + L_{x-1}^{\mathsf{T}} \psi(x-1) \,, \quad x \geq 1$$

(see e.g. [4] for details).

Theorem 4. *Fix a Lagrangian subspace $F \subset \mathbb{R}^{2W}$. Under the assumptions (A) and (C) with $L_x \equiv 1$, the spectrum of H_F in any compact interval I is almost surely pure point, and*

$$\mathbb{P} \left\{ \forall (E, \psi) \in \mathcal{E}[H_F] \;\; E \in I \implies \limsup_{x \to \infty} \frac{1}{|x|} \log \|\psi(x)\| \leq -\gamma_W(E) \right\} = 1 \,. \tag{38}$$

Remark 4.1.

1. *For general L_x, the boundary condition has to be prescribed in a different way. However, in the Dirichlet case $F = F_+$ the result holds as stated for general L_x satisfying (A)–(B).*

2. *Combining the proof of Theorem 2 with the additional argument described below, one can also prove dynamical localisation.*

3. *For $W = 1$, a result of Kotani [24] implies that the operator H_F has pure point spectrum for almost every boundary condition F; a similar statement is valid for $W > 1$. As to fixed (deterministic) boundary conditions, the only published reference known to us is the work of Gorodetski–Kleptsyn [20], treating Schrödinger operators in $W = 1$ with Dirichlet boundary conditions.*

4. *The event of full probability provided by Theorem 4 depends on the boundary condition F. And indeed, a result of Gordon [19] implies that (almost surely) there exists a residual set of initial conditions F for which the spectrum of H_F is not pure point (and in fact has only isolated eigenvalues).*

Sketch of proof of Theorem 4. We indicate the necessary modifications with respect to the proof of Theorem 1. First, we modify the definition (11) of $\mathrm{Res}(\tau, E, N)$ as follows: $x \geq N+1$ is said to be (τ, E, N)-non-resonant ($x \notin \mathrm{Res}(\tau, E, N)$) under the same condition

$$\|G_E[H_{[x-N,x+N]}](x, x \pm N)\| \leqslant e^{-(\gamma_W(E)-\tau)N} \ , \tag{39}$$

while $x \in \{1, \cdots, 2N\}$ is said to be (τ, E, N)-non-resonant if

$$\det(\pi_F \Phi_N(E)^* \Phi_N(E) \pi_F^*) \geq e^{2(\gamma_1(E)+\cdots+\gamma_W(E)-\tau)N} \tag{40}$$

(this condition does not depend on x, and only depends on the restriction of the operator to $[1, N]$). We claim that Proposition 2.1 is still valid:

$$\mathbb{P}\left\{\max_{E \in I} \mathrm{diam}(\mathrm{Res}(\tau, E, N) \cap [1, N^2]) > 2N\right\} \leqslant Ce^{-cN} \ . \tag{41}$$

To prove this estimate, it suffices to show that for any $1 \leq x < y \leq N^2$ with $|y - x| > 2N$ one has

$$\mathbb{P}\{\exists E \in I : \ x, y \in \mathrm{Res}(\tau, E, N)\} \leqslant Ce^{-cN} \ . \tag{42}$$

The case $x, y > 2N$ is covered by the current Proposition 2.1. If $x \leq 2N$ and $y > 2N$ the events $x \in \mathrm{Res}(\tau, E, N)$ and $y \in \mathrm{Res}(\tau, E, N)$ are independent; the probability that $x \in \mathrm{Res}(\tau, E, N)$ is exponentially small due to the large deviation estimate (9), and the collection of E violating (40) is the union of $\leq N^{2W}$ intervals. From this point the proof of (42) mimics the argument in the proof of Proposition 2.1.

As in the proof of Theorem 1, let ψ be a generalised solution at energy $E \in I$ given by Schnol's lemma, $x^{-1} \log \|\psi(x)\| \to 0$. Letting $u_x = \binom{\psi(x)}{\psi(x-1)}$, we have

$$\|\Phi_N(E)u_1\| \leq e^{\tau N} \ ,$$

hence for sufficiently large N one has

$$s_W(\Phi_N(E)\pi_F^*) \leq e^{2\tau N} \ .$$

On the other hand, on an event of full probability one has for all $E \in I$ and all sufficiently large N

$$(s_1 \cdots s_{W-1})(\Phi_N(E)\pi_F^*) \leq (s_1 \cdots s_{W-1})(\Phi_N(E)) \leq e^{(\gamma_1(E)+\cdots+\gamma_{W-1}(E)+\tau)N}$$

due to a version of the Craig–Simon theorem [7] (cf. [18, Lemma 2.2]). This implies

$$(s_1 \cdots s_W)(\Phi_N(E)\pi_F^*) \geq e^{(\gamma_1(E)+\cdots+\gamma_{W-1}(E)+3\tau)N} \ ,$$

which contradicts (40) when $\tau > 0$ is small enough. Thus for N large enough

$$\mathrm{Res}(\tau, E, N) \cap [N + 1, N^2] = \varnothing \ ,$$

and thus ψ decays exponentially as in the proof of Theorem 1. $\qquad\square$

References

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

[2] Aizenman, M.; Graf, G. M. Localization bounds for an electron gas. J. Phys. A 31 (1998), no. 32, 6783–6806.

[3] Aizenman, M.; Warzel, S. Random operators. Disorder effects on quantum spectra and dynamics. Graduate Studies in Mathematics, 168. American Mathematical Society, Providence, RI, 2015. xiv+326 pp.

[4] Atkinson, F. V. Discrete and continuous boundary problems. Mathematics in Science and Engineering, Vol. 8. Academic Press, New York-London 1964 xiv+570 pp.

[5] Bucaj, V.; Damanik, D.; Fillman, J.; Gerbuz, V.; VandenBoom, T.; Wang, F.; Zhang, Z. Localization for the one-dimensional Anderson model via positivity and large deviations for the Lyapunov exponent. Trans. Amer. Math. Soc. 372 (2019), no. 5, 3619—3667.

[6] Carmona, René; Klein, Abel; Martinelli, Fabio. Anderson localization for Bernoulli and other singular potentials. Comm. Math. Phys. 108 (1987), no. 1, 41-66.

[7] Craig, W.; Simon, B. Subharmonicity of the Lyaponov index. Duke Math. J. 50 (1983), no. 2, 551–560.

[8] Duarte, Pedro; Klein, Silvius. Lyapunov exponents of linear cocycles. Continuity via large deviations. Atlantis Studies in Dynamical Systems, 3. Atlantis Press, Paris, 2016. xiii+263 pp.

[9] Elgart, A.; Shamis, M.; Sodin, S. Localisation for non-monotone Schrödinger operators. J. Eur. Math. Soc. (JEMS) 16 (2014), no. 5, 909–924.

[10] Fröhlich, Jürg; Spencer, Thomas Absence of diffusion in the Anderson tight binding model for large disorder or low energy. Comm. Math. Phys. 88 (1983), no. 2, 151–184.

[11] Furstenberg, H.; Kesten, H. Products of random matrices. Ann. Math. Statist. 31 (1960), 457–469.

[12] Furstenberg, H.; Kifer, Y. Random matrix products and measures on projective spaces. Israel J. Math. 46 (1983), no. 1–2, 12–32.

[13] Ge, Lingrui; Zhao, Xin. Exponential dynamical localization in expectation for the one dimensional Anderson model. J. Spectr. Theory 10 (2020), no. 3, 887–904.

[14] Goldsheid, I. Ya. Structure of the spectrum of the Schrödinger random difference operator. (Russian) Dokl. Akad. Nauk SSSR 255 (1980), no. 2, 273–277.

[15] Goldsheid, I. Ya. Zariski closure of subgroups of the symplectic group and Lyapunov exponents of the Schrödinger operator on the strip. Comm. Math. Phys. 174 (1995), no. 2, 347–365.

[16] Goldsheid, I. Ya.; Margulis, G. A. Lyapunov exponents of a product of random matrices. (Russian) Uspekhi Mat. Nauk 44 (1989), no. 5(269), 13–60; translation in Russian Math. Surveys 44 (1989), no. 5, 11–71.

[17] Goldsheid, I. Ya.; Molchanov, S. A.; Pastur, L. A. A random homogeneous Schrödinger operator has a pure point spectrum. (Russian) Funkcional. Anal. i Priložen. 11 (1977), no. 1, 1–10, 96.

[18] Goldsheid, I.; Sodin, S. Lower bounds on Anderson-localised eigenfunctions on a strip. *To appear in* Comm. Math. Phys. `https://arxiv.org/abs/2012.03017`

[19] Gordon, A. Ya. Instability of dense point spectrum under finite rank perturbations. Comm. Math. Phys. 187 (1997), no. 3, 583–595.

[20] Gorodetski, A., Kleptsyn, V. Parametric Furstenberg theorem on random products of **SL(2, ℝ)** matrices. Adv. Math. 378 (2021), 107552.

[21] Han, Rui. Shnol's theorem and the spectrum of long range operators. Proc. Amer. Math. Soc. 147 (2019), no. 7, 2887–2897

[22] Jitomirskaya, Svetlana; Zhu, Xiaowen. Large deviations of the Lyapunov exponent and localization for the 1D Anderson model. Comm. Math. Phys. 370 (2019), no. 1, 311–324.

[23] Klein, Abel; Lacroix, Jean; Speis, Athanasios. Localization for the Anderson model on a strip with singular potentials. J. Funct. Anal. 94 (1990), no. 1, 135–155.

[24] Kotani, S. Lyapunov exponents and spectra for one-dimensional random Schrödinger operators. Random matrices and their applications (Brunswick, Maine, 1984), 277–286, Contemp. Math., 50, Amer. Math. Soc., Providence, RI, 1986.

[25] Kunz, Hervé; Souillard, Bernard. Sur le spectre des opérateurs aux différences finies aléatoires. Comm. Math. Phys. 78 (1980/81), no. 2, 201–246.

[26] Lacroix, Jean. Localisation pour l'opérateur de Schrödinger aléatoire dans un ruban. (French) [Localization for random Schrödinger operators in a strip] Ann. Inst. H. Poincaré Sect. A (N.S.) 40 (1984), no. 1, 97-116.

[27] Lacroix, Jean, The random Schrödinger operator in a strip. Probability measures on groups, VII (Oberwolfach, 1983), 280-297, Lecture Notes in Math., 1064, Springer, Berlin, 1984.

[28] Lacroix, Jean. L'opérateur de Schrödinger aléatoire dans un ruban. Diss. Paris 13, 1990, `https://www.lpsm.paris//pageperso/lacroix/ruban.pdf`

[29] Molchanov, S. A. Structure of the eigenfunctions of one-dimensional unordered structures. (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 42 (1978), no. 1, 70–103, 214.

[30] Rangamani, Nishant. Singular-unbounded random Jacobi matrices. J. Math. Phys. 60 (2019), no. 8, 081904.

[31] Shapiro, J. Incomplete localization for disordered chiral strips. `https://arxiv.org/abs/2108.10978`