



Schrödinger operators with potentials generated by hyperbolic transformations: I—positivity of the Lyapunov exponent

Artur Avila^{1,2} · David Damanik³ · Zhenghe Zhang⁴

Received: 7 December 2020 / Accepted: 25 August 2022 /

Published online: 5 September 2022

© The Author(s) 2022

Abstract We consider discrete one-dimensional Schrödinger operators whose potentials are generated by sampling along the orbits of a general hyperbolic transformation. Specifically, we show that if the sampling function is a non-constant Hölder continuous function defined on a subshift of finite type with a fully supported ergodic measure admitting a local product structure and a fixed point, then the Lyapunov exponent is positive away from a discrete set of energies. Moreover, for sampling functions in a residual subset of the space of Hölder continuous functions, the Lyapunov exponent is positive everywhere.

D.D. was supported in part by NSF grants DMS–1067988, DMS–1301582, DMS–1700131, and DMS–2054752, Simons Fellowship #669836, and by an Alexander von Humboldt Foundation research award.

Z. Z. was supported in part by NSF grant DMS–1764154.

✉ Zhenghe Zhang
zhenghe.zhang@ucr.edu

Artur Avila
artur.avila@math.uzh.ch

David Damanik
damanik@rice.edu

¹ Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

² IMPA, Estrada Dona Castorina, 110, Rio de Janeiro 22460-320, Brazil

³ Department of Mathematics, Rice University, Houston, TX 77005, USA

⁴ Department of Mathematics, University of California, Riverside, CA 92521, USA

If we consider locally constant or globally fiber bunched sampling functions, then the Lyapunov exponent is positive away from a finite set. Moreover, for sampling functions in an open and dense subset of the space in question, the Lyapunov exponent is uniformly positive. Our results can be applied to any subshift of finite type with ergodic measures that are equilibrium states of Hölder continuous potentials. In particular, we apply our results to Schrödinger operators defined over expanding maps on the unit circle, hyperbolic automorphisms of a finite-dimensional torus, and Markov chains.

Contents

1	Introduction	853
1.1	Statement of results	853
1.2	Strategy of proofs	858
2	Preliminaries	859
2.1	The setting	859
2.1.1	The base space and the base transformation	859
2.1.2	Measures with a local product structure	861
2.1.3	$SL(2, \mathbb{R})$ -cocycles and their projectivization	865
2.1.4	Reduction to a topologically mixing subshift	865
2.1.5	Stable and unstable holonomies	866
2.1.6	Invariant measures of projective cocycles	867
2.1.7	Schrödinger operators and cocycles	868
2.2	Periodic potentials	869
3	Large deviations	872
4	Invariance principle and conformal barycenter	881
4.1	Measurable holonomies resulting from small exponents	881
4.2	Invariance principle and su -states	889
4.3	Application of conformal barycenter	893
4.4	Local su -invariance	895
5	Positivity of the Lyapunov exponent I	899
5.1	General case: positivity away from a discrete set	899
5.2	Special cases: positivity away from a finite set	904
5.2.1	Fiber bunching and existence of holonomies	905
5.2.2	Locally constant cocycles	905
5.2.3	Energies admitting an su -state	906
5.2.4	Finiteness of the set of energies admitting su -states	907
5.2.5	Proof of Theorem 1.3	909
6	Positivity of the Lyapunov exponent II	910
6.1	Special cases: uniform positivity in a dense open set	910
6.2	General case: full positivity for generic sampling functions	914
7	Applications	917
	References	925

1 Introduction

1.1 Statement of results

In this series of papers, we are mainly concerned with the Anderson localization phenomenon for one-dimensional discrete Schrödinger operators H_ω in $\ell^2(\mathbb{Z})$ acting by

$$[H_\omega \psi](n) = \psi(n+1) + \psi(n-1) + V_\omega(n)\psi(n). \quad (1.1)$$

Here we assume Ω to be any compact metric space, $T : \Omega \rightarrow \Omega$ a homeomorphism, and $f : \Omega \rightarrow \mathbb{R}$ be continuous. We consider potentials $V_\omega : \mathbb{Z} \rightarrow \mathbb{R}$ defined by $V_\omega(n) = f(T^n \omega)$ for $\omega \in \Omega$ and $n \in \mathbb{Z}$. For general background on Schrödinger operators in $\ell^2(\mathbb{Z})$ with dynamically generated potentials of this form, we refer the reader to [19–21].

Spectral properties of the operators H_ω can be investigated by studying the behavior of the solutions to the difference equation

$$u(n+1) + u(n-1) + V_\omega(n)u(n) = Eu(n), \quad n \in \mathbb{Z} \quad (1.2)$$

with E real (or complex, depending on the problem in question). These solutions in turn can be described with the help of the Schrödinger cocycle (T, A^E) with the cocycle map $A^E : \Omega \rightarrow \mathrm{SL}(2, \mathbb{R})$ (resp., $\mathrm{SL}(2, \mathbb{C})$) being defined as

$$A^E(\omega) = A^{(E-f)}(\omega) := \begin{pmatrix} E - f(\omega) & -1 \\ 1 & 0 \end{pmatrix}, \quad (1.3)$$

where we often leave the dependence on $f : \Omega \rightarrow \mathbb{R}$ implicit as it will be fixed most of the time.

Such cocycles describe the transfer matrices associated with Schrödinger operators. Specifically, $u = u(n)$ solves (1.2) if and only if

$$\begin{pmatrix} u(n) \\ u(n-1) \end{pmatrix} = A_n^E(\omega) \begin{pmatrix} u(0) \\ u(-1) \end{pmatrix}, \quad n \in \mathbb{Z}, \quad (1.4)$$

where

$$A_n(\omega) = \begin{cases} A(T^{n-1}\omega) \dots A(\omega), & n \geq 1; \\ [A_{-n}(T^n\omega)]^{-1}, & n \leq -1, \end{cases} \quad (1.5)$$

and we set $A_0(\omega)$ to be the identity matrix.

The Lyapunov exponent (LE) of the Schrödinger cocycle plays a key role in the spectral analysis of the operators. Let μ be a T -ergodic probability measure

on Ω . The Lyapunov exponent is given by

$$\begin{aligned} L(A^E, \mu) &= \lim_{n \rightarrow \infty} \frac{1}{n} \int \log \|A_n^E(\omega)\| d\mu(\omega) \\ &= \inf_{n \geq 1} \frac{1}{n} \int \log \|A_n^E(\omega)\| d\mu(\omega). \end{aligned} \quad (1.6)$$

For simplicity, we write $L(E) = L(A^E, \mu)$. By Kingman's subadditive ergodic theorem, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_n^E(\omega)\| = L(E)$$

for μ -almost every $\omega \in \Omega$. In particular, certain uniform positivity and uniform large deviation estimates (LDT) for the LE are strong indications of Anderson localization, which in its spectral formulation states that for μ -almost every $\omega \in \Omega$, the operator H_ω has pure point spectrum with exponentially decaying eigenfunctions.

On the other hand, positivity and LDT estimates for the LE are extensively studied topics in dynamical systems. In general, the more random the base dynamics (Ω, T, μ) is, the more likely it is that one has positivity and LDT for the LE. For instance, for the well-known Anderson model, where V_ω is a realization of independent identically distributed random variables, one does have uniform positivity and uniform LDT on any compact set of energies E . These are classic results that go back to the seminal work of Furstenberg [27]. Combining this with a certain elimination of double resonance argument, these two properties indeed lead to a localization result for the Anderson model; see, for example, [16] for recent proofs of all these results mentioned above.

The Anderson model may be put into the context of the present paper as follows. We consider the Anderson model whose single site measure is an atomic measure supported on a finite number of points, which is the most difficult case. Let $\mathcal{A} = \{1, 2, \dots, \ell\}$ with $\ell \geq 2$ and let $\tilde{\mu}$ be a fully supported probability measure on $\mathcal{A}$. Let $\Omega = \mathcal{A}^{\mathbb{Z}}$ be the full shift space and consider the left shift $T : \Omega \rightarrow \Omega$ defined by $(T\omega)_n = \omega_{n+1}$ for $\omega \in \mathcal{A}^{\mathbb{Z}}$ and $n \in \mathbb{Z}$. Let $\mu = \tilde{\mu}^{\mathbb{Z}}$, which is strongly mixing with respect to T . The Anderson model may be generated by setting $V_\omega = f(T^n \omega)$ where $f : \Omega \rightarrow \mathbb{R}$ depends only on ω_0 . The potentials generated in this way are the most random among the potentials studied in this paper. It is natural to ask what can be said if the potentials, or rather the base dynamics (Ω, T, μ) , are less random. In the language of mathematical physics, what if the V_ω 's are weakly correlated? Or in the language of dynamical systems, what if (Ω, T, μ) is a mixing system such as the Arnold cat map or the doubling map? Or more generally, a subshift

of finite type with measure of maximal entropy? It turns out that such systems are much more difficult to analyze.

To further explain what this paper accomplishes, we consider a general framework of the base dynamics that includes most of the systems mentioned above as special classes. Let (Ω, T) be a subshift of finite type. Let μ be a T -ergodic measure that is fully supported on Ω . Moreover, we further assume that μ admits a local product structure (a detailed definition may be found in Sect. 2.1.2). Let $f : \Omega \rightarrow \mathbb{R}$ be α -Hölder continuous for some $0 < \alpha \leq 1$ and non-constant. We define

$$\mathcal{Z}_f = \{E : L(E) = 0\}. \quad (1.7)$$

In the present paper, we address the following question.

Problem *Let Ω , T , μ , and f be described as above. How large is $\mathcal{Z}_f$? In particular, when is it discrete, finite, or even empty?*

Note that the discreteness of $\mathcal{Z}_f$ can be taken as a starting point to show full spectral localization for the corresponding operators; see, for example, the proof of localization in [16, Proof of Theorem 1.3]. We comment on this point in more detail in Remark 1.6 below. Earlier partial results along this line may be found for example in [6, 9, 12, 18, 22, 35, 36, 44], where either the base dynamics or the choice of f are quite restricted, or $\mathcal{Z}_f$ is still quite large. The main theorem of this paper is:

Theorem 1.1 *Suppose (Ω, T) is a subshift of finite type and μ is a fully supported T -ergodic probability measure that has a local product structure that is fully supported on Ω . Suppose T has a fixed point and f is Hölder continuous and non-constant. Then the set $\mathcal{Z}_f$ is discrete.*

Remark 1.2 Let us mention that the concept of local product structure is recalled in detail in Sect. 2.1. If μ has a local product structure, then its topological support $\text{supp } \mu$ is a subshift of finite type (see, e.g., [7, Lemma 1.2]) and hence the assumption in Theorem 1.1 that μ is fully supported is not a restriction. If the support is not the whole space, then we can replace (Ω, T) by $(\text{supp } \mu, T_{\text{supp } \mu})$. This remark applies whenever we assume in this paper that μ is fully supported. Conversely, given any subshift of finite type, the unique equilibrium state associated with a Hölder continuous potential always has a local product structure, see [14, 34] or [8, Section 2.2]. In particular, measures with maximal entropy do have a local product structure.

It is clear that we can add a coupling constant λ to f in the statement of Theorem 1.1. This further indicates that such systems do behave like the Anderson model, as the Anderson model is always localized as long as $\lambda > 0$. If we restrict the choice of f so that it is locally constant or $\|f\|_\infty$ is small,

then we can improve the result as follows. Let $C^\alpha(\Omega, \mathbb{R})$, $0 < \alpha \leq 1$ be the space of α -Hölder continuous functions.

Theorem 1.3 *Let (Ω, T, μ) be as in Theorem 1.1. Suppose $f \in C^\alpha(\Omega, \mathbb{R})$ is globally bunched or locally constant. Assume further that f is non-constant and T has a fixed point. Then $\mathcal{Z}_f$ is finite.*

A detailed definition of global bunching may be found at Sect. 5.2. In particular, f is globally bunched if $\|f\|_\infty$ is small. A possible explicit choice of a smallness condition on $\|f\|_\infty$ may be found in (7.7). We can again add a coupling constant λ to f in Theorem 1.3 if f is locally constant. If f is globally bunched, then as λ becomes large, $\mathcal{Z}_{\lambda f}$ might become a discrete set that is no longer guaranteed to be finite. This is because we will lose global bunching as λ becomes large and we have to apply Theorem 1.1 then. In Sect. 7, we shall show that Theorems 1.1 and 1.3 are sharp in the sense that $\mathcal{Z}_f$ may indeed be nonempty for a suitable locally constant f . Thus another natural question is: when can we remove the discrete or finite set $\mathcal{Z}_f$? We have the following results.

Theorem 1.4 *Suppose (Ω, T) is a subshift of finite type and μ is a fully supported T -ergodic measure that has a local product structure. Then there is a residual subset $\mathcal{G}$ of $C^\alpha(\Omega, \mathbb{R})$ such that for each $f \in \mathcal{G}$, $\mathcal{Z}_f$ is empty.*

Again, if we restrict the choice of f so that it is either locally constant or globally bunched, then we can obtain a uniform lower bound of the $L(E)$ for an even wider class of choices:

Theorem 1.5 *Suppose (Ω, T) is a subshift of finite type and μ is a fully supported T -ergodic measure that has a local product structure. Consider the subspaces of $C^\alpha(\Omega, \mathbb{R})$ consisting of globally bunched or locally constant functions. For each of them, there is an open and dense subset $\mathcal{G}$ such that for every $f \in \mathcal{G}$, we have $\inf\{L(E) : E \in \mathbb{R}\} > 0$.*

Applications of Theorems 1.1–1.5 to more concrete base dynamics such as the doubling map, Arnold's cat map, and Markov shifts may be found in Sect. 7.

Remark 1.6 (a) Let us emphasize that from the perspective of a spectral analysis of the operator family $\{H_\omega\}_{\omega \in \Omega}$, and in particular when seeking a proof of spectral localization for this family, the discreteness of $\mathcal{Z}_f$ is in general the appropriate first milestone towards the eventual goal. It then needs to be combined with control of the Lyapunov exponent away from $\mathcal{Z}_f$ (the connected components of $\mathcal{Z}_f^c$ need to be exhausted by intervals on which the Lyapunov exponent is uniformly bounded away from zero;

this is often established by proving the continuity of $L(E)$ in E whenever possible), suitable large deviation estimates, and an argument that rules out the presence of infinitely many double resonances for almost every ω . It then follows for μ -almost every $\omega \in \Omega$ that spectrally almost every energy in $\mathcal{Z}_f^c$ admits an exponentially decaying eigenfunction for H_ω . As the discrete set $\mathcal{Z}_f$ almost surely carries no weight with respect to the spectral measures of H_ω , this then shows that for μ -almost every $\omega \in \Omega$, the operator H_ω admits a basis consisting of exponentially decaying eigenfunctions, and the desired spectral localization statement then follows.

- (b) One is nevertheless interested in obtaining stronger results on the size of the exceptional set $\mathcal{Z}_f$, such as finiteness or emptiness, whenever possible, as this leads to stronger versions of the dynamical version of an Anderson localization statement. Here, one is interested in showing that the solutions of the time-dependent Schrödinger equation $i\partial_t\psi = H_\omega\psi$ are localized. In other words, one seeks to prove good off-diagonal estimates for the matrix elements of e^{-itH_ω} relative to the standard basis of $\ell^2(\mathbb{Z})$, uniformly in the time parameter t . Energies E in $\mathcal{Z}_f$ present an obstacle for proving this and one generally simply projects away from these exceptional energies and considers $\chi_I(H_\omega)e^{-itH_\omega}$ with a set $I \subseteq \mathcal{Z}_f^c$ that has positive distance from $\mathcal{Z}_f$. In fact, it has been shown that dynamical localization can actually fail, even when spectral localization holds, if one does not project away from $\mathcal{Z}_f$; compare, for example, [23, 29]. Clearly, it is then desirable to show that $\mathcal{Z}_f$ is empty whenever this can be expected to be true. Of course, as pointed out earlier, this will not always be the case.
- (c) Let us emphasize that the road map to spectral localization described in part (a) of this remark is applicable in the general setting of ergodic Schrödinger operators, and it has been implemented for special cases ranging from the Anderson model to potentials generated by torus translations, the standard skew-shift, the doubling map, or the Arnold cat map. While the literature is vast, let us just mention a few representative papers, [10–12, 16], and refer the reader to [19–21] for more information. Regarding the base transformations considered in this paper, the absence of a suitable general and global result showing the discreteness of $\mathcal{Z}_f$ was the primary obstacle in attempting to implement this road map. Thus, the present paper fills precisely this gap and opens the door to a localization proof, which we intend to work out in detail in the second part of this series [1].

1.2 Strategy of proofs

One of the main tools we use to prove our results is the so-called *invariance principle* as coined in [3]. The first version of the invariance principle goes back to Ledrappier [33] and it was later generalized in [3]. The version we adopt in this paper is due to Bonatti et al. [7, 39]. A detailed statement of the invariance principle may be found in Proposition 4.5. It says that if the Lyapunov exponent $L(A, \mu)$ of a cocycle (T, A) is 0 and A depends only on the future or the past, then any (T, A) -invariant measure m on $\Omega \times \mathbb{RP}^1$ admits a disintegration $\{m_\omega : \omega \in \Omega\}$ that depends only on the future or the past, respectively.

Another main tool we use is given by the so-called stable and unstable holonomies, which are defined along the stable or unstable sets of ω , respectively; see Sect. 2.1.5 for a detailed definition. If $L(A, \mu) = 0$, we can define a measurable family of stable and unstable holonomies for μ -almost every ω . Then one can use the stable or unstable holonomies to conjugate the cocycle (T, A) to one that depends only on the future or the past, respectively.

Combining the two steps above, one can show that the family of invariant measures $\{m_\omega\}$ are invariant with respect to the stable and unstable holonomies as well. We call such a family an *su-state*.

It turns out that the existence of *su-states* is a very rare event in the sense that they can be easily perturbed away by modifying the data of the cocycle map A at certain periodic points. Roughly speaking, this is how [7, 8, 39] show the positivity of the Lyapunov exponent for certain typical C^α -cocycles. More precisely, [7, 8] did it in case the cocycle is fiber bunched or is locally constant while [39] did it for the general case.

However, to prove Theorems 1.1–1.3, we need to consider Schrödinger cocycles with fixed sampling functions. They are basically fixed cocycle maps parametrized by the energy parameter $E \in \mathbb{R}$. So we are not allowed to perturb the cocycle maps to get typicality. Hence, the above strategy is not sufficient to yield the discreteness or finiteness of $\mathcal{Z}_f$ as stated in Sect. 1.1. It turns out that in addition we need to deploy certain tools from spectral theory. In particular, we will consider the spectra associated with certain periodic orbits and invoke a result from inverse spectral theory for periodic operators. Moreover, to make use of the periodic data, among other things, we also need to show that periodic orbits with small Lyapunov exponent belong to the topological support of the sets where one can define continuous holonomies. Finally, to use the periodic data to prove the main results, we have to combine the conformal barycenter concept due to Douady and Earle [25], Bowen's specification property [13], and Kalinin's theorem regarding approximating $L(E)$ by the Lyapunov exponent along periodic orbits [30]. In short, the proof

is based on a fusion of ideas and results from both dynamical systems and spectral theory.

The structure of the remainder of the paper is as follows. In Sect. 2, we state some necessary preliminaries and lay out our context. In Sect. 3, we give a proof of an additive version of a large deviation estimate for Hölder continuous functions defined on Ω and for slightly more restricted measures μ . These large deviation estimates may be of independent interest. Moreover, they will play a key role in the second paper of this series [1]. In Sect. 4, we introduce our main tools such as the invariance principle and the conformal barycenter, and we also give detailed proofs of certain lemmas. We prove Theorems 1.1–1.3 in Sect. 5 and Theorems 1.4–1.5 in Sect. 6. In Sect. 7, we apply our general Theorems 1.1–1.5 to several concrete models such as the doubling map, Arnold cat map, and Markov chains. In particular, the class of Markov chains includes general locally constant Schrödinger potentials defined on the full shift space as a special case, which yields a generalization of the classical Furstenberg theorem. Many of the results are the first of their kind. We also compute an explicit choice of $\lambda_0 > 0$ so that $\|f\|_\infty \leq \lambda_0$ is sufficient for f to be globally bunched. Finally, we present an example where we show the finite set $\mathcal{Z}_f$ appearing in the statement of Theorem 1.3 may not be removed in general, so that our results are sharp in a suitable sense.

2 Preliminaries

2.1 The setting

In this section we describe the setting we will work in. We have chosen subshifts of finite type with appropriate ergodic measures as base transformations as a compromise between concreteness and generality. Other possible choices would have been concrete classes of smooth hyperbolic transformations and expanding maps. For background and discussion of the material presented below, we refer the reader to [7, 8, 39].

2.1.1 The base space and the base transformation

Let $\mathcal{A} = \{1, 2, \dots, \ell\}$ with $\ell \geq 2$ be equipped with the discrete topology. Consider the product space $\mathcal{A}^{\mathbb{Z}}$, whose topology is generated by the cylinder sets, which are the sets of the form

$$[n; j_0, \dots, j_k] = \{\omega \in \mathcal{A}^{\mathbb{Z}} : \omega_{n+i} = j_i, 0 \leq i \leq k\}$$

with $n \in \mathbb{Z}$ and $j_0, \dots, j_k \in \mathcal{A}$. The topology is metrizable and for definiteness we fix the following metric d on $\mathcal{A}^{\mathbb{Z}}$. Set $d(\omega, \omega) = 0$ for $\omega \in \mathcal{A}^{\mathbb{Z}}$ and

$$d(\omega, \tilde{\omega}) = e^{-N(\omega, \tilde{\omega})} \quad (2.1)$$

for $\omega, \tilde{\omega} \in \mathcal{A}^{\mathbb{Z}}$ with $\omega \neq \tilde{\omega}$, where

$$N(\omega, \tilde{\omega}) = \max\{N \geq 0 : \omega_n = \tilde{\omega}_n \text{ for all } |n| < N\}. \quad (2.2)$$

We consider the left shift $T : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ defined by $(T\omega)_n = \omega_{n+1}$ for $\omega \in \mathcal{A}^{\mathbb{Z}}$ and $n \in \mathbb{Z}$. Let $\text{Orb}(\omega) = \{T^n \omega : n \in \mathbb{Z}\}$ be the orbit of ω under the dynamics T .

Definition 2.1 Let $Q = (q_{ij})_{1 \leq i, j \leq \ell}$ be an $\ell \times \ell$ matrix with $q_{ij} \in \{0, 1\}$ and let Ω be the subshift of finite type associated to the matrix Q ,

$$\Omega = \{(\omega_n)_{n \in \mathbb{Z}} : q_{\omega_n \omega_{n+1}} = 1 \text{ for all } n \in \mathbb{Z}\}.$$

Consider the topological dynamical system (Ω, T) .

We say that a finite word $j_0 j_1 \dots j_k$, where $j_i \in \{1, \dots, \ell\}$ for $0 \leq i \leq k$, is *admissible* if it occurs in some $\omega \in \Omega$, that is, there are $\omega \in \Omega$ and $n \in \mathbb{Z}$ such that $\omega_{n+i} = j_i$ for all $0 \leq i \leq k$.

The *local stable set* of a point $\omega \in \Omega$ is defined by

$$W_{\text{loc}}^s(\omega) = \{\tilde{\omega} \in \Omega : \omega_n = \tilde{\omega}_n \text{ for } n \geq 0\} \quad (2.3)$$

and the *local unstable set* of ω is defined by

$$W_{\text{loc}}^u(\omega) = \{\tilde{\omega} \in \Omega : \omega_n = \tilde{\omega}_n \text{ for } n \leq 0\}. \quad (2.4)$$

A set is called *s-locally saturated* (resp., *u-locally saturated*) if it is a union of local stable (resp., local unstable) sets of the form above.

For each $j \in \mathcal{A}$ and each pair of points $\omega, \tilde{\omega} \in [0; j]$, we denote the unique point in $W_{\text{loc}}^u(\omega) \cap W_{\text{loc}}^s(\tilde{\omega})$ by $\omega \wedge \tilde{\omega}$. Throughout this paper, we fix for each $1 \leq j \leq \ell$ a choice of $\omega^{(j)} \in [0; j]$ so that the maps

$$\omega \mapsto \omega^{(\omega_0)} \wedge \omega, \quad \omega \mapsto \omega \wedge \omega^{(\omega_0)} \quad (2.5)$$

are well-defined and continuous on Ω and are constant on local stable and unstable sets, respectively.

2.1.2 Measures with a local product structure

Let the subshift Ω be equipped with the Borel σ -algebra and let μ be a probability measure on Ω that is ergodic with respect to T . We define

$$\Omega^+ = \{(\omega_n)_{n \geq 0} : \omega \in \Omega\}, \quad (2.6)$$

$$\Omega^- = \{(\omega_n)_{n \leq 0} : \omega \in \Omega\} \quad (2.7)$$

to be the spaces of one-sided right and left infinite sequences, respectively, associated with Ω . Metrics for $\Omega^\pm$ can be defined in a way similar to the definition of the metric for $\mathcal{A}^\mathbb{Z}$ in Sect. 2.1.1. Abusing notation slightly, we still let d denote their metrics. Let π^+ be the projection from Ω to Ω^+ and $\mu^+ = \pi_*^+(\mu)$ be the pushforward measure of μ on Ω^+ . Similarly, we let π^- be the projection to Ω^- and μ^- be the pushforward measure on Ω^- . Let T_+ be the left shift operator on Ω^+ and T_- be the right shift on Ω^- . For $n \geq 0$, we let $[n; j_0, \dots, j_k]^+$ denote the cylinder sets in Ω^+ ; for $n \leq -k$, we let $[n; j_0, \dots, j_k]^-$ denote the cylinder sets in Ω^- . Let $\omega^\pm$ denote points in $\Omega^\pm$, respectively.

For simplicity, for each $1 \leq j \leq \ell$, we set $\mu_j = \mu|_{[0; j]}$. Similarly, we set $\mu_j^\pm = \mu^\pm|_{[0; j]^\pm}$, respectively.

Note that we do not have $\Omega = \Omega^- \times \Omega^+$. However, for each $1 \leq j \leq \ell$ we have a natural homeomorphism

$$P : [0; j] \rightarrow [0; j]^- \times [0; j]^+ \text{ where } P(\omega) = (\pi^-\omega, \pi^+\omega).$$

Thus, abusing the notation a bit, we may just write $[0; j] = [0; j]^- \times [0; j]^+$. Moreover, we have for all $\omega \in \Omega$,

$$(\pi^+)^{-1}(\pi^+\omega) = W_{\text{loc}}^s(\omega), \quad (\pi^-)^{-1}(\pi^-\omega) = W_{\text{loc}}^u(\omega). \quad (2.8)$$

Definition 2.2 We say μ has a *local product structure* if there is a $\psi : \Omega \rightarrow (0, \infty)$ such that for each $1 \leq j \leq \ell$, $\psi \in L^1([0; j], \mu_j^- \times \mu_j^+)$ and

$$d\mu_j = \psi \cdot d(\mu_j^- \times \mu_j^+). \quad (2.9)$$

The local product structure of μ amounts to saying that $\mu_j^- \times \mu_j^+$ is equivalent to μ_j . Indeed, (2.9) clearly implies that μ_j is absolutely continuous with respect to $\mu_j^- \times \mu_j^+$. On the other hand, if $\mu_j(E) = 0$, then we must have $(\mu_j^- \times \mu_j^+)(E) = 0$ since $\psi(\omega) > 0$ for all $\omega \in \Omega$. In particular, we may draw the following conclusion. If $E \subset [0; j]$ is u -locally saturated with $\mu(E) > 0$

and $F \subset [0; j]$ is s -locally saturated with $\mu(F) > 0$, we have

$$\begin{aligned}(\mu_j^- \times \mu_j^+)(E \cap F) &= (\mu_j^- \times \mu_j^+)(\pi^- E \times \pi^+ F) \\&= \mu_j^-(\pi^- E) \cdot \mu_j^+(\pi^+ F) \\&= \mu(E) \cdot \mu(F) \\&> 0,\end{aligned}$$

which implies that

$$\mu(E \cap F) = \mu_j(E \cap F) > 0. \quad (2.10)$$

Conversely, if $\mu_j^- \times \mu_j^+$ is equivalent to μ_j for each $1 \leq j \leq \ell$, then $d\mu_j = \psi \cdot d(\mu_j^- \times \mu_j^+)$, where $\psi \in L^1([0; j], \mu_j^- \times \mu_j^+)$ is the Radon–Nikodym derivative of μ_j with respect to $\mu_j^- \times \mu_j^+$. Note $1/\psi \in L^1([0; j], \mu_j)$ is the Radon–Nikodym derivative of $\mu_j^- \times \mu_j^+$ with respect to μ_j . Hence we must have that $\psi(\omega) > 0$ for all j and for μ_j -a.e. ω . We can of course modify ψ so that it is positive everywhere.

Definition 2.3 A *Jacobian* of the measure μ^+ with respect to T_+ on Ω^+ is a measurable function $J_+ : \Omega^+ \rightarrow \mathbb{R}_+$ such for each $j \in \{1, \dots, \ell\}$, we have

$$d\mu^+(T_+\omega^+) = J_+(\omega^+) \cdot d((T_+)_*(\mu^+|_{[0; j]^+}))(T_+\omega^+). \quad (2.11)$$

A *Jacobian* of μ^- with respect to T_- can be defined similarly.

One consequence of the local product structure of μ is that $\mu^\pm$ admit Jacobians with respect to $T_\pm$ on $[0; j]^\pm$ for each $1 \leq j \leq \ell$, respectively. The following lemma is essentially contained in [8]. While in [8, Lemma 2.2], ψ is assumed to be continuous, we note that the same proof can be applied to obtain the following lemma.

Lemma 2.4 The measures $\mu^\pm$ admit positive Jacobians $J_\pm \in L^1([0; j]^\pm, d\mu_j^\pm)$ with respect to $T_\pm$ on $[0; j]^\pm$, respectively, for each $1 \leq j \leq \ell$.

For $\underline{l} = (l_1, \dots, l_n) \in \{1, \dots, \ell\}^n$, we write the cylinder $[0; l_1, \dots, l_n, j]$ as $[0; \underline{l}, j]$ and set $|\underline{l}| := n$. We use a similar notation for spaces of one-sided sequences. For a cylinder $[0; \underline{l}, j]^+ \subset \Omega^+$, we clearly have a Jacobian for $T_+^{|\underline{l}|} : [0; \underline{l}, j]^+ \rightarrow [0; j]^+$, which is denoted by $J_+^{(\underline{l}, j)} : [0; l_1, \dots, l_n, j]^+ \rightarrow (0, \infty)$ and is given by the formula

$$J_+^{(\underline{l}, j)}(\omega^+) = \prod_{k=0}^{n-1} J_+(T_+^k \omega^+).$$

By the definition of a Jacobian, we have for any integrable function $f : \Omega^+ \rightarrow \mathbb{R}$ and any $[0; \underline{l}, j]^+ \subset \Omega^+$ that

$$\int_{[0; j]^+} f(\eta) d\mu^+(\eta) = \int_{[0; \underline{l}, j]^+} f(T_+^{|\underline{l}|} \omega^+) J_+^{(\underline{l}, j)}(\omega^+) d\mu^+(\omega^+). \quad (2.12)$$

We first have the following immediate consequence of Lemma 2.4, which will be used in Sect. 5.

Corollary 2.5 *Let $D \subset \Omega^+$ be such that $\mu^+(D \cap [0; j]^+) > 0$. Then for all $[0; \underline{l}, j]^+ \subset \Omega^+$, we have*

$$\mu^+(T_+^{-|\underline{l}|}(D) \cap [0; \underline{l}, j]^+) > 0.$$

Similarly, if $\mu^-(D \cap [0; j]^-) > 0$ for some $D \subset \Omega^-$, then for all $[-|\underline{l}|; j, \underline{l}]^- \subset \Omega^-$, we have

$$\mu^-(T_-^{-|\underline{l}|}(D) \cap [-|\underline{l}|; j, \underline{l}]^-) > 0.$$

Proof We only consider the case for (Ω^+, T_+, μ^+) ; the case with (Ω^-, T_-, μ^-) can be handled similarly.

Without loss of generality, we may just consider a Borel set $D \subset [0; j]^+$ with positive measure. By (2.12), we have

$$\begin{aligned} 0 &< \mu^+(D) \\ &= \int_{[0; j]} \chi_D(\eta) d\mu^+(\eta) \\ &= \int_{[0; \underline{l}, j]} \chi_D(T_+^{|\underline{l}|} \omega^+) J_+^{(\underline{l}, j)}(\omega^+) d\mu^+(\omega^+) \\ &= \int_{[0; \underline{l}, j] \cap (T_+^{-|\underline{l}|} D)} J_+^{(\underline{l}, j)}(\omega^+) d\mu^+(\omega^+), \end{aligned}$$

which implies that $\mu^+([0; \underline{l}, j] \cap (T_+^{-|\underline{l}|} D)) > 0$. □

For some results we will need the measure μ to obey a quantitative version of local product structure, which is defined as follows.

Definition 2.6 We say that μ satisfies the *bounded distortion property* if there is $C \geq 1$ such that for all cylinders $[n; j_0, \dots, j_k] \subset \Omega$ and $[l; i_0, \dots, i_m] \subset \Omega$, where $l > n + k$ and $[n; j_0, \dots, j_k] \cap [l; i_0, \dots, i_m] \neq \emptyset$, we have

$$C^{-1} \leq \frac{\mu([n; j_0, \dots, j_k] \cap [l; i_0, \dots, i_m])}{\mu([n; j_0, \dots, j_k]) \cdot \mu([l; i_0, \dots, i_m])} \leq C. \quad (2.13)$$

Note that by T -invariance of μ and by the definition of $\mu^\pm$, μ has the bounded distortion property if and only if μ^+ or μ^- has the bounded distortion property. For instance, the bounded distortion property of μ^+ means that for all $n \geq 0$, $l > n + k$, and $[n; j_0, \dots, j_k]^+ \cap [l; i_0, \dots, i_m]^+ \neq \emptyset$, we have

$$C^{-1} \leq \frac{\mu^+([n; j_0, \dots, j_k]^+ \cap [l; i_1, \dots, i_m]^+)}{\mu^+([n; j_0, \dots, j_k]^+) \cdot \mu^+([l; i_1, \dots, i_m]^+)} \leq C. \quad (2.14)$$

In fact, given any subshift of finite type, the unique equilibrium state associated with a Hölder continuous potential always has the bounded distortion property; see Lemma 3.4.

It is not difficult to see that every measure satisfying the bounded distortion property has a local product structure. Indeed, for every cylinder $[-k; j_{-k}, \dots, j_{-1}, j_0, \dots, j_k] \subset \Omega$, we have by (2.13)

$$\begin{aligned} & (\mu_{j_0}^- \times \mu_{j_0}^+)([-k; j_{-k}, \dots, j_{-1}, j_0, \dots, j_m]) \\ &= \mu_{j_0}^-([-k; j_{-k}, \dots, j_{-1}, j_0]^-) \cdot \mu_{j_0}^+([0; j_0, \dots, j_m]^+) \\ &= \mu([-k; j_{-k}, \dots, j_{-1}, j_0]) \cdot \mu([0; j_0, \dots, j_m]) \\ &\leq \mu([-k; j_{-k}, \dots, j_{-1}]) \cdot \mu([0; j_0, \dots, j_m]) \\ &\leq C \mu([-k; j_{-k}, \dots, j_{-1}, j_0, \dots, j_m]) \\ &= C \mu_{j_0}([-k; j_{-k}, \dots, j_{-1}, j_0, \dots, j_m]). \end{aligned}$$

Similarly, we can obtain such estimates for all other cylinders. Since every Borel set can be approximated by cylinder sets, these estimates clearly imply that $\mu_j^- \times \mu_j^+$ is absolutely continuous with respect to μ_j . On the other hand,

$$\begin{aligned} & \mu_{j_0}([-k; j_{-k}, \dots, j_{-1}, j_0, \dots, j_m]) \\ &= \mu_{j_0}([-k; j_{-k}, \dots, j_{-1}, j_0, \dots, j_m]) \\ &\leq C \mu_{j_0}^-([-k; j_{-k}, \dots, j_{-1}]^-) \cdot \mu_{j_0}^+([0; j_0, \dots, j_m]^+) \\ &\leq \frac{C^2}{\mu_{j_0}^-([0; j_0]^-)} \mu_{j_0}^-([-k; j_{-k}, \dots, j_{-1}, j_0]^-) \cdot \mu_{j_0}^+([0; j_0, \dots, j_m]^+) \\ &\leq \tilde{C}(\mu_{j_0}^- \times \mu_{j_0}^+)([-k; j_{-k}, \dots, j_{-1}, j_0, \dots, j_m]), \end{aligned}$$

where $\tilde{C} = \max\{\frac{C^2}{\mu([0; j_0]^-)} : 1 \leq j \leq \ell\}$ is independent of the choice of the cylinder sets. Note in the fourth line above, we use the bounded distortion property of μ^- as:

$$\mu_{j_0}^-([-k; j_{-k}, \dots, j_{-1}]^-) \cdot \mu_{j_0}^-([0; j_0]^-) \leq C \mu_{j_0}^-([-k; j_{-k}, \dots, j_{-1}, j_0]^-).$$

It clearly implies that μ_j is absolutely continuous with respect to $\mu_j^- \times \mu_j^+$ for each $1 \leq j \leq \ell$.

2.1.3 $\mathrm{SL}(2, \mathbb{R})$ -cocycles and their projectivization

A continuous map $A : \Omega \rightarrow \mathrm{SL}(2, \mathbb{R})$ gives rise to the cocycle $(T, A) : \Omega \times \mathbb{R}^2 \rightarrow \Omega \times \mathbb{R}^2$, $(\omega, v) \mapsto (T\omega, A(\omega)v)$. For $n \in \mathbb{Z}$, we let $(T, A)^n = (T^n, A_n)$. In particular, we have

$$A_n(\omega) = \begin{cases} A(T^{n-1}\omega) \cdots A(\omega), & n \geq 1; \\ I_2, & n = 0; \\ [A_{-n}(T^n\omega)]^{-1}, & n \leq -1, \end{cases}$$

where I_2 is the identity matrix. Now let μ be a T -ergodic probability measure with topological support equal to Ω . The Lyapunov exponent is given by

$$\begin{aligned} L(A, \mu) &= \lim_{n \rightarrow \infty} \frac{1}{n} \int \log \|A_n(\omega)\| d\mu(\omega) \\ &= \inf_{n \geq 1} \frac{1}{n} \int \log \|A_n(\omega)\| d\mu(\omega). \end{aligned}$$

By Kingman's subadditive ergodic theorem, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_n(\omega)\| = L(A, \mu)$$

for μ -a.e. ω . By linearity and invertibility of each $A(\omega)$, we can projectivize the second component and consider $(T, A) : \Omega \times \mathbb{RP}^1 \rightarrow \Omega \times \mathbb{RP}^1$.

2.1.4 Reduction to a topologically mixing subshift

We need to reduce to the case where $T : \Omega \rightarrow \Omega$ is topologically mixing and collect some standard facts. One may find a detailed discussion of the results stated in this section in [31, Section 1.9].

One says that (Ω, T) is *topologically mixing* if for any pair of nonempty open sets $U, V \subset \Omega$, there is an $N \geq 1$ such that $T^n(U) \cap V \neq \emptyset$ for all $n \geq N$.

Note that a general subshift of finite type (Ω, T) might not be topologically mixing. But any subshift of finite type that has a dense positive semiorbit has the following decomposition. By the spectral decomposition theorem for hyperbolic basic sets, we can decompose Ω as $\Omega = \bigsqcup_{l=1}^s \Omega_l$ for some $s \geq 1$ and for closed subsets Ω_l , so that the following holds true: $T(\Omega_l) = \Omega_{l+1}$ for $1 \leq l < s$ and $T(\Omega_s) = (\Omega_1)$, and $T^s|_{\Omega_l}$ is a topologically mixing subshift

of finite type for each $1 \leq l \leq s$. In particular, if (Ω, T) has a fully supported ergodic measure, then (Ω, T) has such a decomposition. Moreover, the normalized restriction μ_l of μ to Ω_l is a T^s -invariant ergodic, fully supported probability measure with local product structure or bounded distortion property, provided the same property is true for μ on Ω . One may also see [17, Section 3.2] for such facts.

Then for a cocycle map $A : \Omega \rightarrow \mathrm{SL}(2, \mathbb{R})$, we consider $A_s : \Omega_l \rightarrow \mathrm{SL}(2, \mathbb{R})$ as $A_s(\omega)$, which may be considered a cocycle map defined over the base dynamics $T^s : \Omega_l \rightarrow \Omega_l$. Clearly, $L(A_s, \mu_l) > 0$ for some $1 \leq l \leq s$ implies that $L(A, \mu) > 0$. Since the present paper is only concerned with the positivity of the Lyapunov exponent, we assume from now on that (Ω, T) is topologically mixing.

Note that $\mathrm{supp}(\mu) = \Omega$ and ergodicity of μ together already imply that $\overline{\mathrm{Orb}(\omega)} = \Omega$ for μ -almost every $\omega \in \Omega$.

Topological mixing has additional consequences, which are needed in the present paper. First, it implies that the set of periodic orbits is dense in Ω . Moreover, we have the following more quantitative behavior of periodic points, which is called the *specification property*. It concerns shadowing finite pieces of segments of different orbits by a single orbit, in particular, by a periodic orbit. It was first introduced by Bowen [13]. The following version for subshifts of finite type is due to Sigmund [37]. For $a < b \in \mathbb{Z}$, we let $[a, b] \subset \mathbb{Z}$ denote the indicated interval of integers. In other words, $[a, b] = \{n \in \mathbb{Z} : a \leq n \leq b\}$.

Proposition 2.7 *Let (Ω, T) be a topologically mixing subshift of finite type. For each $\epsilon > 0$, there is an integer $r = r(\epsilon) > 0$ such that for any choice of points $p^{(i)} \in \Omega$ and intervals of integers $I_i = [a_i, b_i]$, $i = 1, 2$, with $a_2 - b_1 > r$ and any $n > b_2 - a_1 + r$, there exists a periodic point p with period n such that*

$$d(T^j p, T^j p^{(i)}) < \epsilon \text{ for } j \in I_i, i = 1, 2.$$

2.1.5 Stable and unstable holonomies

Given (Ω, T, μ) as above, consider $A : \Omega \rightarrow \mathrm{SL}(2, \mathbb{R})$ and the projective cocycle $(T, A) : \Omega \times \mathbb{RP}^1 \rightarrow \Omega \times \mathbb{RP}^1$. We will denote the fiber $\{\omega\} \times \mathbb{RP}^1$ by $\mathcal{E}_\omega$.

Definition 2.8 A *stable holonomy* h^s for A is a family of homeomorphisms $h_{\omega, \omega'}^s : \mathcal{E}_\omega \rightarrow \mathcal{E}_{\omega'}$, defined whenever ω and ω' belong to the same local stable set, satisfying the following properties:

- (i) $h_{\omega', \omega''}^s \circ h_{\omega, \omega'}^s = h_{\omega, \omega''}^s$ and $h_{\omega, \omega}^s = \mathrm{id}$,
- (ii) $A(\omega') \circ h_{\omega, \omega'}^s = h_{T\omega, T\omega'}^s \circ A(\omega)$,

- (iii) $(\omega, \omega') \mapsto h_{\omega, \omega'}^s(\phi)$ is continuous when ω, ω' belong to the same local stable set, uniformly in ϕ .

An *unstable holonomy* $h_{\omega, \omega'}^u : \mathcal{E}_\omega \rightarrow \mathcal{E}_{\omega'}$ is defined analogously for pairs of points in the same unstable set.

By property (i), we have $h_{\omega, \omega'}^\tau = (h_{\omega', \omega}^\tau)^{-1}$ for any $\omega' \in W_{\text{loc}}^\tau(\omega)$, where $\tau \in \{s, u\}$.

These projective holonomies $h_{\omega, \omega'}^s, h_{\omega, \omega'}^u$ typically arise via projectivization of $H_{\omega, \omega'}^s, H_{\omega, \omega'}^u \in \text{SL}(2, \mathbb{R})$ that are obtained as follows:

$$\begin{aligned} H_{\omega, \omega'}^s &= \lim_{n \rightarrow \infty} A_n(\omega')^{-1} A_n(\omega), \\ H_{\omega, \omega'}^u &= \lim_{n \rightarrow \infty} A_{-n}(\omega')^{-1} A_{-n}(\omega) \end{aligned} \quad (2.15)$$

for ω, ω' in the same stable (resp., unstable) set. Conditions need to be placed on the cocycle to ensure convergence in (2.15); see, for example, the proof of Lemma 4.2. The analogues of the properties (i)–(iii) for $H_{\omega, \omega'}^s, H_{\omega, \omega'}^u$ follow directly from the construction and this in turn implies (i)–(iii) for $h_{\omega, \omega'}^s, h_{\omega, \omega'}^u$ by projectivization. Holonomies that arise from (2.15) are called *canonical holonomies* of A .

2.1.6 Invariant measures of projective cocycles

Consider a projective cocycle $(T, A) : \Omega \times \mathbb{RP}^1 \rightarrow \Omega \times \mathbb{RP}^1$ that has stable and unstable holonomies.

Definition 2.9 Suppose we are given a (T, A) -invariant probability measure m on $\Omega \times \mathbb{RP}^1$ that projects to μ in the first component. A *disintegration* of m along the fibers is a measurable family $\{m_\omega : \omega \in \Omega\}$ of conditional probabilities on $\mathbb{RP}^1$ such that $m = \int m_\omega d\mu(\omega)$, that is,

$$m(D) = \int_\Omega m_\omega(\{z \in \mathbb{RP}^1 : (\omega, z) \in D\}) d\mu(\omega) \quad (2.16)$$

for each measurable set $D \subset \Omega \times \mathbb{RP}^1$.

By Rokhlin's disintegration theorem, such a disintegration exists. Moreover, $\{\tilde{m}_\omega : \omega \in \Omega\}$ is another disintegration of m if and only if $m_\omega = \tilde{m}_\omega$ for μ -almost every $\omega \in \Omega$. By a straightforward calculation one checks that $\{A(\omega)_* m_\omega : \omega \in \Omega\}$ is a disintegration of $(T, A)_* m$, where $A(\omega)_* m_\omega$ is the measure on the fiber $\{T\omega\} \times \mathbb{RP}^1$. In particular, the (T, A) -invariance of m implies $A(\omega)_* m_\omega = m_{T\omega}$ for μ -almost every $\omega \in \Omega$. Conversely, if $\{\tilde{m}_\omega : \omega \in \Omega\}$ is a family of probability measures where $\tilde{m}_\omega$ is defined on $\{\omega\} \times \mathbb{RP}^1$,

then we may define a measure $\tilde{m}$ on $\Omega \times \mathbb{RP}^1$ via the right side of (2.16) by replacing m_ω with $\tilde{m}_\omega$. Then $\tilde{m}$ is (T, A) -invariant if $A(\omega)_* \tilde{m}_\omega = \tilde{m}_{T\omega}$ for μ -a.e. ω .

We say m is an s -state (resp., a u -state) if it is in addition invariant under the stable (resp., unstable) holonomies. That is, the disintegration $\{m_\omega : \omega \in \Omega\}$ satisfies that $(h_{\omega, \omega'}^s)_* m_\omega = m_{\omega'}$ for μ -almost every $\omega \in \Omega$ and for every $\omega' \in W_{\text{loc}}^s(\omega)$ (resp., $(h_{\omega, \omega'}^u)_* m_\omega = m_{\omega'}$ for μ -almost every $\omega \in \Omega$ and for every $\omega' \in W_{\text{loc}}^u(\omega)$). In this case, we say that $\{m_\omega\}$ is s -invariant (resp. u -invariant). A measure that is both an s -state and a u -state is called an su -state.

2.1.7 Schrödinger operators and cocycles

In this subsection let us initially assume that Ω is a compact metric space, $T : \Omega \rightarrow \Omega$ is a homeomorphism, and $f : \Omega \rightarrow \mathbb{R}$ is continuous. We consider potentials $V_\omega : \mathbb{Z} \rightarrow \mathbb{R}$ defined by $V_\omega(n) = f(T^n \omega)$ for $\omega \in \Omega$ and $n \in \mathbb{Z}$, and associated Schrödinger operators H_ω in $\ell^2(\mathbb{Z})$ acting by

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

The spectrum $\sigma(H_\omega)$ is defined as

$$\sigma(H_\omega) = \{E \in \mathbb{C} : H_\omega - E \text{ does not have a bounded inverse}\}.$$

For a subset S of a metric space (X, d) and $\delta > 0$, the open δ -neighborhood of S is given by $B_\delta(S) = \{x \in X : d(x, s) < \delta \text{ for some } s \in S\}$. In particular, $B_\delta(x)$ denotes the open ball centered at the point $x \in X$. We need the following uniform estimate that relates the spectrum $\sigma(H_\omega)$ with the orbit $\text{Orb}(\omega) = \{T^n(\omega), n \in \mathbb{Z}\}$; see, for example, [43, Theorem 6].

Proposition 2.10 *For each $\varepsilon > 0$, there exists a $\delta > 0$, depending on ε only, so that the following holds true. If the orbit $\text{Orb}(\omega_0)$ of some $\omega_0 \in \Omega$ satisfies*

$$\text{Orb}(\omega_0) \cap B_\delta(\omega) \neq \emptyset$$

for some $\omega \in \Omega$, then

$$\sigma(H_\omega) \subset B_\varepsilon[\sigma(H_{\omega_0})].$$

Proposition 2.10 implies that if $\text{Orb}(\omega_0)$ is dense in Ω , then $\sigma(H_\omega) \subseteq \sigma(H_{\omega_0})$ for all $\omega \in \Omega$. In this case, we set

$$\Sigma = \sigma(H_{\omega_0}).$$

Let us now return to the main scenario of this paper, where T is a topologically mixing shift operator on a subshift of finite type Ω with an ergodic measure μ satisfying $\text{supp}(\mu) = \Omega$. Let $\text{Per}(T)$ be the set of periodic points of T . Recall that $\text{Per}(T) = \Omega$. Recall we have that $\overline{\text{Orb}(\omega)} = \Omega$ for μ -almost every ω . All these facts together with Proposition 2.10 imply for μ -almost every ω that

$$\Sigma = \sigma(H_\omega) = \overline{\bigcup_{\omega_p \in \text{Per}(T)} \sigma(H_{\omega_p})}. \quad (2.17)$$

Spectral properties of the operators H_ω can be investigated in terms of the behavior of the solutions to the difference equation

$$u(n+1) + u(n-1) + V_\omega(n)u(n) = Eu(n), \quad n \in \mathbb{Z}, \quad (2.18)$$

with E real or complex (depending on the problem in question). These solutions in turn can be described with the help of the Schrödinger cocycle (T, A^E) with the cocycle map $A^E : \Omega \rightarrow \text{SL}(2, \mathbb{R})$ (resp., $\text{SL}(2, \mathbb{C})$) when $E \in \mathbb{C} \setminus \mathbb{R}$ being defined as

$$A^E(\omega) = A^{(E-f)}(\omega) := \begin{pmatrix} E - f(\omega) & -1 \\ 1 & 0 \end{pmatrix},$$

where we often leave the dependence on $f : \Omega \rightarrow \mathbb{R}$ implicit as it will be fixed most of the time. Such cocycles describe the transfer matrices associated with Schrödinger operators with dynamically defined potentials. Specifically, u solves (2.18) if and only if

$$\begin{pmatrix} u(n) \\ u(n-1) \end{pmatrix} = A_n^E(\omega) \begin{pmatrix} u(0) \\ u(-1) \end{pmatrix}, \quad n \in \mathbb{Z}.$$

For the Schrödinger cocycle (T, A^E) , we set $L(E) = L(A^E, \mu)$. One of the main questions in the spectral analysis of the ergodic family of Schrödinger operators $\{H_\omega\}_{\omega \in \Omega}$ (with respect to the ergodic measure μ) is for how many $E \in \Sigma$ we have $L(E) > 0$.

2.2 Periodic potentials

A periodic point ω of T gives rise to a periodic potential, that is, if $T^p \omega = \omega$, then, $V_\omega(n+p) = V_\omega(n)$ for every $n \in \mathbb{Z}$. Since much of our work below will involve the study of periodic points and the associated potentials, let us recall

some basic properties of Schrödinger operators with periodic potentials; see [24, 38] for proofs of the results stated in this subsection.

Consider a Schrödinger operator

$$[H\psi](n) = \psi(n+1) + \psi(n-1) + V(n)\psi(n).$$

in $\ell^2(\mathbb{Z})$ with a p -periodic potential, $V(n+p) = V(n)$ for every $n \in \mathbb{Z}$. Define, for $E \in \mathbb{C}$, the *monodromy matrix*

$$M(E) = \prod_{j=p-1}^0 \begin{pmatrix} E - V(j) & -1 \\ 1 & 0 \end{pmatrix}$$

and the *discriminant* $\Delta(E) = \text{Tr}(M(E))$, where $\text{Tr}(B)$ is the trace of B . The function $\Delta(\cdot)$ is a monic polynomial of degree p .

Proposition 2.11 *The set $\Delta^{-1}((-2, 2))$ consists of p disjoint open intervals and on each of them, Δ is strictly monotone. Moreover, $\sigma(H) = \Delta^{-1}((-2, 2)) = \Delta^{-1}([-2, 2])$.*

This shows that the spectrum of H consists of a finite union of closed intervals and, in fact, the number of connected components of the spectrum is bounded by the period of the potential. This suggests an interesting inverse problem. Suppose we are given a set that has such a form, that is, it has finitely many connected components, each being a closed interval. Suppose further that we know that the set is the spectrum of a periodic Schrödinger operator. Can we say anything about the period of the potential?¹

Proposition 2.12 *Suppose $V : \mathbb{Z} \rightarrow \mathbb{R}$ is periodic. Denote the spectrum of the associated Schrödinger operator by σ .*

(a) *For a probability measure m on σ , consider its potential energy*

$$\mathcal{E}(m) = \iint \log(|E - E'|^{-1}) \, dm(E) \, dm(E') \in \mathbb{R} \cup \{\infty\}. \quad (2.19)$$

Then there is a unique measure, m_σ , which minimizes the potential energy among all probability measures on σ , and in fact $\mathcal{E}(m_\sigma) = 0$.

(b) *The measure m_σ assigns rational weight to each connected component of σ .*

¹ The more natural inverse problem would lead us too far afield. Namely, one should rather ask, given a finite union of closed intervals, identify within a suitable class of operators those that have the given set as their spectrum. The theory is classical but one needs to pass to the more general class of finite-gap Jacobi matrices to study this question in the appropriate setting.

- (c) *The potential V is p -periodic if and only if the weight of each connected component of σ with respect to m_σ is an integer multiple of $\frac{1}{p}$.*

This result shows that the shape of the spectrum of a periodic Schrödinger operator determines the period of the potential. An immediate consequence is the fact that the spectrum of a periodic Schrödinger operator is connected if and only if the period is one, that is, the potential is constant. Another characterization of constant potentials is the following:

Proposition 2.13 *Suppose $V : \mathbb{Z} \rightarrow \mathbb{R}$ is periodic. Then the spectrum σ of the associated Schrödinger operator has Lebesgue measure at most 4. Moreover, the Lebesgue measure of σ is equal to 4 if and only if V is constant.*

Finally, we note the following standard facts. For each $E \in \mathbb{C}$ such that $\Delta(E) \neq \pm 2$, there are exactly two eigendirections $s(E)$ and $u(E)$ in $\mathbb{CP}^1$ of the monodromy matrix $M(E)$, which are actually the so-called Weyl–Titchmarsh m -functions associated with the operator. Moreover, $s(E) \neq u(E)$ are real if and only if $E \in \mathbb{R} \setminus \sigma(H_V)$, and they are the stable and unstable directions of the real hyperbolic matrix $M(E)$. Here we always set $s(E)$ to be the stable direction and $u(E)$ to be the unstable direction. If E is in the upper or lower-half plane or is such that $E \in \mathbb{R}$ and $|\Delta(E)| < 2$, then $s(E)$ and $u(E)$ are not real. In the latter case, we have $s(E) = \overline{u(E)}$. For $\Delta(E) = \pm 2$, we let $I \subseteq \mathbb{R}$ be the connected component of $\sigma(H_V)$ containing E . If E belongs to the boundary of I , then $M(E)$ has a unique real invariant direction. We may think of this case as $s(E) = u(E)$. If E is a point at which a spectral gap is collapsed (or, in other words, at which two different components of $\Delta^{-1}(-2, 2)$ touch), then $M(E) = \pm I_2$, in which case all directions are invariant.

Based on the description above, we may consider two functions s and u which are holomorphic on the upper or lower half plane $\mathbb{H}$ and $\mathbb{C} \setminus \mathbb{H}$, respectively. When restricted to the real line $\mathbb{R}$, they both are continuous functions. Moreover, they are analytic on each spectral gap or in the interior of each connected component of $\sigma(H_V)$. If E_0 is on the boundary of some connected component of $I \subseteq \sigma(H_V)$, then s and u are locally like $g(\sqrt{\pm(E - E_0)})$ near E_0 for some choice of g that is real-analytic near E_0 . Here the choice of g depends on s or u , and the sign of $(E - E_0)$ is determined by whether E_0 is the right or left endpoint of I . Moreover, $s(E)$ and $u(E)$ are real only when $\sqrt{\pm(E - E_0)}$ is real. Thus, we can find an open disk $D \subseteq \mathbb{C}$ centered at E_0 and a ramified (at E_0) double cover $\pi : \tilde{D} \rightarrow D$ of D so that $s(\tilde{E})$ and $u(\tilde{E})$ are holomorphic in $\tilde{E} \in \tilde{D}$. Moreover, when $\pi(\tilde{E}) \in D \cap \mathbb{R}$, $s(\tilde{E})$ and $u(\tilde{E})$ are real only when $\sqrt{\pm(\pi(\tilde{E}) - E_0)}$ is real.

3 Large deviations

The main goal of this section is to prove the following large deviation theorem. Let $C^\alpha(\Omega, \mathbb{R})$, $0 < \alpha \leq 1$, be the space of α -Hölder continuous functions. In other words, $f \in C^\alpha(\Omega, \mathbb{R})$ if there are $C > 0$ such that

$$|f(\omega) - f(\omega')| < C \cdot d(\omega, \omega')^\alpha \text{ for all } \omega, \omega' \in \Omega.$$

Note that here $C^1(\Omega, \mathbb{R})$ is the space of Lipschitz continuous functions, not the space of functions with continuous derivatives. Similarly, we can define the space $C^\alpha(\Omega^+, \mathbb{R})$. Throughout this section, μ , or equivalently μ^+ , will be assumed to have the bounded distortion property.

Theorem 3.1 *Let (Ω, T) be a topologically mixing subshift of finite type. Let μ be a T -ergodic probability measure that has the bounded distortion property. Let $f \in C^\alpha(\Omega, \mathbb{R})$ for some $0 < \alpha \leq 1$. Then, for each $\varepsilon > 0$, there exist $C, c > 0$, depending on f, α , and ε , such that*

$$\mu \left\{ \omega \in \Omega : \left| \frac{1}{n} \sum_{k=0}^{n-1} f(T^k \omega) - \int_{\Omega} f d\mu \right| \geq \varepsilon \right\} < C e^{-cn}, \quad \forall n \geq 1.$$

Theorem 3.1 will be a consequence of the following version of large deviations. Recall we have the spaces $(\Omega^\pm, T_\pm, \mu^\pm)$ of one-sided infinite sequences with nonnegative/nonpositive indices.

Theorem 3.2 *Let (Ω^+, T_+, μ^+) be a topologically mixing one-sided subshift of finite type and suppose that μ^+ is T^+ -ergodic and has the bounded distortion property. Let $f \in C^\alpha(\Omega^+, \mathbb{R})$ for some $0 < \alpha \leq 1$. Then for each $\varepsilon > 0$, there exist $C, c > 0$, depending on f, α , and ε such that*

$$\mu^+ \left\{ \omega^+ \in \Omega^+ : \left| \frac{1}{n} \sum_{k=0}^{n-1} f(T_+^k \omega^+) - \int_{\Omega^+} f d\mu^+ \right| \geq \varepsilon \right\} < C e^{-cn}, \quad \forall n \geq 1. \quad (3.1)$$

We first derive Theorem 3.1 from Theorem 3.2. Let us write $S_n f := \sum_{k=0}^{n-1} f \circ T^k$ for the Birkhoff sums. We let $\varphi(\omega) = \omega^{(\omega_0)} \wedge \omega$ which we defined in (2.5). Note that $\varphi(\omega)$ is continuous on Ω and constant on local stable sets.

Proof of Theorem 3.1 Let $f \in C^\alpha(\Omega, \mathbb{R})$. Since f is Hölder continuous and $\varphi(\omega) \in W_{\text{loc}}^s(\omega)$, a straightforward computation shows that

$$h^s(\omega) := \sum_{n=0}^{\infty} [f(T^n \omega) - f(T^n \varphi(\omega))]$$

converges uniformly, and hence is continuous. We define

$$f^+(\omega) := f(\omega) + h^s(T\omega) - h^s(\omega).$$

In particular, we have

$$\int_{\Omega} f \, d\mu = \int_{\Omega} f^+ \, d\mu \text{ and } \|S_n f - S_n f^+\|_{\infty} < 2\|h^s\|_{\infty},$$

where $\|\cdot\|_{\infty}$ denotes the supremum norm. It is straightforward to see that

$$f^+(\omega) = f(\varphi(\omega)) + \sum_{n=0}^{\infty} [f(T^n T\varphi(\omega)) - f(T^n \varphi(T\omega))],$$

which implies that f^+ is constant on $W_{\text{loc}}^s(\omega)$ for all $\omega \in \Omega$. Moreover, we claim that $f^+ \in C^{\frac{\alpha}{2}}(\Omega, \mathbb{R})$. Indeed, take ω and $\omega' \in \Omega$. Without loss of generality, we may assume $N(\omega, \omega')$ is large and take $k = \lfloor \frac{N}{2} \rfloor$. Then we have

$$\begin{aligned} f^+(\omega) - f^+(\omega') &= \sum_{n=0}^k [f(T^n \varphi(\omega)) - f(T^n \varphi(\omega'))] \\ &\quad + \sum_{n=0}^{k-1} [f(T^n \varphi(T\omega')) - f(T^n \varphi(T\omega))] \\ &\quad + \sum_{n=k}^{\infty} [f(T^n T\varphi(\omega)) - f(T^n \varphi(T\omega))] \\ &\quad - \sum_{n=k}^{\infty} [f(T^n T\varphi(\omega')) - f(T^n \varphi(T\omega'))], \end{aligned}$$

where the absolute values of the first two terms may be bounded by

$$C \sum_{i=1}^k e^{-\alpha(N-i)} \leq C e^{-\alpha \frac{N}{2}} = C d(\omega, \omega')^{\frac{\alpha}{2}},$$

and the absolute values of the last two terms may be bounded by

$$Ce^{-\alpha k} \leq Ce^{-\alpha \frac{N}{2}} = Cd(\omega, \omega')^{\frac{\alpha}{2}}.$$

Thus $f^+ \in C^{\frac{\alpha}{2}}(\Omega, \mathbb{R})$. Since f^+ is constant on local stable sets, it descends to a function in $C^{\frac{\alpha}{2}}(\Omega^+, \mathbb{R})$. Abusing notation slightly, let f^+ denote its descended function as well. Clearly, we have $\int_{\Omega} f^+ d\mu = \int_{\Omega^+} f^+ d\mu^+$ and $S_n f^+(\omega) = S_n f^+(\pi^+ \omega)$. Fix any $\varepsilon > 0$ and define

$$\mathcal{B}_n^+(\varepsilon) := \left\{ \omega^+ \in \Omega^+ : \left| \frac{1}{n} S_n f^+(\omega^+) - \int_{\Omega^+} f^+ d\mu^+ \right| > \varepsilon \right\}.$$

By Theorem 3.2, there are $C, c > 0$, depending on f^+, α , and ε , such that

$$\mu^+(\mathcal{B}_n^+) < Ce^{-cn}, \quad \forall n \geq 0.$$

Combining the relations of f and f^+ above, if we choose $N = N(\varepsilon)$ so that $4\|h^s\|_{\infty} < N\varepsilon$, then we have

$$\left\{ \omega \in \Omega : \left| \frac{1}{n} S_n f(\omega) - \int_{\Omega} f d\mu \right| > \varepsilon \right\} \subseteq (\pi^+)^{-1} \mathcal{B}_n^+(\varepsilon/2), \quad \forall n \geq N.$$

Changing C, c if necessary, we then have for all $n \geq 1$,

$$\begin{aligned} \mu \left\{ \omega \in \Omega : \left| \frac{1}{n} S_n f(\omega) - \int_{\Omega} f d\mu \right| > \varepsilon \right\} &\leq \mu[(\pi^+)^{-1} \mathcal{B}_n^+(\varepsilon/2)] \\ &= \mu^+(\mathcal{B}_n^+(\varepsilon/2)) \\ &< Ce^{-cn}, \end{aligned}$$

as desired. $\square$

To prove Theorem 3.2, we first need the following lemma. For $\underline{l} = (l_1, \dots, l_n)$ where $l_1 \dots l_n$ is admissible (in this case we also just say that $\underline{l}$ is admissible), we set $\Omega_{\underline{l}}^+ := [0; l_1, l_2, \dots, l_n]^+$, $|\underline{l}| := n$, and

$$\mu_{\underline{l}}^+ = \frac{1}{\mu^+(\Omega_{\underline{l}}^+)} T_*^{|\underline{l}|} \mu^+|_{\Omega_{\underline{l}}^+}. \quad (3.2)$$

In other words, $\mu_{\underline{l}}^+$ is the normalized push-forward of μ^+ under the injective map $T^{|\underline{l}|} : \Omega_{\underline{l}}^+ \rightarrow \Omega^+$. Note that $\mu_{\underline{l}}^+$ is concentrated on $T_+^{|\underline{l}|}(\Omega_{\underline{l}}^+)$. By

definition, it clearly holds that

$$\int_{\Omega^+} f d\mu_{\underline{l}} = \frac{1}{\mu(\Omega_{\underline{l}}^+)} \int_{\Omega_{\underline{l}}^+} f \circ T_+^{|\underline{l}|} d\mu^+.$$

If we view $T_+^{-|\underline{l}|}$ as a map from $T_+^{|\underline{l}|}(\Omega_{\underline{l}}^+)$ to $\Omega_{\underline{l}}^+$, we obtain

$$\mu(\Omega_{\underline{l}}^+) \int_{T_+^{|\underline{l}|}(\Omega_{\underline{l}}^+)} f \circ T_+^{-|\underline{l}|} d\mu_{\underline{l}}^+ = \int_{\Omega_{\underline{l}}^+} f d\mu^+.$$

Since $\mu_{\underline{l}}^+$ is concentrated on $T_+^{|\underline{l}|}(\Omega_{\underline{l}}^+)$, we may simply write the equation above as

$$\mu(\Omega_{\underline{l}}^+) \int f \circ T_+^{-|\underline{l}|} d\mu_{\underline{l}}^+ = \int_{\Omega_{\underline{l}}^+} f d\mu^+. \quad (3.3)$$

Recall we also write $[n; \underline{l}]^+ = [n; l_1, \dots, l_n]^+$.

Lemma 3.3 *Consider a topologically mixing one-sided subshift of finite type (Ω^+, T_+, μ^+) , where μ^+ has the bounded distortion property. There exists a $C \geq 1$ so that, uniformly for all admissible $\underline{l}$, we have*

$$\frac{d\mu_{\underline{l}}^+}{d\mu^+}(\omega^+) \leq C \text{ for } \mu\text{-a.e. } \omega^+, \quad (3.4)$$

where $\frac{d\mu_{\underline{l}}^+}{d\mu^+}$ is the Radon–Nikodym derivative of $\mu_{\underline{l}}^+$ with respect to μ^+ . In particular, we have for all positive measurable functions f and all admissible $\underline{l}$,

$$\int f d\mu_{\underline{l}}^+ \leq C \int f d\mu^+. \quad (3.5)$$

Proof Fix an admissible $\underline{l} = (l_1, \dots, l_n)$. Clearly, (3.4) is equivalent to the existence of a $C \geq 1$, independent of $\underline{l}$, such that for every $[n; \underline{i}]^+ = [n; i_1, \dots, i_m]^+ \subseteq \Omega^+$ (which implies $n \geq 0$), we have

$$\frac{\mu_{\underline{l}}^+([n; \underline{i}])}{\mu^+([n; \underline{i}])} \leq C. \quad (3.6)$$

By definition of $\mu_{\underline{l}}^+$, we have

$$\frac{\mu_{\underline{l}}^+([n; \underline{l}])}{\mu^+([n; \underline{l}])} = \frac{\mu^+([0; \underline{l}] \cap [n + |\underline{l}|; \underline{l}])}{\mu^+([0; \underline{l}])\mu^+([n; \underline{l}])}.$$

Hence, (3.6) is equivalent to

$$\frac{\mu^+([0; \underline{l}] \cap [n + |\underline{l}|; \underline{l}])}{\mu^+([0; \underline{l}])\mu^+([n; \underline{l}])} \leq C.$$

By T^+ -invariance of μ^+ , the above estimate is then guaranteed by (2.14). Indeed, if $[0; \underline{l}] \cap [n + |\underline{l}| + r_0; \underline{l}]$, then it is trivial. If $[0; \underline{l}] \cap [n + |\underline{l}| + r_0; \underline{l}] \neq \emptyset$, then it is a consequence of the second inequality of (2.14). $\square$

We are now ready to prove Theorem 3.2. We adopt the strategy of [2, Section 6.1].

Proof of Theorem 3.2 We split the proof into two parts. First, we show

$$\mu^+ \left\{ \omega^+ \in \Omega^+ : \frac{1}{n} S_n f(\omega^+) - \int f d\mu^+ \geq \varepsilon \right\} < C e^{-cn}, \quad \forall n \geq 1. \quad (3.7)$$

For simplicity, we write $I_n(\omega^+) = \frac{1}{n} S_n f(\omega^+)$ and $\gamma = \int f d\mu^+$. Fix a $\varepsilon > 0$. By the Birkhoff Ergodic Theorem, $I_n(\omega)$ converges to $\int f d\mu^+$ pointwise almost everywhere and in L^1 . Thus we have

$$\lim_{n \rightarrow \infty} \mu^+ \{ \omega^+ : |I_n(\omega^+) - \gamma| > \varepsilon \} = 0.$$

By (3.5) of Lemma 3.3, we have

$$\lim_{n \rightarrow \infty} \sup_{\underline{l}} \mu_{\underline{l}}^+ \{ \omega^+ : |I_n(\omega^+) - \gamma| > \varepsilon \} = 0, \quad (3.8)$$

where the supremum is taking over the set of all admissible $\underline{l}$. Fix a $0 < \varepsilon' < \varepsilon$. Let $\mathcal{B}_n = \{ \omega^+ \in \Omega^+ : I_n(\omega^+) > \gamma + \varepsilon' \}$ and $\kappa = \frac{\varepsilon - \varepsilon'}{2}$. By (3.8) (replacing ε by ε'), we have for all admissible $\underline{l}$ and all large n that

$$\begin{aligned} \sup_{\underline{l}} \int (I_n(\omega^+) - \gamma - \varepsilon) d\mu_{\underline{l}}^+ &= \sup_{\underline{l}} \left(\int_{\mathcal{B}_n} + \int_{\mathcal{B}_n^c} \right) (I_n(\omega^+) - \gamma - \varepsilon) d\mu_{\underline{l}}^+ \\ &\leq C \sup_{\underline{l}} \mu_{\underline{l}}^+(\mathcal{B}_n) + (\varepsilon' - \varepsilon) \inf_{\underline{l}} \mu_{\underline{l}}^+(\mathcal{B}_n^c) \\ &< -\kappa. \end{aligned} \quad (3.9)$$

Fix a large N so that (3.9) holds true. For any $\delta' > 0$, by (3.5) and boundedness of the integrand below, there clearly exists a $C' = C'(\delta', N) > 0$ such that for all $|t| < \delta'$ and all $1 \leq n \leq N$, we have

$$\sup_{\underline{l}} \int e^{tn(I_n(\omega^+) - \gamma - \varepsilon)} d\mu_{\underline{l}}^+ \leq C'. \quad (3.10)$$

Hence $\phi_{N,\underline{l}}(t) := \int e^{tN(I_N(\omega^+) - \gamma - \varepsilon)} d\mu_{\underline{l}}^+$ are uniformly bounded holomorphic functions on $\{t \in \mathbb{C} : |t| < \delta'\}$ for all $\underline{l}$ admissible. In particular, $\{\phi_{N,\underline{l}}(t)\}_{\underline{l}}$ is a normal family on the open disk $\{t \in \mathbb{C} : |t| < \delta'\}$. Shrinking δ' if necessary, we see that $\{\phi'_{N,\underline{l}}(t)\}_{\underline{l}}$ is a normal family on the open disk $\{t \in \mathbb{C} : |t| < \delta'\}$ as well. Note that

$$\phi_{N,\underline{l}}(0) = 1 \text{ and } \phi'_{N,\underline{l}}(0) = \int N(I_N(\omega^+) - \gamma - \varepsilon) d\mu_{\underline{l}}^+.$$

By (3.9) and shrinking δ' if necessary, we must have that $\phi'_{N,\underline{l}}(t) < -N\kappa$ for all $t \in (-\delta', \delta')$ and all admissible $\underline{l}$. This implies that

$$(\log \phi_{N,\underline{l}})'(t) = \frac{\phi'_{N,\underline{l}}(t)}{\phi_{N,\underline{l}}(t)} < -(C')^{-1}N\kappa$$

for all $t \in (-\delta', \delta')$ and all admissible $\underline{l}$. Since $(\log \phi_{N,\underline{l}})(0) = 0$, we then have

$$\sup_{\underline{l}} \{ \log \phi_{N,\underline{l}}(t) \} \leq -(C')^{-1}N\kappa t$$

for all $0 \leq \delta < \delta'$. Hence, it holds for all $0 \leq \delta < \delta'$ that

$$\sup_{\underline{l}} \int e^{\delta N(I_N(\omega^+) - \gamma - \varepsilon)} d\mu_{\underline{l}}^+ < e^{-(C')^{-1}N\delta\kappa}. \quad (3.11)$$

Now we want to extend the estimate above to all $n \geq 1$ via the bounded distortion property of μ^+ .

Since $f \in C^\alpha(\Omega^+, \mathbb{R})$ and $nI_n(\omega) = (S_n f)(\omega)$ is the Birkhoff sum, it is straightforward to see that

$$|nI_n(\omega^+) - nI_n(\tilde{\omega}^+)| \leq C \sum_{k=0}^{n-1} d(T_+^k \omega^+, T_+^k \tilde{\omega}^+)^\alpha \leq C_1, \quad (3.12)$$

provided $\omega^+, \tilde{\omega}^+ \in \Omega_{\underline{l}}^+$, where $|\underline{l}| = n$. Note here that C_1 depends only on α and f . We choose $\omega_{\underline{l}, \max}^+, \omega_{\underline{l}, \min}^+ \in \Omega_{\underline{l}}^+$ so that

$$I_n(\omega_{\underline{l}, \max}^+) = \max_{\omega^+ \in \Omega_{\underline{l}}^+} \{I_n(\omega^+)\} \text{ and } I_n(\omega_{\underline{l}, \min}^+) = \min_{\omega^+ \in \Omega_{\underline{l}}^+} \{I_n(\omega^+)\}.$$

In particular we have

$$n(I_n(\omega_{\underline{l}, \max}^+) - I_n(\omega_{\underline{l}, \min}^+)) < C_1. \quad (3.13)$$

Since $(n + N)I_{n+N}(\omega^+) = NI_N(T_+^n \omega^+) + nI_n(\omega^+)$, we have for all $n \geq 1$ that

$$\begin{aligned} & \int e^{\delta(n+N)(I_{n+N}(\omega^+) - \gamma - \varepsilon)} d\mu^+ \\ &= \int e^{\delta n(I_n(\omega^+) - \gamma - \varepsilon)} e^{\delta N(I_N(T_+^n \omega^+) - \gamma - \varepsilon)} d\mu^+ \\ &= \sum_{|\underline{l}|=n} \int_{\Omega_{\underline{l}}^+} e^{\delta n(I_n(\omega^+) - \gamma - \varepsilon)} e^{\delta N(I_N(T_+^n \omega^+) - \gamma - \varepsilon)} d\mu^+ \\ &\leq \sum_{|\underline{l}|=n} e^{\delta n(I_n(\omega_{\underline{l}, \max}^+) - \gamma - \varepsilon)} \int_{\Omega_{\underline{l}}^+} e^{\delta N(I_N(T_+^n \omega^+) - \gamma - \varepsilon)} d\mu^+ \\ &= \sum_{|\underline{l}|=n} \mu^+(\Omega_{\underline{l}}^+) e^{\delta n(I_n(\omega_{\underline{l}, \max}^+) - \gamma - \varepsilon)} \int e^{\delta N(I_N(\omega^+) - \gamma - \varepsilon)} d\mu_{\underline{l}}^+ \\ &\leq \left(\sup_{\underline{l}} \int_{\Omega^+} e^{\delta N(I_N(\omega^+) - \gamma)} d\mu_{\underline{l}}^+ \right) \\ &\quad \cdot \sum_{|\underline{l}|=n} \left(e^{\delta n[I(\omega_{\underline{l}, \max}^+) - I(\omega_{\underline{l}, \min}^+)]} \cdot \int_{\Omega_{\underline{l}}^+} e^{\delta n(I_n(\omega^+) - \gamma - \varepsilon)} d\mu^+ \right) \\ &\leq e^{C_1 \delta} \left(\sup_{\underline{l}} \int_{\Omega^+} e^{\delta N(I_N(\omega^+) - \gamma - \varepsilon)} d\mu_{\underline{l}}^+ \right) \cdot \int_{\Omega^+} e^{\delta n(I_n(\omega^+) - \gamma - \varepsilon)} d\mu^+, \end{aligned}$$

where the third identity follows from (3.3) and the last inequality follows from (3.13). We choose N large so that $(C')^{-1}N\kappa > 2C_1$ and set $c = \frac{1}{2}(C')^{-1}\delta\kappa$. Then by (3.11) we have for all $n \geq 1$,

$$\int_{\Omega^+} e^{\delta(n+N)(I_{n+N}(\omega^+) - \gamma - \varepsilon)} d\mu^+ \leq e^{-cN} \int_{\Omega^+} e^{\delta n(I_n(\omega^+) - \gamma - \varepsilon)} d\mu^+. \quad (3.14)$$

Now, given any $n \geq 1$, we may apply the Euclidean division $n = kN + r$. Using (3.14) several times, we obtain for all $n \geq 1$ that

$$\mu^+ \{ \omega^+ : I_n(\omega^+) - \gamma \geq \varepsilon \} \leq \int_{\Omega^+} e^{\delta n(I_n(\omega^+) - \gamma - \varepsilon)} d\mu^+ \leq C e^{-cn}. \quad (3.15)$$

Note that the estimate for small n is absorbed into the constant C due to (3.10). This gives one part of the estimates (3.7).

To prove the second part which is

$$\mu^+ \{ \omega^+ : \gamma - I_n(\omega) \geq \varepsilon \} \leq \int_{\Omega^+} e^{\delta n(\gamma - I_n(\omega^+) - \varepsilon)} d\mu^+ \leq C e^{-cn}, \quad (3.16)$$

we just need to replace $I_n(\omega) - \gamma - \varepsilon$ by $\gamma - I_n(\omega) - \varepsilon$ and run the same proof of (3.7) above. The only difference is that in this case we set $\mathcal{B}_n = \{ \omega^+ : \gamma - I_n(\omega) > \varepsilon' \}$ for some $0 \leq \varepsilon' < \varepsilon$. All other steps are exactly the same.

The two estimates (3.7) and (3.16) clearly imply the desired large deviation estimate as stated in Theorem 3.2. $\square$

The fact that an equilibrium state of a Hölder continuous potential has local product structure may be found in [8, 14]. Here we show that equilibrium states of Hölder continuous potentials have the bounded distortion property as defined in (2.13), which also implies that such a μ has a local product structure. In particular this shows that Theorem 3.1 holds true for such measures. Equivalently, we may consider (Ω^+, μ^+, T_+) , where μ^+ is an equilibrium state of a Hölder continuous potential and show that such a μ^+ has the bounded distortion property. Indeed, equilibrium states of Hölder continuous potentials defined over (Ω, T) are lifts of equilibrium states of Hölder continuous potentials defined over (Ω^+, T_+) ; see, for example, [8]. According to [8, 14], such a μ^+ has a Hölder continuous Jacobian with respect to T_+ . So it suffices to prove the following lemma:

Lemma 3.4 *Let (Ω^+, T_+, μ^+) be a one-sided subshift of finite type, where μ^+ is a T_+ -ergodic measure that has a Hölder continuous Jacobian. Then μ^+ satisfies the bounded distortion property as defined in (2.14).*

Proof To get (2.14), we fix any $[0; \underline{l}]^+ \subset \Omega^+$ and set $n = |\underline{l}|$. Choose any $[k; \underline{j}]^+ \subset \Omega^+$ such that $k \geq n$ and $[0; \underline{l}]^+ \cap [k; \underline{j}]^+ \neq \emptyset$.

Let $J_+ \in C^\alpha(\Omega, \mathbb{R}_+)$ be the Jacobian of μ^+ with respect to T_+ . Since it is positive and continuous on Ω^+ , we have $\inf_{\omega^+ \in \Omega^+} |J_+(\omega^+)| > c > 0$, which implies that $\log J_+ \in C^\alpha(\Omega^+, \mathbb{R}_+)$. Consider the map $T^n : [0; \underline{l}] \rightarrow \Omega^+$ and

let J_+^l be its Jacobian. Then we have

$$J_+^l(\omega^+) = \prod_{i=0}^{n-1} J_+(T^i \omega^+).$$

Suppose $\omega^+, \tilde{\omega}^+ \in [0; \underline{l}]$ for some $|\underline{l}| = n$. Then we have

$$\begin{aligned} |\log J_+^l(\omega^+) - \log J_+^l(\tilde{\omega}^+)| &= \sum_{i=0}^{n-1} |\log J_+(T^i \omega^+) - \log J_+(T^i \tilde{\omega}^+)| \\ &< C \cdot d(T_+^i \omega^+, T_+^i \tilde{\omega}^+)^{\alpha} \\ &< C, \end{aligned}$$

where C is independent of $\underline{l}$, ω^+ , and $\tilde{\omega}^+$. Thus we have

$$C^{-1} < \left| \frac{J_+^l(\omega^+)}{J_+^l(\tilde{\omega}^+)} \right| < C \text{ for all } \omega^+, \tilde{\omega}^+ \in [0; \underline{l}]^+. \quad (3.17)$$

Now by the definition of the Jacobian, we have

$$\int_{\Omega^+} \chi_{[k-n; \underline{j}]^+}(\eta) d\mu^+(\eta) = \int_{[0; \underline{l}]^+} \chi_{[k-n; \underline{j}]^+}(T_+^n \omega^+) J_+^l(\omega^+) d\mu^+(\omega^+),$$

which implies that

$$J_+^l(\omega_{\underline{l}, \min}^+) \leq \frac{\mu([k; \underline{j}]^+)}{\mu^+([0; \underline{l}]^+ \cap [k; \underline{j}]^+)} \leq J_+^l(\omega_{\underline{l}, \max}^+).$$

Here $\omega_{\underline{l}, \min}^+$ and $\omega_{\underline{l}, \max}^+$ are chosen as in the proof of Theorem 3.2. Using

$$1 = \int_{\Omega^+} 1 d\mu^+ = \int_{[0; \underline{l}]^+} J_+^l(\omega^+) d\mu^+, \text{ we obtain}$$

$$\frac{1}{J_+^l(\omega_{\underline{l}, \max}^+)} \leq \mu^+([0; \underline{l}]^+) \leq \frac{1}{J_+^l(\omega_{\underline{l}, \min}^+)}.$$

Combining the two estimates above with (3.17), we clearly get

$$C^{-1} \leq \frac{\mu^+([0; \underline{l}]^+) \cdot \mu^+([k; \underline{j}]^+)}{\mu^+([0; \underline{l}]^+ \cap [k; \underline{j}]^+)} \leq C,$$

which is (2.14). □

Remark 3.5 There are many early works concerning large deviation estimates for functions defined on hyperbolic dynamical systems; see, for example, [41]. But we could not find a proof that applies in our framework. In Sect. 4, we shall show that if μ has the bounded distortion property, then Theorem 3.1 yields relatively global versions of all the techniques needed. We hope they will be of independent interest. Most importantly, the results, ideas, and proofs given in this section will be used in the second paper [1] of this series.

4 Invariance principle and conformal barycenter

In this section, we develop some tools that are needed to prove our main results in the next two sections. Our main objective is to consider cocycles that have zero Lyapunov exponent. First, we show that a small Lyapunov exponent gives rise to a measurable family of holonomies, which will be integrable if μ has the bounded distortion property. In the case of a zero Lyapunov exponent, we shall introduce techniques originally developed in [33], and then generalized in [3, 7, 39], which are referred to as the *invariance principle*. We will use the invariance principle to show the existence of a continuous su -state on suitable sets. Concretely, in case of a zero Lyapunov exponent and bounded distortion, we will construct an su -state that is continuous on the support of a certain full measure set. In case we have only local product structure, we will construct a local su -state that is continuous on the support of some positive measure set. Then we show that periodic points with small Lyapunov exponents belong to the support in question. Finally, we will construct an su -invariant family of δ -measures by using the conformal barycenter.

4.1 Measurable holonomies resulting from small exponents

For the remainder of this subsection, we fix $0 < \alpha \leq 1$ and consider the space of α -Hölder continuous cocycles $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$, that is, $A \in C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ if

$$\|A(\omega) - A(\tilde{\omega})\| < C \cdot d(\omega, \tilde{\omega})^\alpha \text{ for all } \omega, \omega' \in \Omega. \quad (4.1)$$

Thus for every $\omega, \tilde{\omega} \in \Omega$ and $n \geq 0$, we have

$$\begin{cases} \|A(T^n \omega) - A(T^n \tilde{\omega})\| \leq C e^{-\alpha n} & \text{if } \tilde{\omega} \in W_{\text{loc}}^s(\omega), \\ \|A(T^{-n} \omega) - A(T^{-n} \tilde{\omega})\| \leq C e^{-\alpha n} & \text{if } \tilde{\omega} \in W_{\text{loc}}^u(\omega). \end{cases} \quad (4.2)$$

Our goal is to show that a small Lyapunov exponent produces stable and unstable holonomies. Unfortunately, we can not show that these holonomies

satisfy all properties of Definition 2.8. Therefore we introduce the following weaker version of Definition 2.8.

Definition 4.1 A measurable stable holonomy h^s for A is a measurable family of homeomorphisms $h^s_{\omega, \omega'} : \mathcal{E}_\omega \rightarrow \mathcal{E}_{\omega'}$, defined for μ -almost every ω and every $\omega' \in W^s_{\text{loc}}(\omega)$, satisfying the following properties:

- (i) $h^s_{\omega', \omega''} \circ h^s_{\omega, \omega'} = h^s_{\omega, \omega''}$ and $h^s_{\omega, \omega} = \text{id}$,
- (ii) $A(\omega') \circ h^s_{\omega, \omega'} = h^s_{T\omega, T\omega'} \circ A(\omega)$.

A measurable unstable holonomy $h^u_{\omega, \omega'} : \mathcal{E}_\omega \rightarrow \mathcal{E}_{\omega'}$ is defined analogously for μ -almost every ω and every $\omega' \in W^u_{\text{loc}}(\omega)$.

Using measurable stable and unstable holonomies, we can define the notions of s -state, u -state, and su -state in the same way as in Sect. 2.1.6 since the disintegration $\{m_\omega : \omega \in \Omega\}$ of the invariant measure m is only measurable anyway.

In Lemma 4.2 below we will indeed show that small Lyapunov exponents ensure the existence of measurable stable and unstable holonomies. In fact, these maps will arise via projectivization from canonical holonomies defined on suitably defined subsets.

The following subsets will play a key role. We define for $N \in \mathbb{Z}_+$ and $\delta > 0$,

$$K_s(N, \delta) = \{\omega : \|A_n(\omega)\|^2 \leq e^{(\alpha-\delta)n} \text{ for every } n \geq N\}, \quad (4.3)$$

$$K_u(N, \delta) = \{\omega : \|A_{-n}(\omega)\|^2 \leq e^{(\alpha-\delta)n} \text{ for every } n \geq N\}. \quad (4.4)$$

Lemma 4.2 Assume that $L(A, \mu)$ and $\delta > 0$ are such that $2L(A, \mu) < \alpha - \delta$. Then, for each $N \in \mathbb{Z}_+$, the limit

$$H^s_{\omega, \tilde{\omega}} = \lim_{n \rightarrow \infty} A_n(\tilde{\omega})^{-1} A_n(\omega)$$

exists uniformly for each $\omega \in K_s(N, \delta)$ and $\tilde{\omega} \in W^s_{\text{loc}}(\omega)$. Similarly, for each $N \in \mathbb{Z}_+$, the limit

$$H^u_{\omega, \tilde{\omega}} = \lim_{n \rightarrow \infty} A_{-n}(\tilde{\omega})^{-1} A_{-n}(\omega)$$

exists uniformly for each $\omega \in K_u(N, \delta)$ and $\tilde{\omega} \in W^u_{\text{loc}}(\omega)$. Moreover, if μ has bounded distortion, then the following integrability conditions hold:

$$\int_{\Omega} \log \|H^s_{\omega(\omega_0) \wedge \omega, \omega}\| d\mu(\omega) < \infty, \quad (4.5)$$

$$\int_{\Omega} \log \|H^u_{\omega \wedge \omega(\omega_0), \omega}\| d\mu(\omega) < \infty. \quad (4.6)$$

Since each of the sets $\bigcup_N K_s(N, \delta)$ and $\bigcup_N K_u(N, \delta)$ has full measure and is T -invariant, it follows in particular that there are measurable stable and unstable holonomies in the sense of Definition 4.1.

Proof In this proof, C will denote a positive constant that depends only on $\|A\|_\infty$ and δ . Conditions on it will be placed in several places, which leads to finitely many adjustments.

Let $H_{\omega, \tilde{\omega}}^{s, n} = A_n(\tilde{\omega})^{-1} A_n(\omega)$ for $\tilde{\omega} \in W_{\text{loc}}^s(\omega)$. Define

$$\delta_{\omega, \tilde{\omega}}^{s, n} = \left(H_{\omega, \tilde{\omega}}^{s, n} \right)^{-1} \left(H_{\omega, \tilde{\omega}}^{s, n+1} - H_{\omega, \tilde{\omega}}^{s, n} \right), \quad (4.7)$$

so that

$$H_{\omega, \tilde{\omega}}^{s, n} \left(\text{Id} + \delta_{\omega, \tilde{\omega}}^{s, n} \right) = H_{\omega, \tilde{\omega}}^{s, n+1}.$$

We first will estimate $\|\delta_{\omega, \tilde{\omega}}^{s, n}\|$ as follows. In the case where $\|A_n(\omega)\|^2 \leq e^{(\alpha-\delta)n}$, we have

$$\delta_{\omega, \tilde{\omega}}^{s, n} = A_n(\omega)^{-1} \left(A(T^n \tilde{\omega})^{-1} A(T^n \omega) - \text{Id} \right) A_n(\omega)$$

and therefore

$$\|\delta_{\omega, \tilde{\omega}}^{s, n}\| \leq C e^{(\alpha-\delta)n} e^{-\alpha n} = C e^{-\delta n}. \quad (4.8)$$

The fact $\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_n(\omega)\| = L(A, \mu)$ for μ -almost every $\omega \in \Omega$ implies that $\frac{1}{n} \log \|A_n(\omega)\|$ converges to $L(A, \mu)$ in measure. Thus, the sets $K_s(N, \delta)$ defined in (4.3) are compact and increasing in N , and their union over N has full measure.

For $\omega \in K_s(N, \delta)$, we have the uniform summability statement

$$\sum_{n=N}^{\infty} \|\delta_{\omega, \tilde{\omega}}^{s, n}\| \leq C.$$

Changing C if necessary, the estimate above in turn implies that

$$\begin{aligned} \|H_{\omega, \tilde{\omega}}^{s, n}\| &= \left\| H_{\omega, \tilde{\omega}}^{s, N} \prod_{k=N}^n (H_{\omega, \tilde{\omega}}^{s, k})^{-1} \cdot H_{\omega, \tilde{\omega}}^{s, k+1} \right\| \\ &\leq \|H_{\omega, \tilde{\omega}}^{s, N}\| \prod_{k=N}^n \|(H_{\omega, \tilde{\omega}}^{s, k})^{-1} \cdot H_{\omega, \tilde{\omega}}^{s, k+1}\| \end{aligned}$$

$$\begin{aligned}
&\leq e^{CN} \exp \left(\sum_{k=N}^n \log(1 + \|\delta_{\omega, \tilde{\omega}}^{s,k}\|) \right) \\
&\leq e^{CN} \exp \left(C \sum_{k=N}^n \|\delta_{\omega, \tilde{\omega}}^{s,k}\| \right) \\
&\leq e^{CN},
\end{aligned} \tag{4.9}$$

which implies for all $n \geq N$:

$$\|H_{\omega, \tilde{\omega}}^{s,n} - H_{\omega, \tilde{\omega}}^{s,n+1}\| = \|H_{\omega, \tilde{\omega}}^{s,n} \delta_{\omega, \tilde{\omega}}^{s,n}\| \leq C e^{CN} e^{-\delta n}.$$

Hence $\{H_{\omega, \tilde{\omega}}^{s,n}\}_{n \geq 0}$ is Cauchy sequence in $\text{SL}(2, \mathbb{R})$ and is thus convergent. Let us define

$$H_{\omega, \tilde{\omega}}^s := \lim_{n \rightarrow \infty} H_{\omega, \tilde{\omega}}^{s,n}$$

where the convergence is uniform on $K_s(N, \delta)$. In particular, $H_{\omega, \tilde{\omega}}^s$ depends continuously on $\omega \in K_s(N, \delta)$ and $\tilde{\omega} \in W_{\text{loc}}^s(\omega)$. Changing C if necessary, (4.9) implies for the same ω and $\tilde{\omega}$ that

$$\|A_n(\tilde{\omega})^{-1} A_n(\omega)\| < e^{CN} \tag{4.10}$$

for all $n \geq 1$.

To get the integrability condition (4.5), we define $\phi(\omega) = \log \|A(\omega)\|$ and assume without loss of generality that

$$2 \int_{\Omega} \phi(\omega) d\mu < \alpha - \delta, \tag{4.11}$$

because otherwise we may instead consider $\phi(\omega) = \frac{1}{k} \log \|A_k(\omega)\|$ for some large k , which must satisfy the condition above since $\int_{\Omega} \frac{1}{k} \log \|A_k(\omega)\| d\mu$ converges to $L(A, \mu)$. It is straightforward to see that $\phi \in C^\alpha(\Omega, \mathbb{R})$ since A is α -Hölder continuous and $\|A(\omega)\| \geq 1$ for all $\omega \in \Omega$.

We want to estimate for some $C > 0$ the measure of the following set,

$$\mathcal{B}_N = \{\omega : \log \|H_{\omega, \tilde{\omega}}^s\| > CN \text{ for some } \tilde{\omega} \in W_{\text{loc}}^s(\omega)\}.$$

Recall $S_n \phi(\omega) = \sum_{j=0}^{n-1} \phi(T^j \omega)$. We define for $\delta' = \frac{\delta}{2}$,

$$\mathcal{Z}_m = \left\{ \omega : \frac{1}{n} S_n \phi(\omega) < \frac{\alpha - \delta'}{2} \text{ for all } n \geq m \right\}$$

If $\omega \in \mathcal{Z}_1$, then $2S_n\phi(\omega) < n(\alpha - \delta')$ for all $n \geq 0$, which clearly implies that $\|A_n(\omega)\|^2 < e^{(\alpha - \delta')n}$ for all $n \geq 0$. Thus by the computation leading to (4.10), we have for all $\omega \in \mathcal{Z}_1$,

$$\|A_n(\tilde{\omega})^{-1}A_n(\omega)\| \leq C \text{ for all } \tilde{\omega} \in W_{\text{loc}}^s(\omega) \text{ and for all } n \geq 1.$$

If $\omega \in \mathcal{Z}_m$ with $m > 1$, then we set $0 < k < m$ to be the largest integer for which $\omega \notin \mathcal{Z}_k$. A direct computation then shows that $T^k\omega \in \mathcal{Z}_1$, which implies that for all $\tilde{\omega} \in W_{\text{loc}}^s(T^k\omega)$,

$$\|A_n(\tilde{\omega})^{-1}A_n(T^k\omega)\| \leq C \text{ for all } n \geq 1.$$

Combining $\|A_n(\omega)\| < e^{Cm}$ for all $1 \leq n \leq m$ and for all ω , we have for all $\omega \in \mathcal{Z}_m$ and all $\tilde{\omega} \in W_{\text{loc}}^s(\omega)$ that

$$\|A_n(\tilde{\omega})^{-1}A_n(\omega)\| \leq e^{Cm} \text{ for all } n \geq 1,$$

which clearly implies that $\log \|H_{\omega, \tilde{\omega}}^s\| < Cm$. Thus by choosing C appropriately, we have

$$\mathcal{B}_N \subset \Omega \setminus \mathcal{Z}_N.$$

However, by (4.11) it is clear that

$$\Omega \setminus \mathcal{Z}_N \subseteq \bigcup_{n=N}^{\infty} \left\{ \omega : \left| \frac{1}{n} S_n f(\omega) - \int_{\Omega} \phi d\mu \right| > \frac{\delta}{4} \right\}.$$

Suppose that μ has bounded distortion. Note that $\phi \in C^\alpha(\Omega, \mathbb{R})$. Hence, by Theorem 3.1, there exist $C' > 0$ and $\eta > 0$ such that

$$\mu \left\{ \omega : \left| \frac{1}{n} S_n f(\omega) - \int_{\Omega} \phi d\mu \right| > \frac{\delta}{4} \right\} < C' e^{-\eta n} \text{ for all } n \geq 1.$$

Clearly, this implies that $\mu(\mathcal{B}_N) < C' e^{-\eta N}$ for all $N \geq 1$, which in turn implies the integrability statement (4.5) since

$$\begin{aligned} \int_{\Omega} \log \|H_{\omega^{(\omega_0)} \wedge \omega}^s\| d\mu &= \int_{\Omega} \log \|H_{\omega, \omega^{(\omega_0)} \wedge \omega}^s\| d\mu \\ &= \sum_{N=1}^{\infty} \int_{\mathcal{B}_N \setminus \mathcal{B}_{N+1}} \log \|H_{\omega, \omega^{(\omega_0)} \wedge \omega}^s\| d\mu \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{N=1}^{\infty} \mu(\mathcal{B}_N) C N \\
&\leq \sum_{N=1}^{\infty} C' C e^{-\eta N} N \\
&< \infty.
\end{aligned}$$

The case of $K_u(N, \delta)$ can be done similarly after replacing $A_n(\omega)$ by $A_{-n}(\omega)$. $\square$

Definition 4.3 For a periodic point p with period n , we let $L(A, p) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_n(p)\|$ be the individual Lyapunov exponent of A at p . We say p is γ -bunched if $2L(A, p) < \gamma \leq \alpha$.

Next, we show the following result, which says that α -bunched periodic points are in the support of $K_s(N, \delta)$ for suitable $\delta > 0$ and for large N . We note that a similar result has appeared in [17].²

Lemma 4.4 Suppose (Ω, T) is a subshift of finite type and μ has a local product structure. Assume that $2L(A, \mu) < \alpha$. Let p be an α -bunched periodic point. Then for every $0 < \delta_0 < \min\{\alpha - 2L(A, \mu), \alpha - 2L(A, p)\}$, there exists $N_0 \in \mathbb{Z}_+$ such that

$$p \in \text{supp}(\mu|_{K_s(N_0, \delta_0) \cap K_u(N_0, \delta_0)}). \quad (4.12)$$

Proof Fix any number δ so that $0 < \delta < \min\{\alpha - 2L(A, \mu), \alpha - 2L(A, p)\}$. Then we have $2L(A, p) < \alpha - \delta$ and $2L(A, \mu) < \alpha - \delta$. Assume the period of p is r .

Consider the family $K_s(N, \delta)$ as in Proposition 4.2. By (4.10), we have for all $\omega \in K_s(N, \delta)$, all $\tilde{\omega} \in W_{\text{loc}}^s(\omega)$, and all $n \geq 1$,

$$\|A_n(\tilde{\omega})^{-1} A_n(\omega)\| < e^{CN}.$$

Recall (2.8) says $(\pi^+)^{-1}(\pi^+\omega) = W_{\text{loc}}^s(\omega)$. Thus by definition of $K_s(N, \delta)$ and the estimate above, for each $0 < \delta_1 < \delta$, there exists $N_1 > N$ such that for all $\omega \in (\pi^+)^{-1}(\pi^+[K_s(N, \delta)])$ and all $n \geq N_1$, we have

$$\|A_n(\omega)\|^2 < e^{(\alpha - \delta_1)n}. \quad (4.13)$$

Fix such a choice of δ_1 and N_1 . By choosing N large, we may assume that $\mu(K_s(N, \delta) \cap [0; i]) > 0$ for each $1 \leq i \leq \ell$, which in turn implies that

$$\mu^+[\pi^+(K_s(N, \delta)) \cap [0; i]^+] = \mu(K_s(N, \delta) \cap [0; i]) > 0.$$

² We would like to thank Clark Butler for pointing this out to us.

By Corollary 2.5, for each $n \geq 0$, we have

$$\mu^+(T_+^{-n}[\pi^+(K_s(N, \delta))] \cap [0; p_0, \dots, p_n]^+) > 0. \quad (4.14)$$

By the same argument as the one leading to (4.10), $2L(A, p) < \alpha - \delta$ implies that for each $0 < \delta_2 < \delta$, we can find $m \in \mathbb{Z}_+$ large enough so that $\|A_{rm}(\omega)\|^2 \leq e^{(\alpha-\delta_2)rm}$ for $\omega \in T^{-rm}W_{\text{loc}}^u(T^{rm}p)$. By periodicity of p , we have for each $l \geq 1$ and each $1 \leq k < l$,

$$\|A_{rm}(T^{krm}\omega)\|^2 \leq e^{(\alpha-\delta_2)rm} \text{ for all } \omega \text{ with } T^{-lrm}\omega \in W_{\text{loc}}^u(T^{lrm}p),$$

which in turn implies that for each $1 \leq k \leq l$,

$$\|A_{krm}(\omega)\|^2 < \prod_{j=0}^{k-1} \|A_{rm}(T^{jrm})(\omega)\|^2 \leq e^{(\alpha-\delta_2)krm}. \quad (4.15)$$

For each $l \in \mathbb{Z}_+$, we define the following s -locally saturated set,

$$\mathcal{D}_+^l = (\pi^+)^{-1}(T_+^{-lrm}[\pi^+(K_s(N, \delta))] \cap [0; p_0, \dots, p_{lrm}]^+).$$

By (4.14), we have $\mu(\mathcal{D}_+^l) > 0$ and $\mathcal{D}_+^l \subset [0; p_0, \dots, p_{lrm}]$. For each $0 < \delta_3 < \delta_2$, we can fix a $N' \in \mathbb{Z}_+$ large enough so that the following holds true. For all l large and for each $\omega \in \mathcal{D}_+^l$, we have for all $N' \leq n \leq lrm + N_1$ that

$$\|A_n(\omega)\|^2 < \|A_{krm}(\omega)\|^2 \cdot \|A_{n-krm}(T^{krm}\omega)\|^2 < e^{(\alpha-\delta_3)n},$$

where k is so chosen that $0 \leq n - N_1 - krm < rm$. On the other hand, if $n > lrm + N_1$, then we have

$$\begin{aligned} \|A_n(\omega)\|^2 &\leq \|A_{lrm}(\omega)\|^2 \cdot \|A_{n-lrm}(T^{lrm}\omega)\|^2 \\ &\leq \|A_{lrm}(\omega \wedge p)^{-1} \cdot A_{lrm}(\omega)\|^2 \cdot \|A_{lrm}(\omega \wedge p)\|^2 \\ &\quad \cdot \|A_{n-lrm}(T^{lrm}\omega)\|^2. \end{aligned}$$

We estimate each factor in the product above. First we consider the last factor. The fact that $\omega \in \mathcal{D}_+^l$ implies that $T^{krm}(\omega) \in W_{\text{loc}}^s(\omega')$ for some $\omega' \in K_s(N, \delta)$. Thus by (4.13) and the fact that $n - lrm > N_1$, we have

$$\|A_{n-lrm}(T^{lrm}\omega)\|^2 < e^{(\alpha-\delta_1)(n-lrm)}.$$

For the second factor, the fact $\omega \in \mathcal{D}_+^l$ implies that $T^{-lrm}(\omega \wedge p) \in W_{\text{loc}}^u(T^{lrm}p)$. Thus by (4.15), we have for each $1 \leq k \leq l$

$$\|A_{krm}(\omega \wedge p)\|^2 \leq e^{(\alpha - \delta_2)krm}.$$

In particular, the second factor is taken care of by choosing $k = l$. Combining the fact $\omega \in W_{\text{loc}}^u(\omega \wedge p)$ with the estimate above and using

$$A_{lrm}(\omega) = \prod_{j=l-1}^0 A_{rm}((T^{rm})^j \omega),$$

the same argument getting (4.10) yields

$$\|A_{lrm}(\omega \wedge p)^{-1} \cdot A_{lrm}(\omega)\|^2 < C.$$

Thus by setting $\delta' = \min\{\delta_1, \delta_3\}$, we have for all large l and $n \geq lrm + N_1$ that

$$\|A_n(\omega)\|^2 < e^{(\alpha - \delta')n}.$$

Combining the estimates in the case of $N' \leq n \leq lrm + N_1$, we obtain for all l large, all $\omega \in \mathcal{D}_+^l$, and all $n \geq N'$ that

$$\|A_n(\omega)\|^2 < e^{(\alpha - \delta')n},$$

which implies that $\mathcal{D}_+^l \subset K_s(N', \delta') \cap [0; p_0, \dots, p_{lrm}]$ for all large l . Note that $0 < \delta_i < \delta$, $i = 1, 2, 3$, are arbitrarily chosen, hence δ' can be any number in $(0, \delta)$. In particular, we have for all $\delta_0 \in (0, \delta)$, there is a N' such that $\mathcal{D}_+^l \subset K_s(N', \delta_0) \cap [0; p_0, \dots, p_{lrm}]$ for all large l .

Similarly, for each $0 < \delta_0 < \delta$, we can find a $N'' \in \mathbb{Z}_+$ and a sequence of u -locally saturated $\mathcal{D}_-^l \subset K_u(N'', \delta_0) \cap [-lrm; p_{-lrm}, \dots, p_{-1}, p_0]$ with $\mu(\mathcal{D}_-^l) > 0$. Taking $N_0 = \max\{N', N''\}$, we have for all large l ,

$$\begin{aligned} \mathcal{D}_-^l \cap \mathcal{D}_+^l &\subset K_s(N_0, \delta_0) \cap K_u(N_0, \delta_0), \\ \mathcal{D}_-^l \cap \mathcal{D}_+^l &\subset [-lrm; p_{-lrm}, \dots, p_{lrm}], \end{aligned}$$

where the second line implies that $\mathcal{D}_-^l \cap \mathcal{D}_+^l$ is contained in arbitrarily small neighborhood of p as l tends to infinity. Finally, combining that $\mathcal{D}_+^l$ is s -locally saturated in $[0; p_0]$, $\mathcal{D}_-^l$ is u -locally saturated in $[0; p_0]$, and (2.10), we have for all l large,

$$\mu(\mathcal{D}_-^l \cap \mathcal{D}_+^l) > 0,$$

which then implies that $p \in \text{supp}(\mu|_{K_s(N_0, \delta_0) \cap K_u(N_0, \delta_0)})$. This concludes the proof since the choices of δ_0 and δ such that $0 < \delta_0 < \delta < \min\{\alpha - 2L(A, \mu), \alpha - 2L(A, p)\}$ are all arbitrary. $\square$

4.2 Invariance principle and su -states

Let $\mathcal{M}$ be the Borel σ -algebra of the subshift of finite type (Ω, T, μ) , where μ has a local product structure. Let $A : \Omega \rightarrow \text{SL}(2, \mathbb{R})$ be a measurable map. Then the following *invariance principle* is due to Ledrappier [33], see also [3, 39]:

Proposition 4.5 *Let $\mathcal{B} \subseteq \mathcal{M}$ be a σ -algebra such that*

- (1) $T^{-1}\mathcal{B} \subseteq \mathcal{B} \bmod 0$ and $\{T^n\mathcal{B} : n \in \mathbb{Z}\}$ generates $\mathcal{M} \bmod 0$.
- (2) *the σ -algebra generated by A is contained in $\mathcal{B} \bmod 0$.*

If $L(A, \mu) = 0$, then for any (T, A) -invariant measure m on $\Omega \times \mathbb{R}^1$ that projects to μ in the first component, the disintegration $\{m_\omega\}_{\omega \in \Omega}$ is $\mathcal{B}$ -measurable mod 0.

Definition 4.6 We say that a function defined on Ω *only depends on the future* (resp., *past*) if it is constant on every local stable (resp., unstable) set.

The following consequence of Proposition 4.5 is due to [7]. We sketch a proof for the reader's convenience.

Proposition 4.7 *Suppose A only depends on the future and $L(A, \mu) = 0$. Then for every (T, A) -invariant measure m on $\Omega \times \mathbb{R}^1$ that projects to μ in the first component, its disintegration only depends on the future for μ -almost every $\omega \in \Omega$.*

Proof Let $\mathcal{B} \subseteq \mathcal{M}$ be the σ -algebra generated by sets $\{W_{\text{loc}}^s(\omega) : \omega \in \Omega\}$. It is clear that the sets $W_{\text{loc}}^s(\omega)$ are mutually disjoint. Thus, $D \in \mathcal{B}$ if and only if for each $\omega \in \Omega$, either $W_{\text{loc}}^s(\omega) \cap D = \emptyset$ or $W_{\text{loc}}^s(\omega) \subseteq D$. Since $T\mathcal{B}$ is the σ -algebra generated by $\{TW_{\text{loc}}^s(\omega) : \omega \in \Omega\}$, it is clear that $\mathcal{B} \subseteq T\mathcal{B}$, or equivalently $T^{-1}\mathcal{B} \subseteq \mathcal{B}$. More generally, $T^n\mathcal{B}$ is generated by $\{T^n W_{\text{loc}}^s(\omega) : \omega \in \Omega\}$. Now for any cylinder $[n; l] \subset \Omega$, it is clear that it is $T^n\mathcal{B}$ -measurable for some large $n \in \mathbb{Z}_+$. Since $\mathcal{M}$ is generated by cylinders, we then have that $\{T^n\mathcal{B} : n \in \mathbb{Z}\}$ generates $\mathcal{M} \bmod 0$. The result then follows from Proposition 4.5 and the straightforward fact that A is $\mathcal{B}$ -measurable if and only if A depends on the future. $\square$

An immediate consequence of Proposition 4.7 is that if A is constant along the local stable set and $L(A, \mu) = 0$, then for every (T, A) -invariant measure m on $\Omega \times \mathbb{R}^1$ that projects to μ in the first component, its disintegration is

constant on the local stable set $W_{\text{loc}}^s(\omega)$ for μ -almost every $\omega \in \Omega$. Indeed, we just need to define $\omega' = \phi(\omega)$ to be the sequence for which $\omega'_n = \omega_{-n}$ for all $n \in \mathbb{Z}$ and set

$$\Omega' := \{\omega' = \phi(\omega) : \omega \in \Omega\}.$$

Then μ is again an ergodic measure of (Ω', T) which has a local product structure. Set $A'(\omega') = A(\phi(\omega))$ so that A' depends only on the past. Then it is a standard result that $L(A', \mu) = L(A, \mu) = 0$ and m is (T, A') -invariant if it is (T, A) -invariant. Now the conclusion follows from Proposition 4.7.

We have the following consequence regarding the existence of su -states.

Proposition 4.8 *Suppose the cocycle map A is measurable, satisfies the integrability condition $\int_{\Omega} \log \|A(\omega)\| d\mu < \infty$, and admits measurable canonical stable and unstable holonomies which satisfy the integrability conditions (4.5) and (4.6). If $L(A, \mu) = 0$, then every (T, A) -invariant measure m on $\Omega \times \mathbb{R}P^1$ that projects to μ in the first component has a disintegration that is almost surely invariant under the stable and unstable holonomies.*

Proof First we consider the s -invariance. For simplicity, we define $\varphi(\omega) = \omega^{(\omega_0)} \wedge \omega$, which depends only on the future. We define a new cocycle map as follows:

$$\tilde{A}(\omega) := H_{T\omega, \varphi(T\omega)}^s \cdot A(\omega) \cdot H_{\varphi(\omega), \omega}^s. \quad (4.16)$$

It is clear that $\tilde{A}$ is conjugate to A via the stable holonomy. By the condition (4.5) and the definition of $\tilde{A}$, we then obtain $\int_{\Omega} \log \|\tilde{A}(\omega)\| d\mu < \infty$ and $L(\tilde{A}, \mu) = 0$. On the other hand, by conditions (i)–(ii) of the definition of stable holonomy, we have that

$$\begin{aligned} \tilde{A}(\omega) &= H_{T\omega, \varphi(T\omega)}^s \cdot A(\omega) \cdot H_{\varphi(\omega), \omega}^s \\ &= H_{T\omega, \varphi(T\omega)}^s \cdot H_{T\varphi(\omega), T\omega}^s \cdot A(\varphi(\omega)) \\ &= H_{T\varphi(\omega), \varphi(T\omega)}^s \cdot A(\varphi(\omega)), \end{aligned}$$

which implies that $\tilde{A}(\omega)$ depends only on the future. Thus Proposition 4.7 implies that we have for every $(T, \tilde{A})$ -invariant measure m that projects to μ in the first component, its disintegration only depends on the future.

Now let m be a (T, A) -invariant measure that projects to μ in the first component. Let $\{m_{\omega} : \omega \in \Omega\}$ be a disintegration of m . Thus $A(\omega)_* m_{\omega} = m_{T\omega}$ for μ -almost every ω . We define

$$\tilde{m}_{\omega} = (H_{\omega, \varphi(\omega)}^s)_* m_{\omega}, \quad \omega \in \Omega. \quad (4.17)$$

One readily checks that $\tilde{A}(\omega)_* \tilde{m}_\omega = \tilde{m}_{T(\omega)}$. Thus the family of conditional measures $\{\tilde{m}_\omega : \omega \in \Omega\}$ is a disintegration of a $(T, \tilde{A})$ -invariant measure $\tilde{m}$. Thus $\tilde{m}_\omega$ depends only on the future. In other words, for μ -almost every ω , we have for each $\omega' \in W_{\text{loc}}^s(\omega)$ that

$$(H_{\omega, \varphi(\omega)}^s)_* m_\omega = (H_{\omega', \varphi(\omega')}^s)_* m_{\omega'}. \quad (4.18)$$

Since $\varphi(\omega) = \varphi(\omega')$, by condition (i) of the definition of stable holonomy we have

$$m_\omega = (H_{\varphi(\omega'), \omega}^s \cdot H_{\omega', \varphi(\omega')}^s)_* m_{\omega'} = (H_{\omega', \omega}^s)_* m_{\omega'}. \quad (4.19)$$

In other words, $\{m_\omega : \omega \in \Omega\}$ is s -invariant μ -almost everywhere, concluding the proof of the s -invariance.

As for the u -invariance, we just need to conjugate A to a new $\tilde{A}$ via the unstable holonomy so that $\tilde{A}$ is constant along the local unstable set $W_{\text{loc}}^u(\omega)$ for μ -almost every $\omega \in \Omega$. Then by repeating the same argument above and using the remark following Proposition 4.7, we obtain that m_ω is u -invariant μ -almost everywhere. This completes the proof. $\square$

Lemma 4.9 *Assume that $L(A, \mu) = 0$ and μ has the bounded distortion property. Then there exists a full measure set $K \subset \Omega$ on which one has measurable stable and unstable holonomies. Moreover, every (T, A) -invariant measure m on $\Omega \times \mathbb{RP}^1$ that projects to μ in the first component has a continuous, su -invariant disintegration over $\text{supp}(\mu|_K) \cap K$.*

Proof Since $L(A, \mu) = 0$, Lemma 4.2 applies and yields for each δ with $0 < \delta < \alpha$, the sets $K_s(N, \delta)$, $K_u(N, \delta)$ along with continuous families of holonomies satisfying the estimates required to apply Proposition 4.8. Thus, applying Proposition 4.8, choose a (T, A) -invariant measure m on $\Omega \times \mathbb{RP}^1$ and consider its disintegration $\{m_\omega\}$, which is invariant almost everywhere with respect to the stable and unstable holonomies. Recall that both $\bigcup_{N>0} K_s(N, \delta)$ and $\bigcup_{N>0} K_u(N, \delta)$ have full measure. We let

$$K_\delta = \left(\bigcup_{N>0} K_s(N, \delta) \right) \cap \left(\bigcup_{N>0} K_u(N, \delta) \right). \quad (4.20)$$

As in [7], we can now produce a disintegration $\{\tilde{m}_\omega\}$ over $\text{supp}(\mu|_{K_\delta}) \cap K_\delta$, which is holonomy-invariant and continuous. For the reader's convenience we give the argument. In the following argument, we will work in the full measure set K_δ so that the stable and unstable holonomies are defined on the local stable and unstable set of $\omega \in K_\delta$, respectively.

For each $\omega \in K_\delta$, if $m_{\omega'}$ already exists for some $\omega' \in W_{\text{loc}}^s(\omega)$ from the original disintegration of m above, then we may define $m_{\omega''}^s$ via $H^s(\omega', \omega'')_* m_{\omega'}$ for each $\omega'' \in W_{\text{loc}}^s(\omega)$. If $m_{\omega'}$ does not exist from the original disintegration of m for all $\omega' \in W_{\text{loc}}^s(\omega)$, then $W_{\text{loc}}^s(\omega)$ is a μ -zero measure set and we may pick any probability measure m_ω^s and extend via $m_{\omega''}^s = H^s(\omega, \omega'')_* m_\omega^s$ for each $\omega'' \in W_{\text{loc}}^s(\omega)$. Clearly, the new family m_ω^s is invariant under the stable holonomy at every $\omega \in K_\delta$. On the other hand, due to the almost sure invariance of the original disintegration m_ω of m , the new family m_ω^s coincides with m_ω for μ -almost every ω in K_δ , and hence it also coincides with m_ω for μ -almost every ω . In particular, m_ω^s is again a disintegration of m . Similarly, we may construct another disintegration m_ω^u of m which is invariant under the unstable holonomy at every $\omega \in K_\delta$. Note that the set $\tilde{K} = \{\omega \in K_\delta : m_\omega^s = m_\omega^u\}$ has full μ -measure.

Clearly, for each $1 \leq j \leq \ell$, $[0; j] \cap \tilde{K}$ has full μ -measure in $[0; j]$. By the local product structure of μ , $[0; j] \cap \tilde{K}$ has full $\mu^- \times \mu^+$ -measure in $[0; j]$. Thus by Fubini's theorem, for μ^- -almost every $\omega^- \in [0; j]^-$, we have that $\pi^+[(\{\omega^-\} \times [0; j]^+) \cap \tilde{K}]$ has full μ^+ -measure in $[0; j]^+$. Note for each $\omega \in [0; j]$ with $\pi^-(\omega) = \omega^-$, we have $\omega^- \times [0; j]^+ = W_{\text{loc}}^u(\omega)$. Thus for each $1 \leq j \leq \ell$, we may choose an $\omega^{(j)} \in [0; j] \cap K_\delta$ such that

$$\mu^+ \left(\pi^+ \left(W_{\text{loc}}^u(\omega^{(j)}) \cap \tilde{K} \right) \right) = \mu^+([0; j]^+).$$

By the definition of μ^+ , we then have that

$$\begin{aligned} & \mu \left((\pi^+)^{-1} \left[\pi^+ \left(\bigcup_{j=1}^{\ell} (W_{\text{loc}}^u(\omega^{(j)}) \cap \tilde{K}) \right) \right] \right) \\ &= \mu^+ \left(\pi^+ \left(\bigcup_{j=1}^{\ell} (W_{\text{loc}}^u(\omega^{(j)}) \cap \tilde{K}) \right) \right) \\ &= \sum_{j=1}^{\ell} \mu^+([0; j]^+) = 1. \end{aligned} \tag{4.21}$$

In other words, for μ -almost every $\omega \in K_\delta$, we have $\omega^{(\omega_0)} \wedge \omega \in W_{\text{loc}}^u(\omega^{(\omega_0)}) \cap \tilde{K}$. Now for each $\omega \in K_\delta$, we define

$$\tilde{m}_\omega^s = H_{\omega^{(\omega_0)} \wedge \omega, \omega}^s \cdot m_{\omega^{(\omega_0)} \wedge \omega}^u = H_{\omega^{(\omega_0)} \wedge \omega, \omega}^s \cdot H_{\omega^{(\omega_0)}, \omega^{(\omega_0)} \wedge \omega}^u \cdot m_{\omega^{(\omega_0)}}^u.$$

Recall that by the proof of Lemma 4.7, the stable and unstable holonomies are continuous on each local stable and unstable set, respectively. Thus the

equalities in the definition of $\tilde{m}_\omega$ above imply that for each $1 \leq j \leq \ell$, we have that $\tilde{m}_\omega^s$ is continuous in ω at $[0; j] \cap K_\delta$. Thus $\tilde{m}_\omega^s$ is continuous on K_δ . Clearly, the construction also implies that $\tilde{m}_\omega^s$ is s -invariant. On the other hand, by invariance with respect to stable holonomies, we have that for each ω such that $\omega^{(\omega_0)} \wedge \omega \in W_{\text{loc}}^u(\omega^{(\omega_0)}) \cap \tilde{K}_{\omega_0}$, we have

$$\tilde{m}_\omega^s = H_{\omega^{(\omega_0)} \wedge \omega, \omega}^s \cdot m_{\omega^{(\omega_0)} \wedge \omega}^u = H_{\omega^{(\omega_0)} \wedge \omega, \omega}^s \cdot m_{\omega^{(\omega_0)} \wedge \omega}^s = m_\omega^s.$$

Thus, we have $\tilde{m}_\omega^s = m_\omega^s$ for μ -almost every $\omega \in K_\delta$, and we obtain an s -invariant and continuous disintegration $\{\tilde{m}_\omega^s\}$ of m . Producing in an analogous fashion a u -invariant and continuous disintegration $\{\tilde{m}_\omega^u\}$, we find that $\tilde{m}_\omega^s = \tilde{m}_\omega^u$ in $\text{supp}(\mu|_{K_\delta}) \cap K_\delta$ by continuity and almost everywhere coincidence. This produces an su -invariant continuous disintegration $\{\tilde{m}_\omega\}$ over $\text{supp}(\mu|_{K_\delta}) \cap K_\delta$ by setting $\tilde{m}_\omega = \tilde{m}_\omega^s$. By continuity, we also have invariance under (T, A) , that is, $A(\omega)_* \tilde{m}_\omega = \tilde{m}_{T\omega}$ for every $\omega \in \text{supp}(\mu|_{K_\delta}) \cap K_\delta$. Clearly, any K_δ can be chosen to be the desired K . $\square$

4.3 Application of conformal barycenter

Let $\mathbb{H} \subseteq \mathbb{C}$ be the upper-half plane, $\mathbb{D}$ the open unit disk, and $S^1 = \partial\mathbb{D}$ the unit circle. It is a standard result that the Möbius transformation associated with an element of the group $\text{SU}(1, 1)$ preserves S^1 and $\mathbb{D}$. Here $P = \begin{pmatrix} a & b \\ b & \bar{a} \end{pmatrix} \in \text{SU}(1, 1)$ if $|a|^2 - |b|^2 = 1$ and the Möbius transformation associated with it is $P \cdot z = \frac{az+b}{bz+\bar{a}}$. It is a standard result that $\text{SU}(1, 1)$ is conjugate to $\text{SL}(2, \mathbb{R})$ through the $\text{SL}(2, \mathbb{C})$ -matrix $Q = \frac{-1}{1+i} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$, that is, $Q^* \text{SU}(1, 1) Q = \text{SL}(2, \mathbb{R})$. In fact, we have the following commutative diagram:

$$\begin{array}{ccc} \mathbb{H} & \xrightarrow{(Q^{-1} P Q) \cdot} & \mathbb{H} \\ Q \cdot \downarrow & & \downarrow Q \cdot \\ \mathbb{D} & \xrightarrow{P \cdot} & \mathbb{D}, \end{array} \quad (4.22)$$

where all transformations are Möbius transformations, as well as homeomorphisms. Moreover, Q is a homeomorphism between their boundaries, that is, a homeomorphism from $\mathbb{RP}^1 = \mathbb{R} \cup \{\infty\} = \partial\mathbb{H}$ to $S^1 = \partial\mathbb{D}$. We need the following proposition from [25, Proposition 1].

Proposition 4.10 *For each probability measure ν on the unit circle S^1 containing no atom of mass $\geq \frac{1}{2}$, there is an unique point $B(\nu) \in \mathbb{D}$, called the conformal barycenter of ν , so that the map $\nu \rightarrow B(\nu)$ is invariant under*

the Möbius transformation of $SU(1, 1)$, that is, $B(P_*v) = P \cdot B(v)$ for each $P \in SU(1, 1)$.

Lemma 4.11 *Let (Ω, T, μ) , A , and $K \subset \Omega$ be as in Lemma 4.9. Then there exists a family of A -invariant, su -invariant measures $\hat{m}_\omega$ over $\text{supp}(\mu|_K) \cap K$ such that for each $\omega \in \text{supp}(\mu|_K) \cap K$, $\hat{m}_\omega$ is supported by at most two points of $\mathbb{CP}^1$.*

Proof We start with the continuous disintegration $\{\tilde{m}_\omega : \omega \in \Omega\}$ of m over $\omega \in \text{supp}(\mu|_K) \cap K$ that we constructed as stated in Lemma 4.9. To produce the family of measures $\{\hat{m}_\omega\}$, we divide it into three different cases.

If $\tilde{m}_\omega$ has an atom $z(\omega) \in \mathbb{RP}^1$ of mass $> 1/2$, we let $\hat{m}_\omega = \delta_{z(\omega)}$, that is, the Dirac measure (mass one) supported in this point $z(\omega)$. By invariance of $\tilde{m}_\omega$ under the holonomies, it is clear that if $\tilde{m}_\omega$ has such a point $z(\omega)$, then so does $m_{\omega'}$ for each point $\omega' \in W_{\text{loc}}^s(\omega) \cup W_{\text{loc}}^u(\omega)$. Moreover, $z(\omega') = H_{\omega, \omega'}^*(z(\omega))$ for $* \in \{s, u\}$, which exactly implies that $\delta_{z(\omega)}$ is invariant under the holonomies. Similarly, by invariance of $\tilde{m}_\omega$ under $A(\omega)$, we have that $\tilde{m}_{T^n\omega}$ has such a point mass for all $n \in \mathbb{Z}$ and $A(\omega)_*\delta_{z(\omega)} = \delta_{z(T\omega)}$.

If $\tilde{m}_\omega$ contains two atoms of mass $1/2$ each, we set $\hat{m}_\omega = \tilde{m}_\omega$. Similar to the argument of the case (1) above, we have that $m_{\omega'}$ falls into this case for each point $\omega' \in W_{\text{loc}}^s(\omega) \cup W_{\text{loc}}^u(\omega) \cup \text{Orb}(\omega)$ and $\hat{m}_\omega$ is invariant under the holonomies and $A(\omega)$.

In all other cases, by Proposition 4.10, we define $\hat{m}_\omega$ to be the Dirac measure supported at

$$z(\omega) := Q^{-1} \cdot B(Q_*\tilde{m}_\omega) \in \mathbb{H}, \quad (4.23)$$

where $B(Q_*\tilde{m}_\omega)$ is the conformal barycenter of the measure $Q_*\tilde{m}_\omega$ of the unit circle S^1 . Note again by holonomy invariance, if ω is not in the two cases above, then neither is $m_{\omega'}$ for each point $\omega' \in W_{\text{loc}}^s(\omega) \cup W_{\text{loc}}^u(\omega) \cup \text{Orb}(\omega)$. Moreover, for $\omega' \in W_{\text{loc}}^s(\omega)$, we have $QH_{\omega, \omega'}^s Q^{-1} \in SU(1, 1)$ which together with Proposition 4.10 and the holonomy invariance of $\tilde{m}_\omega$ implies that

$$\begin{aligned} H_{\omega, \omega'}^s \cdot z(\omega) &= H_{\omega, \omega'}^s \cdot (Q^{-1} \cdot B(Q_*\tilde{m}_\omega)) \\ &= Q^{-1} (QH_{\omega, \omega'}^s Q^{-1}) \cdot B(Q_*\tilde{m}_\omega) \\ &= Q^{-1} \cdot B\left((QH_{\omega, \omega'}^s Q^{-1})_* Q_*\tilde{m}_\omega\right) \\ &= Q^{-1} \cdot B\left((QH_{\omega, \omega'}^s Q^{-1} Q)_* \tilde{m}_\omega\right) \\ &= Q^{-1} \cdot B\left(Q_*(H_{\omega, \omega'}^s)_* \tilde{m}_\omega\right) \\ &= Q^{-1} \cdot B(Q_*\tilde{m}_{\omega'}) \\ &= z(\omega'), \end{aligned}$$

which in turn implies that $\hat{m}(\omega)$ is invariant under the stable holonomy. By a similar argument we can establish the invariance under the unstable holonomy and under $A(\omega)$. $\square$

4.4 Local su -invariance

In this subsection, we drop the assumption that μ has bounded distortion and assume only that it has a local product structure. Note that to ensure the existence of the su -state when only assuming that the Lyapunov exponent is small, we need the bounded distortion property of μ so that we have the integrability conditions on the measurable stable and unstable holonomies. But to prove our main results concerning the positivity of the Lyapunov exponent, we actually only need a “local su -state” for which a local product structure suffices.

We adapt the techniques from [39] to produce a certain disintegration of m that has local su -invariance. Throughout this subsection, we assume $A \in C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ to be such that $L(A, \mu) = 0$ and we fix a δ with $\frac{\alpha}{2} < \delta < \alpha$.

We start with the following consequence of the proof of Lemma 4.2. Recall that $K_s(N, \delta)$ was defined in (4.3).

Lemma 4.12 *Let (Ω, T, μ) , A , and δ be above. Then there exists a $\tilde{C} = \tilde{C}(\delta, N)$ so that the following holds true. For all $\omega \in K_s(N, \delta)$, all $\tilde{\omega} \in W_{\text{loc}}^s(\omega)$, and all $j \geq 0$, we have*

$$H_{T^j\omega, T^j\tilde{\omega}}^s := \lim_{n \rightarrow \infty} H_{T^j\tilde{\omega}, T^j\omega}^{s,n} \text{ exists and } \|H_{T^j\tilde{\omega}, T^j\omega}^s\| \leq \tilde{C}. \quad (4.24)$$

Proof By (4.9), we have that

$$\|H_{\tilde{\omega}, \omega}^s\| = \left\| \lim_{n \rightarrow \infty} H_{\tilde{\omega}, \omega}^{s,n} \right\| \leq e^{CN}.$$

A direct computation shows that

$$H_{T^j\tilde{\omega}, T^j\omega}^{s,n} = A_j(\tilde{\omega}) H_{\tilde{\omega}, \omega}^{s, n+j} A_j(\omega)^{-1},$$

which implies the existence of

$$\begin{aligned} H_{T^j\tilde{\omega}, T^j\omega}^s &:= \lim_{n \rightarrow \infty} H_{T^j\tilde{\omega}, T^j\omega}^{s,n} \text{ and} \\ \|H_{T^j\tilde{\omega}, T^j\omega}^s\| &\leq \|H_{T^j\tilde{\omega}, T^j\omega}^{s,n}\| \cdot \|A_j(\omega)\| \cdot \|A_j(\tilde{\omega})\|. \end{aligned}$$

In particular, for all $0 \leq j \leq N$, we have

$$\|H_{T^j\tilde{\omega}, T^j\omega}^s\| \leq e^{3CN}. \quad (4.25)$$

Fix a $j > N$. Then we have $\|A_j(\omega)\|^2 < e^{(\alpha-\delta)j}$ since $\omega \in K_s(N, \delta)$. Using (4.8), a direct computation shows that

$$\begin{aligned}\|\delta_{T^j\tilde{\omega}, T^j\omega}^{s,n}\| &= \|A_j(\omega)\delta_{\tilde{\omega},\omega}^{s,n+j}A_j(\omega)^{-1}\| \\ &\leq Ce^{(\alpha-\delta)j}e^{(\alpha-\delta)(n+j)}e^{-\alpha(n+j)} \\ &= Ce^{(\alpha-2\delta)j}e^{-\delta n} \\ &< Ce^{-\delta n},\end{aligned}$$

where the last inequality follows from the fact $2\delta > \alpha$. Combining (4.24) and the proof of (4.9), we obtain for all $j \geq N$

$$\|H_{T^j\tilde{\omega}, T^j\omega}^s\| \leq \|H_{T^j\tilde{\omega}, T^j\omega}^{s,N}\| \cdot \exp\left(C \sum_{n=N}^{\infty} \|\delta_{\omega, \tilde{\omega}}^{s,n}\|\right) \leq Ce^{2CN}. \quad (4.26)$$

Combining (4.25) and (4.26), we clearly obtain the latter half of (4.24), where we may take $\tilde{C} = \max\{e^{3CN}, Ce^{2CN}\}$. $\square$

So we may choose N large so that $K(N, \delta) = K_s(N, \delta) \cap K_u(N, \delta)$ (which were defined in (4.3)) has measure sufficiently close to 1 and $\mu(K(N, \delta) \cap [0; j]) > 0$ for all $1 \leq j \leq \ell$. We set $K_\tau^j := K_\tau(N, \delta) \cap [0; j]$ for $\tau \in \{s, u\}$ and $K^j = K_s^j \cap K_u^j$.

Lemma 4.13 *Let (Ω, T, μ) , A , δ be as in Lemma 4.12 and let $K(N, \delta)$ be as above. Then for every (T, A) -invariant measure m that projects to μ on the first component, there is a disintegration $\{m_\omega : \omega \in \Omega\}$ of m that is su -invariant for μ -almost every $\omega \in K(N, \delta)$.*

Proof We only consider the case of s -invariance, as u -invariance can be established in a completely analogous way. We break the argument into three steps. *Step I.* As in the proof of Proposition 4.7, the first step is to construct a certain σ -algebra $\mathcal{B}$ to which we can apply the invariance principle as formulated in Proposition 4.5. By (4.9), $K_s(N, \delta)$ is s -saturated. Similarly, $K_u(N, \delta)$ is u -saturated. Fix a $\omega^j \in K^j$ and set $S = W_{\text{loc}}^s(\omega^j) \cap K^j$. For each $\omega' \in S$, we define $r(\omega') = 1$ if $T(W_{\text{loc}}^s(\omega')) \cap W_{\text{loc}}^s(\omega'') = \emptyset$ for some $\omega'' \in S$; otherwise, we define $2 \leq r(\omega') \in \mathbb{Z}_+ \cup \{\infty\}$ be the largest number such that $T^i(W_{\text{loc}}^s(\omega')) \cap W_{\text{loc}}^s(\omega'') = \emptyset$ for all $\omega'' \in S$ and for all $0 < i < r(\omega')$. Now we define the σ -algebra $\mathcal{B} \subseteq \mathcal{M}$ to be the one generated by the family

$$\{T^i(W_{\text{loc}}^s(\omega')) : \omega' \in S, 0 \leq i < r(\omega')\}.$$

By our definition of $r(\omega')$, it is clear that the sets in the family above are mutually disjoint. Thus, $\mathcal{B}$ contains all $B \in \mathcal{M}$ such that for all $\omega' \in S$ and all

$0 \leq i < r(\omega')$, either $B \cap T^i(W_{\text{loc}}^s(\omega')) = \emptyset$ or $T^j(W_{\text{loc}}^s(\omega')) \subset B$. First we claim that $\mathcal{B}$ satisfies condition (1) of Proposition 4.5. The proof is analogous to the one of Proposition 4.7. Indeed, $T\mathcal{B}$ is the σ -algebra generated by

$$\{T^{i+1}(W_{\text{loc}}^s(\omega')) : \omega' \in S, 0 \leq i < r(\omega')\},$$

which is again a family of mutually disjoint sets. Since $T^{r(\omega')}(W_{\text{loc}}^s(\omega')) \subseteq W_{\text{loc}}^s(\omega'')$ for some $\omega'' \in S$, one readily checks that $B \in \mathcal{B}$ implies $B \in T\mathcal{B}$. Hence, we have that $T\mathcal{B}$ contains $\mathcal{B}$, or equivalently, $T^{-1}\mathcal{B} \subseteq \mathcal{B}$. More generally, for all $n \geq 1$, we have that $T^n\mathcal{B}$ is generated by $\{T^{i+n}(W_{\text{loc}}^s(\omega')) : \omega' \in S, 0 \leq i < r(\omega')\}$, which implies that $T^n\mathcal{B}$, $n \geq 1$ generates $\mathcal{M}$ mod 0. Indeed, since $\mathcal{M}$ is generated by cylinders, we just need to show that all $[k; \underline{l}]$ are contained in $T^n\mathcal{B}$ for some large n . Taking any $n \geq |k|$, it is clear that $[k; \underline{l}] \in T^n\mathcal{B}$.

Step II. Similarly to the proof of Proposition 4.8, our second step is to conjugate A to some $\tilde{A}$, which is measurable with respect to $\mathcal{B}$. We define $\tilde{A}$ by

$$\tilde{A}(\omega) := H_{T\omega, T^{i+1}\omega'}^s A(\omega) H_{T^i\omega', \omega}^s = A(T^i\omega') \quad (4.27)$$

if $\omega \in T^i(W_{\text{loc}}^s(\omega'))$ for some $\omega' \in S$ (so that $\omega \in W_{\text{loc}}^s(T^i\omega')$) and $0 \leq i < r(\omega')$; and

$$\tilde{A}(\omega) := A(\omega) \text{ otherwise.} \quad (4.28)$$

Clearly, if we set $B(\omega)$ as

$$B(\omega) = \begin{cases} H_{\omega, T^j\omega'}^s, & \omega \in T^j(W_{\text{loc}}^s(\omega')), \omega' \in S, \text{ and } 0 \leq j < r(\omega') \\ I_2, & \text{otherwise,} \end{cases} \quad (4.29)$$

then $\tilde{A}(\omega) = B(T\omega)A(\omega)B(\omega)^{-1}$. In other words, $\tilde{A}$ is conjugate to A via B . Combining this with the fact that $\omega' \in S \subseteq K_s(N, \delta)$ and Lemma 4.12, we have $\|B(\omega)\| \leq \tilde{C}$ for all $\omega \in \Omega$. In particular, $\int \log \|\tilde{A}\| d\mu < \infty$ and $L(\tilde{A}, \mu) = 0$. By definition, $\tilde{A}$ is constant on $T^i(W_{\text{loc}}^s(\omega'))$ for any $\omega' \in S$ and any $0 \leq i < r(\omega')$, which clearly implies that $\tilde{A}$ is $\mathcal{B}$ -measurable.

Step III. Following the second half of the proof of Proposition 4.8, for any given disintegration $\{m_\omega : \omega \in \Omega\}$ of m , we can set

$$\tilde{m}_\omega = B(\omega)_*^{-1} m_\omega, \quad \omega \in \Omega. \quad (4.30)$$

Then it becomes a disintegration of a $(T, \tilde{A})$ -invariant measure $\tilde{m}$. We can now apply Proposition 4.5 to $(\mathcal{B}, \tilde{A}, \tilde{m}_\omega)$ and obtain that $\{\tilde{m}_\omega\}$ is $\mathcal{B}$ -measurable. In

particular, $\tilde{m}_\omega$ is constant on $T^i(W_{\text{loc}}^s(\omega'))$ for all $\omega' \in S$ and all $0 \leq i \leq r(\omega')$. Taking $i = 0$, then a similar proof to (4.19) yields

$$(H_{\omega, \tilde{\omega}}^s)_* m_\omega = m_{\tilde{\omega}} \text{ for all } \omega, \tilde{\omega} \in W_{\text{loc}}^s(\omega') \text{ and all } \omega' \in S. \quad (4.31)$$

Thus we have obtained s -invariance for all points in

$$\bigcup_{\omega \in S} \{W_{\text{loc}}^s(\omega) : \omega \in W_{\text{loc}}^u(\omega^j)\},$$

which contains

$$\bigcup_{\omega \in W_{\text{loc}}^u(\omega^j) \cap K^j} \{W_{\text{loc}}^s(\omega) : \omega \in W_{\text{loc}}^u(\omega^j) \cap K^j\}.$$

By local product structure of μ and following the proof of (4.21), we have that the set above is a full measure subset of K^j . Since $1 \leq j \leq \ell$ is arbitrarily chosen, we thus obtain s -invariance of $\{m_\omega\}$ on a full measure subset of $K(N, \delta)$. $\square$

Now we can apply the proof of Lemmas 4.9 and 4.11 (replacing K_δ by $K(N, \delta)$) to obtain the following corollary.

Corollary 4.14 *Using the setup of Lemma 4.13, there is a disintegration $\{\tilde{m}_\omega\}$ of m so that $\omega \mapsto \tilde{m}_\omega$ is continuous and su -invariant on $\text{supp}(K(N, \delta)) \cap K(N, \delta)$. Moreover, there is family of measures $\{\hat{m}_\omega\}$ that is su -invariant on $\text{supp}(K(N, \delta)) \cap K(N, \delta)$ and for each ω , $\text{supp}(\hat{m}_\omega)$ contains at most two points.*

The main goal of the present Sect. 4 is to obtain the following corollary.

Corollary 4.15 *Suppose (Ω, T) is a subshift of finite type and μ is a T -ergodic measure that has a local product structure. Let $A : \Omega \rightarrow \text{SL}(2, \mathbb{R})$ be a cocycle map so that $L(A, \mu) = 0$. Then for every periodic point p (of period n) such that $2L(A, p) < \frac{\alpha}{2}$, there exists a set $Z_p \subseteq \mathbb{CP}^1$, invariant under complex conjugation and under $A_n(p)$, and consisting of either one or two points, with the following property. Let q be another periodic point such that $2L(A, q) < \frac{\alpha}{2}$. If $p_0 = q_0$, then*

$$H_{q, q \wedge p}^u(Z_q) = H_{p, q \wedge p}^s(Z_p). \quad (4.32)$$

Proof Since $L(A, \mu) = 0$ and $L(A, p) < \frac{\alpha}{2}$, by Lemma 4.4, we clearly have that $p \in \text{supp}(\mu|_{K(N, \delta)}) \cap K_\delta(N, \delta)$ for some $\frac{\alpha}{2} < \delta < \alpha$. Thus we may apply

Corollary 4.14 to obtain the measure $\hat{m}_p$ defined at p which is $A_n(p)$ -invariant and su -invariant. Hence, if we define

$$Z_p := \{z(p), \overline{z(p)} : z(p) \in \text{supp}(\hat{m}_p)\},$$

then Z_p consists of at most two points. Moreover, it is clear that Z_p is su -invariant, $A(p)$ -invariant, and invariant under complex conjugation.

Since q is also a periodic point such that $L(A, q) < \frac{\alpha}{2}$, we can certainly find $\frac{\alpha}{2} < \delta < \alpha$ so that both p and q belong to $\text{supp}(\mu|_{K(N, \delta)}) \cap K(N, \delta)$. Thus Z_p and Z_q are both defined. By su -invariance of Z_p and $p_0 = q_0$, we have

$$H_{q, q \wedge p}^u(Z_q) = H_{p, q \wedge p}^s(Z_p)$$

as desired. $\square$

5 Positivity of the Lyapunov exponent I

Throughout this section we assume that $\Omega \subseteq \mathcal{A}^{\mathbb{Z}}$ is a subshift of finite type and μ is a T -ergodic probability measure that is fully supported on Ω and has a local product structure. We fix a non-constant $f \in C^\alpha(\Omega, \mathbb{R})$ and consider the one-parameter family of Schrödinger cocycles (T, A^E) . We shall apply the techniques from Sect. 4 to study the positivity property of the Lyapunov exponent.

In Sect. 5.1, we show under a very general condition that the set of energies with zero Lyapunov exponent is a discrete set. In Sect. 5.2, we apply the same techniques to the scenario where we have global existence of holonomies and obtain a stronger result for the corresponding Schrödinger cocycles. Namely, we show that under the same general condition, the set of energies with zero Lyapunov exponent is a finite set. Global existence of the holonomies may be obtained if the $\|\cdot\|_\infty$ norm of the sampling function is small or if the sampling function is locally constant.

5.1 General case: positivity away from a discrete set

Throughout this subsection, we assume that $E_0 \in \mathbb{R}$ is an accumulation point of

$$\mathcal{Z}_f = \{E : L(E) = 0\}.$$

Clearly,

$$E_0 \in \Sigma, \tag{5.1}$$

since $\mathcal{Z}_f \subseteq \Sigma$ and Σ is closed.

Definition 5.1 We say that a periodic point q is γ -bunched at E if

$$2L(A^E, q) < \gamma \leq \alpha. \quad (5.2)$$

We fix p to be a periodic point that is $\frac{\alpha}{2}$ -bunched at E_0 . We let n_p denote the period of p .

Lemma 5.2 $L(A^{E_0}, p) = 0$. In particular, $E_0 \in \sigma(H_p)$.

Proof Let $E_n \rightarrow E_0$, $E_n \neq E_0$ be a sequence in $\mathcal{Z}_f$. Recall from (5.1) that $E_0 \in \Sigma$.

Assume that p is hyperbolic for E_0 (i.e., the matrix $A_{n_p}^{E_0}(p)$, which serves as the monodromy matrix at energy E_0 for the periodic potential associated with p , is hyperbolic). Then p is still hyperbolic and $\frac{\alpha}{2}$ -bunched in a small neighborhood J of E_0 . Let $E_0 \in \sigma(H_{\omega'})$ for some $\omega' \in \Omega$. By Proposition 2.10 and by choosing $\delta > 0$ small, we have $J \cap \sigma(H_\omega) \neq \emptyset$ for any ω such that $\text{Orb}(\omega) \cap B_\delta(\omega') \neq \emptyset$. On the other hand, by Proposition 2.7 there is an $r = r(\delta)$ so that for any $I_1 = [0, n_1] \subseteq \mathbb{Z}$, there is a periodic orbit q with period $n_q = n_1 + r + 1$ so that $d(T^j q, T^j p) < \delta$ for all $1 \leq j \leq n_1$ and $d(T^{n_1+r+1} q, \omega') < \delta$. An immediate consequence is that $\sigma(H_q) \cap J \neq \emptyset$. Moreover, as n_1 goes to infinity, it clearly holds that $L(A^E, q)$ tends to $L(A^E, p)$ uniformly for all $E \in J$. In particular, by choosing n_1 large, we have that q is $\frac{\alpha}{2}$ -bunched for all $E \in J$ as well. We fix such a periodic point q .

Clearly, $p_0 = q_0$. Thus we may define

$$H^E := H_{q \wedge p, q}^{u, E} \cdot H_{p, q \wedge p}^{s, E} \quad (5.3)$$

for each $E \in J$. Here $H_{p, q \wedge p}^{s, E}$ and $H_{q, q \wedge p}^{u, E}$ are the holonomies corresponding to A^E , which are well-defined since both p and q are $\frac{\alpha}{2}$ -bunched through J . Moreover, they are holomorphic on J since they are limits of uniformly convergent sequences of holomorphic functions $H^{s, n}(E)$ or $H^{u, n}(E)$ on J . Thus we have that $E \mapsto H^E$ is analytic. Let $Z_p = Z_p(E_n)$ be as in Corollary 4.15. By passing to a subsequence, we may assume that

$$Z_p(E_n) = \begin{cases} \{s(E_n)\} & \text{for all } n, \\ \{u(E_n)\} & \text{for all } n, \text{ or} \\ \{s(E_n), u(E_n)\} & \text{for all } n. \end{cases} \quad (5.4)$$

Thus, we may extend the definition of $Z_p(E)$ to all $E \in J$ so that $Z_p(E)$ consists of one or two functions that are analytic on J . By Corollary 4.15,

$$Z_q(E) := H^E \cdot Z_p(E) \quad (5.5)$$

is invariant under the monodromy matrix of q for infinitely many E_n . By analyticity, it follows that $Z_q(E)$ is invariant by the monodromy matrix of q for every $E \in J$. Since $Z_p(E)$ is real, so is $Z_q(E)$, which implies that the absolute value of the trace of the monodromy matrix of q cannot become smaller than 2 anywhere in J . Since J is open and $E \in \sigma(H_q)$ cannot be an isolated point, we must have $J \cap \sigma(H_q) = \emptyset$, which contradicts our choice of q . It follows that p is not hyperbolic for E_0 . In particular, $E_0 \in \sigma(H_p)$. $\square$

Choose a small open disk $D \subseteq \mathbb{C}$ around E_0 such that p is $\frac{\alpha}{2}$ -bunched for all energies E in the closed disk $\bar{D}$. Recall that by Proposition 2.11, $\Delta(E) = \text{Tr}(A_{n_p}^E(p))$ is monotonic on each connected component of $\Delta^{-1}(-2, 2)$. Thus we may also assume that D is small enough so that, through $\bar{D} \setminus \{E_0\}$, $\Delta(E)$ is different from $-2, 2, 0$. According to Sect. 2.2, if $E_0 \notin \partial(\sigma(H_p))$, we can then define two holomorphic functions $u, s : D \rightarrow \mathbb{CP}^1$, distinct everywhere, such that $u(E)$ and $s(E)$ are eigendirections of $A_{n_p}^E(p)$; otherwise we can still define holomorphic functions u, s on the ramified (at E_0) double cover of $\pi : \tilde{D} \rightarrow D$, giving (distinct) eigendirections when $\tilde{E} \in \tilde{D} \setminus \{E_0\}$, but taking as value at E_0 the single real eigendirection of $A_{n_p}^{E_0}(p)$. Moreover, for $\pi(\tilde{E}) \in \mathbb{R}$, $s(\tilde{E})$ and $u(\tilde{E})$ are real if and only if $\pi(\tilde{E})$ not in the interior of $\sigma(H_p)$.

Lemma 5.3 *If q is a periodic point that is $\frac{\alpha}{2}$ -bunched through $E \in \bar{D}$ and $q_i = p_j$ for some i, j , then $\sigma(H_p) \cap D = \sigma(H_q) \cap D$.*

Proof Since $\sigma(H_\omega) = \sigma(H_{T^n \omega})$ for any ω and for any n , we may assume that $p_0 = q_0$. Then similarly to the proof of Lemma 5.2, we define $H^E = H_{q \wedge p, q}^{u, E} \cdot H_{p, q \wedge p}^{s, E}$ for $E \in D$. Since $E_0 \in \sigma(H_p)$, we consider two different cases.

If $E_0 \notin \partial(\sigma(H_p))$, then $D \cap \sigma(H_p) = D \cap \mathbb{R}$ by our choice of D , and $Z_p(E_n)$ is a nonempty subset of $\{u(E_n), s(E_n)\}$. Following the same argument that showed that $A_{n_q}^{E_0}(q)$ has a real eigendirection in the proof of Lemma 5.2, we obtain that $A_{n_q}^E(q)$ of the present lemma has a non-real eigendirection for all $E \in D \cap \mathbb{R}$. This implies that $D \cap \mathbb{R} \subseteq \sigma(H_q)$, and the claim follows in this case.

If $E_0 \in \partial(\sigma(H_p))$, then by our choice of D we have $D \cap \Sigma_p$ is either $[E_0, E_+)$ or $(E_-, E_0]$, where

$$(E_-, E_+) = D \cap \mathbb{R}. \quad (5.6)$$

For simplicity, we will assume that $\text{int } D \cap \Sigma = [E_0, E_+)$. Recall $\pi : \tilde{D} \rightarrow D$ is the double cover map of D ramified at E_0 . For each n , choose a preimage $\tilde{E}_n \in \pi^{-1}(E_n)$. Then $Z_p(E_n)$ is a subset of $\{u(\tilde{E}_n), s(\tilde{E}_n)\}$. As in the proof

of Lemma 5.2 and up to replacing E_n by a subsequence, $Z_p(E_n)$ is always of the form $\tilde{Z}_p(\tilde{E}_n)$, where

$$\tilde{Z}_p(\tilde{E}) = \begin{cases} \{s(\tilde{E})\} & \text{for all } \tilde{E} \in \tilde{D}, \\ \{u(\tilde{E})\} & \text{for all } \tilde{E} \in \tilde{D}, \text{ or} \\ \{s(\tilde{E}), u(\tilde{E})\} & \text{for all } \tilde{E} \in \tilde{D}. \end{cases} \quad (5.7)$$

Notice that if $\pi(\tilde{E}) \in (E_-, E_0)$, then $\tilde{Z}_p(\tilde{E})$ consists of real directions; if $\pi(\tilde{E}) \in (E_0, E_+)$, then $\tilde{Z}_p(\tilde{E})$ consists of non-real directions. We again define

$$\tilde{Z}_q(\tilde{E}) := H^{\pi(\tilde{E})} \cdot \tilde{Z}_p(\tilde{E}).$$

Then $\tilde{Z}_p(\tilde{E})$ is invariant under $A_{n_q}^{\pi(\tilde{E})}(q)$ whenever $\tilde{E} = \tilde{E}_n$. By the fact that $H^{\pi(\cdot)}$, u , and s are all holomorphic on $\tilde{D}$, it follows that $\tilde{Z}_p(\tilde{E})$ is invariant under $A_{n_q}^{\pi(\tilde{E})}(q)$ for all $\tilde{E} \in \tilde{D}$. This implies that $A_{n_q}^E(q)$ has at least one real eigendirection for $E \in (E_-, E_0)$ and has at least one non-real eigendirection for $E \in (E_0, E_+)$. This can only happen when $D \cap \sigma(H_q) = [E_0, E_+)$, and the claim follows in this case. $\square$

Lemma 5.4 *If q is any periodic point, then $\sigma(H_q) \cap D = \sigma(H_p) \cap D$.*

Proof Fix an arbitrary periodic point q^* . Let us say that a periodic point q is (ϵ, δ) -good, $0 < \delta < \epsilon < 1$, if it spends at least a $1 - \epsilon$ proportion of its iterates within distance δ of p , and at least a $\epsilon/2$ proportion of its iterates within distance δ of q^* . By Proposition 2.7 and similar to the argument leading to the choice of q in the proof of Lemma 5.2, we see that the set of (ϵ, δ) -good periodic points is not empty for any choice of $0 < \delta < \epsilon < 1$. Moreover, if ϵ is sufficiently small, then an (ϵ, δ) -good q is $\frac{\alpha}{2}$ -bunched for energies $E \in \tilde{D}$. Moreover, since certain iterates of q are close to p , we clearly have $q_i = p_j$ for some i, j . By Lemma 5.3, it then holds that $\sigma(H_p) \cap D = \sigma(H_q) \cap D$ for all such q 's. We fix such a small ϵ for the remainder of this proof.

First we show $\sigma(H_{q^*}) \cap D \subseteq \sigma(H_p) \cap D$. If this is not true, then there is some $E_0 \in (\sigma(H_{q^*}) \cap D) \setminus \sigma(H_p)$. In particular, we have

$$\varepsilon := \min\{d(E_0, \sigma(H_p)), d(E_0, \partial D)\} > 0.$$

Then for an (ϵ, δ) -good periodic point q , we also have that

$$\varepsilon = \min\{d(E_0, \sigma(H_q)), d(E_0, \partial D)\} > 0.$$

By Proposition 2.10 and the fact that $\text{Orb}(q) \cap B_\delta(q^*) \neq \emptyset$ for (ϵ, δ) -good points, we have for sufficiently small δ and an (ϵ, δ) -good periodic point q

that

$$\sigma(H_{q^*}) \subseteq B_{\frac{\varepsilon}{2}}(\sigma(H_q)).$$

Clearly, this implies that $d(E_0, \sigma(H_q)) < \frac{\varepsilon}{2}$ and we obtain a contradiction. So the first part follows.

Now we show $\sigma(H_p) \cap D \subseteq \sigma(H_{q^*}) \cap D$. Suppose this is not the case. By the first part, there is an $E_0 \in (\sigma(H_p) \cap D) \setminus \sigma(H_{q^*})$. In particular, we have

$$\tilde{\varepsilon} := \min\{d(E_0, \sigma(H_{q^*})), d(E_0, \partial D)\} > 0.$$

Notice that $\text{Orb}(q^*) \cap B_\delta(T^m q) \neq \emptyset$ for an (ϵ, δ) -good periodic point p and for some $m \in \mathbb{Z}$. Thus by Proposition 2.10, and by choosing δ small (the smallness of which is independent of q or q^*), we have

$$\sigma(H_{T^m q}) \subseteq B_{\frac{\tilde{\varepsilon}}{2}}(\sigma(H_{q^*})).$$

Since $\sigma(H_{T^m q}) = \sigma(H_q)$ and $E_0 \in \sigma(H_p) \cap D = \sigma(H_q) \cap D$, we obtain

$$d(E_0, \sigma(H_{q^*})) < \frac{\tilde{\varepsilon}}{2},$$

which is a contradiction and the lemma follows. $\square$

By (2.17), the spectrum Σ is the closure of the union of the spectra of periodic points. Thus Lemma 5.4 implies that $\Sigma \cap D = \sigma(H_p) \cap D$. For each T -ergodic measure ν on Ω , we let Σ_ν denote the set such that $\sigma(H_\omega) = \Sigma_\nu$ for ν almost every ω ; see, for example, [32].

Lemma 5.5 *For any T -ergodic measure ν on Ω , we have $\Sigma_\nu \cap D = \sigma(H_p) \cap D$. Moreover, $L(A^E; \nu) = 0$ for all $E \in \sigma(H_p) \cap D$. In particular, $L(E) = 0$ for all such E 's.*

Proof Since we have $\sigma(H_\omega) \subseteq \Sigma$ for each $\omega \in \Omega$ and $\Sigma \cap D = \sigma(H_p) \cap D$, it clearly holds that $\Sigma_\nu \cap D \subseteq \sigma(H_p) \cap D$. On the other hand, if $E \notin \Sigma_\nu$, then the sequence $\{A^E(T^n \omega)\}_{n \in \mathbb{Z}}$ is uniformly hyperbolic for ν -almost every $\omega \in \Omega$, which in turn implies that $L(A^E; \nu) > 0$; see, for example, [43, Theorem 3]. Thus $L(A^E; \nu) = 0$ implies that $E \in \Sigma_\nu$. So we only need to prove the second part of the lemma.

Assume that the statement is false. In other words, we have $L(A^E; \nu) > 0$ for some $E \in \sigma(H_p) \cap D$. By [30, Theorem 3], for each $\epsilon > 0$, there is a periodic point $q \in \Omega$ so that $|L(A^E; \nu) - L(A^E, q)| < \epsilon$. Thus there is periodic point $q \in \Omega$ so that $L(A^E, q) > 0$. In particular, $E \notin \sigma(H_q)$, which contradicts Lemma 5.4, concluding the proof. $\square$

Lemma 5.6 *Let E_0 be an accumulation point of $\mathcal{Z}_f = \{E : L(E) = 0\}$. Assume that there exists a periodic point p that is $\frac{\alpha}{2}$ -bunched at E_0 . Then*

- (1) *The connected component I of E_0 in the spectrum is isolated,*
- (2) *$L(A^E; \nu) = 0$ for all $E \in I$ and all T -ergodic measure ν on Ω .*

Proof By Lemma 5.2, $E_0 \in \sigma(H_p)$. Let I be the connected component of $\sigma(H_p)$ that contains E_0 . Notice that p is $\frac{\alpha}{2}$ -bunched for every $E \in I$ since $L(A^E, p) = 0$ on I . Let S be the set of accumulation points of $\mathcal{Z}_f \cap I$. It is clearly a closed and non-empty subset of I since $E_0 \in S$. Moreover, applying Lemma 5.5 to any $E \in S$, we see that there is open disk D around E so that $L(E) = 0$ on $\sigma(H_p) \cap D$ which contains $I \cap D$. This implies $I \cap D \subset S$. Thus S is open in I as well. Thus we have $S = I$. Clearly $I \subseteq \Sigma$. Applying Lemma 5.5 to the boundary points of I , we obtain $\Sigma \cap D = \sigma(H_p) \cap D$ for some disk D around the boundary points. Thus, I is an isolated component of Σ . Applying Lemma 5.5 to all $E \in I$ again, we obtain that $L(A^E, \nu) = 0$ for all $E \in I$ and for all T -ergodic measure ν on Ω . $\square$

We have now collected all the tools to prove our main theorem.

Proof of Theorem 1.1 Suppose to the contrary that there are $E_0 \in \{E : L(E) = 0\}$ and $E_n \in \{E : L(E) = 0\} \setminus \{E_0\}$, $n \in \mathbb{Z}_+$, such that $E_n \rightarrow E_0$ as $n \rightarrow \infty$. Since $L(E_0) = 0$, we can choose a $\frac{\alpha}{2}$ -bunched periodic point by [30, Theorem 3]. It now follows from Lemma 5.6 that E_0 belongs to a non-degenerate compact interval I , which is a connected component of Σ , as well as of all periodic spectra $\sigma(H_p)$. In particular, for the fixed point of T , the unique connected component of its spectrum is an interval of length 4. Since having such a connected component is only possible for constant periodic potentials by Proposition 2.13, it follows that the potential associated with each periodic point must be constant. This implies that f itself must be constant; contradiction. $\square$

5.2 Special cases: positivity away from a finite set

In this subsection we consider sampling functions $f : \Omega \rightarrow \mathbb{R}$ for which we have global existence of the holonomies in the sense that the cocycle A^E admits *canonical holonomies* as defined in Sect. 2.1.5 for all E in a complex neighborhood of the convex hull $\mathcal{U} \subseteq \mathbb{C}$ of the spectrum Σ_f . Since A^E depends on E holomorphically, we obtain that the holonomies are holomorphic on $\mathcal{U}$ as well. In this case, we are able to improve the result we obtained in Sect. 5.1.

There are two types of f for which we have such global existence of holonomies. One is the set of $f \in C^\alpha(\Omega, \mathbb{R})$ for which A^E is fiber bunched in the sense of Definition 5.7 below for every E in the convex hull of the spectrum Σ . Such an f will be called *globally fiber bunched* (or just *globally bunched*). The other is the set of locally constant f 's.

5.2.1 Fiber bunching and existence of holonomies

Definition 5.7 We say that $A \in C^\alpha(\Omega, \mathrm{SL}(2, \mathbb{C}))$ is *fiber bunched* if there exists $n_0 \geq 1$ such that for every $\omega \in \Omega$, we have

$$\|A_{n_0}(\omega)\|^2 < e^{\alpha n_0}. \quad (5.8)$$

Equivalently, there is $\theta < \alpha$ such that $\|A_{n_0}(\omega)\|^2 < e^{\theta n_0}$ for every $\omega \in \Omega$.

Note that fiber bunching is clearly a C^0 -open condition. A fiber bunched cocycle has canonical holonomies as defined in Sect. 2.1.5. In fact, we can run the proof of Lemma 4.2 to show that for $\omega' \in W_{\mathrm{loc}}^u(\omega)$, $H_{\omega, \omega'}^{u, n} = A_{-n}(\omega')^{-1} \cdot A_{-n}(\omega)$ converges uniformly on Ω to the unstable holonomy $H_{\omega, \omega'}^u$. Similarly for $\omega' \in W_{\mathrm{loc}}^s(\omega)$, we have that $H_{\omega, \omega'}^{s, n} = A_n(\omega')^{-1} A_n(\omega)$ converges uniformly on Ω to the stable holonomy $H_{\omega, \omega'}^s$. Indeed, to obtain the uniform convergence to holonomies, the only condition we used in the proof of Lemma 4.2 is the condition in (4.3), which is exactly the fiber bunching condition (5.8). We also note the following: If $A^t \in C^\alpha(\Omega, \mathrm{SL}(2, \mathbb{C}))$, t in some domain $U \subseteq \mathbb{C}$, is a continuous family such that $t \mapsto A^t(\omega)$ is holomorphic for every $\omega \in \Omega$ and A^t is fiber bunched for every t , then the stable and unstable holonomies depend holomorphically on t . Indeed, in this case, the holonomies are limits of uniformly convergent sequences of holomorphic functions. In particular, we may consider Schrödinger cocycles A^E with sampling function $f \in C^\alpha(\Omega, \mathbb{R})$. If $\|f\|_\infty$ is sufficiently small, then A^E is fiber bunched in a complex neighborhood of the convex hull of the spectrum Σ . To see this, we first see that $\begin{pmatrix} E & -1 \\ 1 & 0 \end{pmatrix}$ is fiber bunched for all $E \in [-2, 2]$ since they are all elliptic or parabolic. By openness of fiber bunching, we then have that A^E is fiber bunched for all E in a complex neighborhood of $[-2, 2]$ provided $\|f\|_\infty$ is sufficiently small. If necessary, we can then choose $\|f\|_\infty$ smaller so that the convex hull of Σ_f is contained in such an open neighborhood. Thus f is globally bunched.

5.2.2 Locally constant cocycles

The other class for which the canonical holonomies exist for obvious reasons is defined as follows.

Definition 5.8 We say that $A : \Omega \rightarrow \mathrm{SL}(2, \mathbb{R})$ is *locally constant* if there exists a n_0 such that for each $\omega \in \Omega$, $A(\omega)$ depends only on the cylinder set $[-n_0; \omega_{-n_0}, \dots, \omega_{n_0}]$.

Evidently, locally constant cocycles are α -Hölder continuous for all $\alpha > 0$. Locally constant cocycles might not be fiber bunched. However, the

holonomies exist trivially. Indeed, if A is locally constant, then there is a $n_0 \in \mathbb{Z}_+$ so that for all ω and all $n > n_0$ we have

$$H_{\omega, \omega^\tau}^{\tau, n} = H_{\omega, \omega^\tau}^{\tau, n_0},$$

where $\tau \in \{s, u\}$ and $\omega^\tau \in W_{\text{loc}}^\tau(\omega)$. Thus $H_{\omega, \omega^\tau}^{\tau, n_0}$ are exactly the holonomies. Now we consider Schrödinger cocycles A^E with potential $f : \Omega \rightarrow \mathbb{R}$. If there is a $n_0 \in \mathbb{Z}_+$ such that $f(\omega)$ depends only on $[-n_0; \omega_{-n_0}, \dots, \omega_{n_0}]$, then A^E is locally constant for all $E \in \mathbb{C}$. In other words, a locally constant sampling function induces locally constant Schrödinger cocycle maps.

5.2.3 Energies admitting an su -state

Again, our objective is to study the energies for which $L(E) = L(A^E, \mu) = 0$. We will point out how the desired statements will follow by simple specialization of the proofs of the lemmas in Sect. 4.

Assume that $A \in C^0(\Omega, \text{SL}(2, \mathbb{R}))$ has canonical holonomies $H_{\omega, \omega'}^\tau$, where $\tau \in \{s, u\}$. Recall that an su -state for A is a (T, A) -invariant measure m with a disintegration $\{m_\omega : \omega \in \Omega\}$ that is invariant under the cocycle and the holonomies. In particular, for μ -almost every $\omega \in \Omega$, we have

- (1) $A(\omega)_* m_\omega = m_{T\omega}$,
- (2) $(H_{\omega, \omega'}^s)_* m_\omega = m_{\omega'}$ for every $\omega' \in W_{\text{loc}}^s(\omega)$.
- (3) $(H_{\omega, \omega'}^u)_* m_\omega = m_{\omega'}$ for every $\omega' \in W_{\text{loc}}^u(\omega)$.

Then we have the following invariance principle:

Proposition 5.9 *If $L(A, \mu) = 0$, then there exists an su -state for A .*

Proof This follows from Proposition 4.7, noting that the canonical holonomies exist and are continuous on Ω , and therefore the conditions (4.5) and (4.6) are automatically satisfied. $\square$

One of the main properties of su -states is the following.

Proposition 5.10 *If m is an su -state, then it admits a disintegration for which the conditional measures m_ω depend continuously on ω and are both s -invariant and u -invariant.*

Proof We take an su -state m . Then we run the proof of Lemma 4.9 where we constructed the disintegration $\tilde{m}$ which is continuous on $\text{supp}(K_\delta) \cap K_\delta$. In the present setting, we have $K_\delta = \Omega$ since we have canonical holonomies. The result follows. $\square$

By continuity and almost everywhere coincidence, all the invariance properties in the definition of su -states may then hold true for every $\omega \in \Omega$. From now on, we always choose such a disintegration for an su -state m .

5.2.4 Finiteness of the set of energies admitting su -states

Now we return to the Schrödinger case. Recall that by general principles, $L(E) = 0$ implies $E \in \Sigma \subseteq \mathbb{R}$. Since each of these real energies gives rise to an su -state for A^E , let us consider the following set (whose dependence on μ and f we leave implicit):

$$\mathcal{F} = \{E \in \Sigma : \text{there is an } su\text{-state for } A^E\}. \quad (5.9)$$

Lemma 5.11 *Suppose that $0 < \alpha \leq 1$ and $f \in C^\alpha(\Omega, \mathbb{R})$ is globally bunched or locally constant. Assume that $\mathcal{F}$ is infinite. Let $p, q \in \Omega$ be two periodic points of T . Then $\sigma(H_p) = \sigma(H_q)$.*

Proof First, we consider the case that $p_i = q_j$ for some $i, j \in \mathbb{Z}$. Since $\sigma(H_\omega) = \sigma(H_{T^n \omega})$ for any ω and for any n , we may assume that $p_0 = q_0$. Recall in this case there is a unique $q \wedge p \in W_{\text{loc}}^u(q) \cap W_{\text{loc}}^s(p)$. Assume that n_p is the period of p and n_q is the one of q . By our choice of f , we may choose $\tilde{\Sigma} \subseteq \mathbb{R}$ to be a compact interval containing the spectrum Σ and $\mathcal{U} \subseteq \mathbb{C}$ to be a complex neighborhood of $\tilde{\Sigma}$ where A^E has canonical holonomies for all $E \in \mathcal{U}$. Recall that under the conditions of the present lemma, the holonomies are holomorphic functions on $\mathcal{U}$.

By the arguments in the proof of Lemma 4.9, Corollary 4.15, and the existence of canonical holonomies, we can find for each periodic $\omega \in \Omega$, a subset $Z_\omega \subseteq \mathbb{CP}^1$ consisting of at most two points that is invariant under $A(\omega)$ and the holonomies. In particular, for the periodic point p with period n_p , Z_p is invariant under $A_{n_p}(p)$ and

$$H_{p, q \wedge p}^s(Z_p) = H_{q, q \wedge p}^u(Z_q) \text{ whenever } q_0 = p_0. \quad (5.10)$$

Note that if $\text{Tr}(A_{n_p}(p)) \neq 0$, then Z_p must be a subset of the eigendirections of $A_{n_p}(p)$. In particular, for $A_{n_p}(p)$ with nonzero trace, $A_{n_p}(p)$ is elliptic if and only if Z_p is non-real. We let $\{s(E), u(E)\}$ denote the pair of eigendirections of $A_{n_p}^E(p)$. Note that both $s(E)$ and $u(E)$ are continuous on $\tilde{\Sigma}$ and analytic on each spectral gap or on the interior of each connected component of $\sigma(H_p)$. We define

$$H^E = H_{q \wedge p, q}^{u, E} \cdot H_{p, q \wedge p}^{s, E}, \quad (5.11)$$

which are holomorphic in E on a complex neighborhood $\mathcal{U}$ of $\tilde{\Sigma}$.

Let E_0 be an accumulation point of $\mathcal{F}$. Then, similarly to the proof of Lemma 5.2 or 5.3, we can find a sequence $\{E_n\}_{n \geq 1}$ in $\mathcal{F}$ so that $E_n \rightarrow E_0$,

$E_n \neq E_0$, and

$$Z_p(E_n) = \begin{cases} \{s(E_n)\} & \text{for all } n \geq 1, \\ \{u(E_n)\} & \text{for all } n \geq 1, \text{ or} \\ \{s(E_n), u(E_n)\} & \text{for all } n \geq 1. \end{cases} \quad (5.12)$$

Thus we may extend the domain of $Z_p(\cdot)$ from $\{E_n, n \geq 1\}$ to $\mathcal{U}$. Then we define

$$Z_q(E) := H^E(Z_p(E)) \quad (5.13)$$

and we get that $Z_q(E_n)$ is invariant under $A_{nq}^{E_n}(q)$ for all $n \geq 1$. By the continuity and analyticity properties of H^E , $s(E)$, and $u(E)$, we obtain the following conclusions: if E_0 is in a spectral gap, then $Z_q(E)$ is invariant under $A_{nq}^{E_n}(q)$ for all E in the closure of that spectral gap; if E_0 is the interior of a connected component of $\sigma(H_p)$, then $Z_q(E)$ is invariant under $A_{nq}^{E_n}(q)$ for all E in that connected component.

Now by the same arguments as in the proof of Lemma 5.2, if E_0 is in a spectral gap of H_p , then it is away from $\sigma(H_q)$ with a uniform distance for all periodic points q . But $E_0 \in \Sigma$ since it is an accumulation point of $\mathcal{F}$. Thus E_0 can be approximated by $\sigma(H_q)$ for a certain choice of q , a contradiction. We may conclude that $E_0 \in \sigma(H_p)$. So we may let $I \subseteq \sigma(H_p)$ be the connected component containing E_0 . Now we claim that

$$Z_q(E) = H^E(Z_p(E)) \text{ is invariant under } A_{nq}^E(q) \text{ for all } E \in \tilde{\Sigma}. \quad (5.14)$$

If E_0 is in the interior of I , we have already obtained that $\tilde{Z}_q(E)$ is invariant under $A_{nq}^E(q)$ for all $E \in I$. If E_0 belongs to the boundary of I , then similarly to the proof of Lemma 5.3, there is an open disk D centered at E_0 with ramified (at E_0) double cover $\pi : \tilde{D} \rightarrow D$ so that s and u are holomorphic on $\tilde{D}$. Thus we may assume $Z_p(E) = \tilde{Z}_p(\tilde{E})$ where $\tilde{E} \in \pi^{-1}(E)$ and

$$\tilde{Z}_p(\tilde{E}) = \begin{cases} \{s(\tilde{E})\} & \text{for all } \tilde{E} \in \tilde{D}, \\ \{u(\tilde{E})\} & \text{for all } \tilde{E} \in \tilde{D}, \text{ or} \\ \{s(\tilde{E}), u(\tilde{E})\} & \text{for all } \tilde{E} \in \tilde{D}. \end{cases} \quad (5.15)$$

Then we define

$$\tilde{Z}_q(\tilde{E}) := H^{\pi(\tilde{E})}(\tilde{Z}_p(\tilde{E})), \quad (5.16)$$

so that $\tilde{Z}_p(\tilde{E})$ is invariant under $A_{n_q}^{\pi(\tilde{E})}(q)$ for infinitely many $\tilde{E}_n \in \tilde{D}$. By the fact that $H^{\pi(\cdot)}$, s , and u are holomorphic on $\tilde{D}$, we obtain that $\tilde{Z}_q(\tilde{E})$ is invariant under $A_{n_q}^{\pi(\tilde{E})}(q)$ for all $\tilde{E} \in \tilde{D}$. Descending to D , we obtain that

$$Z_q(E) := H^E(Z_p(E)) \quad (5.17)$$

is invariant under $A_{n_q}^E(q)$ for all $E \in D$. In particular, $Z_q(E)$ is invariant under $A_{n_q}^E(q)$ for all $E \in (E_0 - \rho, E_0 + \rho)$, where $\rho > 0$ is the radius of D .

By the analysis above, we obtain that no matter whether E_0 belongs to the boundary or to the interior of I , after a finite number of continuations, we get that $Z_q(E)$ is invariant under $A_{n_q}^E(q)$ for all $E \in \tilde{\Sigma}$, as claimed. As in the proof of Lemma 5.3, and by the fact that H^E is real for E real, we obtain that $Z_p(E)$ and $Z_q(E)$ are simultaneously real or non-real for all $E \in \tilde{\Sigma} \supseteq \Sigma$. This clearly implies that

$$\sigma(H_p) = \sigma(H_q) \text{ whenever } p_0 = q_0. \quad (5.18)$$

Now we remove the condition $p_i = q_j$ for some $i, j \in \mathbb{Z}$. As in the proof of Lemma 5.4, we can find a periodic point p' with some iterates very close to p and some very close to q . In particular, $p'_i = p_j$ for some $i, j \in \mathbb{Z}$ and $p'_k = q_m$ for some $k, m \in \mathbb{Z}$. Thus by the first case we consider above, we have

$$\sigma(H_p) = \sigma(H_{p'}) = \sigma(H_q). \quad (5.19)$$

This concludes the proof. $\square$

5.2.5 Proof of Theorem 1.3

Theorem 1.3 is an immediate consequence of the following theorem.

Theorem 5.12 *Suppose $0 < \alpha \leq 1$ and let $f \in C^\alpha(\Omega, \mathbb{R})$ be globally bunched or locally constant. If the periodic spectra associated with periodic points of T in Ω are not all identical, then $\{E : L(E) = 0\}$ is finite.*

Proof As $\{E : L(E) = 0\} \subseteq \mathcal{F}$, the statement follows from Lemma 5.11. $\square$

Remark 5.13 Theorem 5.12 is particularly easy to apply when T has a fixed point, as the latter property ensures the presence of a constant potential and all one needs to do in order to show that not all periodic spectra are the same is to use the non-constancy of the sampling function to produce a non-constant periodic potential. However, there are certainly cases of interest where the

base dynamics given by T is fixed-point-free. In this case Theorem 5.12 still provides a direct tool for proving that $\{E : L(E) = 0\}$ is finite for many globally bunched or locally constant $f \in C^\alpha(\Omega, \mathbb{R})$, one just needs to take a closer look at the resulting periodic spectra.

Remark 5.14 Consider (Ω^+, T_+, μ^+) and assume that we can lift μ^+ to an ergodic measure μ on (Ω, T) that has a local product structure. Then all our main results of this section, in particular Theorem 1.1 and Theorem 1.3, can be applied to $f \in C^\alpha(\Omega^+, \mathbb{R})$. Indeed, such an f can be lifted to an $\tilde{f} \in C^\alpha(\Omega, \mathbb{R})$ that depends only on the future. Then all our results follow since $L(\mu, A^{(E-\tilde{f})}) = L(\mu^+, A^{(E-f)})$.

6 Positivity of the Lyapunov exponent II

We first show that in the scenario of Sect. 5.2, we may remove the finite exceptional set for an open and dense subset of sampling functions. Then we apply similar arguments to the general case discussed in Sect. 5.1 and obtain that for a residual set of sampling functions, the discrete exceptional set can be removed. Throughout this section, we again assume that (Ω, T) is a subshift of finite type with a fully supported ergodic measure μ that has a local product structure. Note that for $0 < \alpha \leq 1$, the space $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ is a Banach space with the C^α norm defined by

$$\|A\|_{0,\alpha} = \|A\|_\infty + \sup_{\omega \neq \omega'} \frac{\|A(\omega) - A(\omega')\|}{d(\omega, \omega')^\alpha}, \quad (6.1)$$

where $\|A\|_\infty$ is the standard C^0 norm $\|A\|_\infty = \sup_{\omega \in \Omega} \|A(\omega)\|$. Similarly, the space $C^\alpha(\Omega, \mathbb{R})$ is a Banach space with a C^α norm that can be defined analogously. We say that a subset of $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ has *codimension infinity* if it is locally contained in finite unions of closed submanifolds with arbitrary codimension. The same notion can be defined when we consider a subspace or an open subset of $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$.

6.1 Special cases: uniform positivity in a dense open set

In this subsection, we assume that $A \in C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ is fiber bunched or locally constant, and hence admits canonical holonomies by our earlier discussion.

We first introduce the follow notion of typical cocycles.

Definition 6.1 We say A is *typical* if there are two periodic points p and q with periods n_p and n_q such that $p_0 = q_0$ and the following properties hold:

- (1) $A_{n_p}(p) \neq I_2$ and $\text{Tr}(A_{n_p}(p)) \neq 0$.
- (2) Let $\{s(p), u(p)\} \subseteq \mathbb{CP}^1$ be the set of eigendirections of $A_{n_p}(p)$. Then there is no $Z_p \subseteq \{s(p), u(p)\}$ so that $H_{q \wedge p, q}^u \cdot H_{p, q \wedge p}^s \cdot Z_p$ is invariant under $A_{n_q}(q)$.

Since the definition involves two periodic points p and q , we may more precisely say that A is *typical with respect to* (p, q) . Note that A might be typical with respect to many other pairs of periodic points as well. Clearly, the defining conditions of a typical cocycle are open in the C^0 topology. Thus they are open in the C^α topology as well.

The notion of a typical cocycle in the present scenario was first introduced in [7, 8]. Our version is slightly different from theirs. It is adapted for the proof of Theorem 1.5 below. In particular, employing the arguments from [7, 8], one can show the following result. We only sketch the proof for the convenience of the reader.

Proposition 6.2 *The set of typical cocycles as defined above forms a C^α -open and dense subset in the set of fiber bunched (resp., locally constant) cocycles. Moreover, the complement of the set of typical cocycles has codimension infinity.*

Proof Following the arguments from [7, 8], for each fixed pair of periodic points p and q with $p_0 = q_0$, the complement of the set cocycles satisfying conditions (1) and (2), denoted by $\mathcal{B}_{p,q}$, is seen to be contained in the union of a finite number of sets of the form

$$\{A : \mathcal{H}(A) = 0\},$$

where each $A \mapsto \mathcal{H}(A)$ is a C^1 submersion when restricted to suitable sets of $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$. Thus for each fixed pair (p, q) , one can show that $\mathcal{B}_{p,q}$ is a submanifold of $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ with positive codimension. Note that the complement of the set of typical cocycles is

$$\bigcap_{p, q \in \text{Per}(T): p_0 = q_0} \mathcal{B}_{p,q}.$$

Since there are infinitely many such pairs (p, q) , the set above is contained in a subset of $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ with codimension infinity. Thus, the complement of the set of typical cocycles has codimension infinity and the set of typical cocycles is open and dense. $\square$

Remark 6.3 Let us mention that one can have the following type of perturbation from [7, 8]: for each fixed pair of periodic points p and q with $p_0 = q_0$, one can modify the values of A at other points without changing its values at

p and q as well as without changing its holonomies on the local stable and unstable sets of these two points.

We first note the following consequence of our proof of Lemma 5.11, which also recovers one of the results in [7]:

Lemma 6.4 *Assume that the fiber bunched or locally constant $A \in C^\alpha(\Omega, \mathrm{SL}(2, \mathbb{R}))$ is typical. Then $L(A, \mu) > 0$. In particular, there is an open and dense subset $\mathcal{G}$ of fiber bunched or locally constant cocycles whose complement has codimension infinity and $L(A, \mu) > 0$ for all $A \in \mathcal{G}$.*

Proof Assume that $L(A, \mu) = 0$. Let p and q be two periodic points satisfying the conditions in the definition of typical cocycles. Then by the proof of Lemma 5.11, we know that there is a set $Z_p \subseteq \mathbb{CP}^1$ consisting of at most two points with the following properties:

- (1) Z_p is invariant under $A_{n_p}(p)$,
- (2) $H_{q \wedge p, q}^u \cdot H_{p, q \wedge p}^s \cdot Z_p$ is invariant under $A_{n_q}(q)$.

Since p and q satisfy the conditions stated in the definition of typical cocycles, we have that $A_{n_p}(p) \neq I_2$ and $\mathrm{Tr}(A_{n_p}(p)) \neq 0$. Thus property (1) implies that Z_p is a subset of $\{s(p), u(p)\}$. As a consequence, property (2) contradicts condition (2) of the definition of typical cocycles, concluding the proof. $\square$

Remark 6.5 Although Proposition 6.2 and Lemma 6.4 are stated for the space of general cocycles, $C^\alpha(\Omega, \mathrm{SL}(2, \mathbb{R}))$, they hold true if one restricts to the subspace of Schrödinger cocycles, that is, cocycles taking the form

$$A(\omega) = \begin{pmatrix} f(\omega) & -1 \\ 1 & 0 \end{pmatrix}.$$

Note that this subspace is equivalent to the space $C^\alpha(\Omega, \mathbb{R})$. Indeed, it is not difficult to see that the perturbation argument used in the proof of Proposition 6.2 works equally well when considering Schrödinger cocycles.

We note the following consequence of [4, Theorem 2.8].

Proposition 6.6 *Suppose (Ω, T) is a subshift of finite type and μ is T -ergodic with a local product structure. Let $f \in C^\alpha(\Omega, \mathbb{R})$ be globally fiber bunched or locally constant. Then $E \mapsto L(E)$ is continuous on $\mathbb{R}$.*

Indeed, [4, Theorem 2.8] implies that the Lyapunov exponent is continuous on the subspace of $C^\alpha(\Omega, \mathrm{SL}(2, \mathbb{R}))$ of globally fiber bunched or locally constant cocycles. If $f \in C^\alpha(\Omega, \mathbb{R})$ is globally fiber bunched or locally constant, then there is a connected compact interval $\hat{\Sigma}$ that contains the spectrum $\Sigma = \Sigma_f$ so that A^E is fiber bunched or locally constant for all $E \in \hat{\Sigma}$. Thus

$L(E)$ is continuous on $\hat{\Sigma}$. On the other hand, $L(E)$ is smooth outside of the spectrum as (T, A^E) is uniformly hyperbolic for $E \notin \Sigma$ and the Lyapunov exponent is pluriharmonic on the set of uniformly hyperbolic cocycles. Thus $L(E)$ is continuous on $\mathbb{R}$.

Proof of Theorem 1.5 We focus on the case where f is globally fiber bunched as the proof in the locally constant case is completely analogous.

Fix an $f \in C^\alpha(\Omega, \mathbb{R})$ that is non-constant and globally fiber bunched. Thus we may find a compact connected interval $\hat{\Sigma}$ whose interior contains the spectrum Σ_f so that $A^{(E-f)}$ is fiber bunched for each $E \in \hat{\Sigma}$. Note that fiber bunching is a C^0 open condition and

$$\forall E : \|A^{(E-f_1)} - A^{(E-f_2)}\|_\infty = \|f_1 - f_2\|_\infty. \quad (6.2)$$

Thus, for any open neighborhood $\mathcal{U}_f \subseteq C^\alpha(\Omega, \mathbb{R})$ of f that is sufficiently small, we have for each $g \in \mathcal{U}_f$ that $\Sigma_g \subseteq \hat{\Sigma}$ and $A^{(E-g)}$ is fiber bunched for all $E \in \hat{\Sigma}$. In the remaining part of the proof, we fix such a $\mathcal{U}_f$ and work inside it.

If $\sigma(H_{p,f}) = \sigma(H_{q,f})$ for all periodic points p and q , then by total disconnectedness of Ω , we can modify the value of f at q without changing its value along the orbit of p . On the other hand, if we choose E on the boundary of the spectrum of $\sigma(H_{q,f})$, we can certainly perturb f to g so that $L(A^{(E-g)}, q) > 0$. Thus we may perturb f to a g that is arbitrarily close to f with the property $\sigma(H_{p,g}) \neq \sigma(H_{q,g})$. Then we can instead work with g . Moreover, when perturbing f to g , we can certainly choose p and q so that $p_0 = q_0$.

Thus, we may assume without loss of generality that f is such that $\sigma(H_{p,f}) \neq \sigma(H_{q,f})$ for suitably chosen periodic points p and q such that $p_0 = q_0$. As described in Sect. 2.2, we again let $\{s(E), u(E)\}_{E \in \Sigma}$ be the pair of functions associated with the eigendirections of $A_{n_p}^{(E-f)}(p)$. Define $H^E = H_{q \wedge p, q}^{u, E} \cdot H_{p, q \wedge p}^{s, E}$. Then by the proof of Lemma 5.11, if we define $Z_p(E)$ to be

$$Z_p(E) = \begin{cases} \{s(E)\} & \text{for all } E \in \hat{\Sigma}, \\ \{u(E)\} & \text{for all } E \in \hat{\Sigma}, \text{ or} \\ \{s(E), u(E)\} & \text{for all } E \in \hat{\Sigma}, \end{cases} \quad (6.3)$$

then the set

$$\left\{ E \in \hat{\Sigma} : A_{n_q}^{(E-f)}(q) \cdot H^E \cdot Z_p(E) = H^E \cdot Z_p(E) \right\} \quad (6.4)$$

is finite. On the other hand, the set

$$\left\{ E \in \hat{\Sigma} : A_{n_p}^{(E-f)}(p) = \pm I_2 \text{ or } \text{Tr}(A_{n_p}^{(E-f)}(p)) = 0 \right\} \quad (6.5)$$

is finite as well. Combining the facts above, we then have that

$$\mathcal{B}_f := \left\{ E \in \hat{\Sigma} : A^{(E-f)} \text{ is not typical} \right\} \quad (6.6)$$

is finite. Note that for all $E \notin \mathcal{B}_f$, $A^{(E-f)}$ is typical with respect to (p, q) . By Remark 6.3 we can modify the values of $A^{(E-f)}$ at different points and keep its values at p and q , as well as their holonomies. In particular, by Remark 6.5, after a finite number of perturbations, we can perturb f to g with the following properties. There is a pair of periodic points (p', q') with $p'_0 = q'_0$ and $A^{(E-g)}$ is typical with respect to (p, q) for all $E \notin \mathcal{B}_f$ and typical with respect to (p', q') for all $E \in \mathcal{B}_f$. Thus we have that $A^{(E-g)}$ is typical for all $E \in \hat{\Sigma}$. By the fact that the defining properties of typical cocycles are open conditions with respect to the C^0 topology, property (6.2), and the compactness of $\hat{\Sigma}$, we obtain a neighborhood $\mathcal{U}_g \subseteq \mathcal{U}_f$ of g so that for each $h \in \mathcal{U}_g$, we have

$$L(A^{(E-h)}, \mu) > 0 \text{ for all } E \in \hat{\Sigma}.$$

By Proposition 6.6, $L(A^{(E-h)})$ is continuous on $\mathbb{R}$. On the other hand, it is well known that $(T, A^{(E-h)})$ is uniformly hyperbolic outside of Σ and $L(A^{(E-h)}, \mu)$ tends to ∞ as $|E|$ tends to ∞ . Combining all these statements, we find that for each $h \in \mathcal{U}_g$, we have

$$\inf_{E \in \mathbb{R}} L(A^{(E-h)}, \mu) > 0.$$

This concludes the proof. $\square$

6.2 General case: full positivity for generic sampling functions

In this subsection, we return to the general setting of Theorem 1.1. Note that in this case we have neither the canonical holonomies, nor global existence of holonomies. Moreover, the discrete set can in principle be infinite. To remove the discrete exceptional set, the price we need to pay is that we can only do it for C^α -generic sampling functions. For the remaining part of the section, we fix $0 < \alpha \leq 1$ and consider the space $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$.

We start with a new definition of typical cocycles that is adapted for the purpose of this section.

Definition 6.7 We say $A \in C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ is *typical* if there are two periodic points p and q with periods n_p and n_q , respectively, such that $p_0 = q_0$ and the following properties hold:

- (1) p and q are $\frac{\alpha}{2}$ -bunched, that is, $2L(A, p) < \frac{\alpha}{2}$ and $2L(A, q) < \frac{\alpha}{2}$.
- (2) $A_{n_p}(p) \neq I_2$ and $\text{Tr}(A_{n_p}(p)) \neq 0$.
- (3) Let $\{s(p), u(p)\} \subseteq \mathbb{CP}^1$ be the set of eigendirections of $A_{n_p}(p)$. Then there is no $Z_p \subseteq \{s(p), u(p)\}$ so that $H_{q \wedge p, q}^u \cdot H_{p, q \wedge p}^s \cdot Z_p$ is invariant under $A_{n_q}(q)$.

Note that the existence of the holonomies of p and q in condition (3) is guaranteed by condition (1). As in the previous subsection, we may also say that A is *typical with respect to* (p, q) , as the definition involves p and q .

Define

$$\mathcal{T}_\alpha := \{A \in C^\alpha(\Omega, \text{SL}(2, \mathbb{R})) : A \text{ is a typical cocycle}\}. \quad (6.7)$$

It is a standard fact that $A \mapsto L(A, \mu)$ is upper-semicontinuous on $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$. In particular, the set

$$\mathcal{L}_\alpha = \left\{A \in C^\alpha(\Omega, \text{SL}(2, \mathbb{R})) : 2L(A, \mu) < \frac{\alpha}{2}\right\} \quad (6.8)$$

is open in $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$. Again by [30, Theorem 3], if $2L(A, \mu) < \frac{\alpha}{2}$, there exists a periodic point p such that $2L(A, p) < \frac{\alpha}{2}$, that is, p is $\frac{\alpha}{2}$ -bunched. Then, as in the proof of Lemma 5.2, we may use the specification property to produce infinitely many pairs of $\frac{\alpha}{2}$ -bunched periodic points (p, q) so that $p_0 = q_0$. In particular, similarly to Proposition 6.2, we have the following:

Proposition 6.8 *Suppose (Ω, T) is a subshift of finite type and μ is a T -ergodic measure that has a local product structure. Consider the space $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ for $\alpha > 0$ and let $\mathcal{T}_\alpha$ and $\mathcal{L}_\alpha$ be defined as above. $\mathcal{T}_\alpha \cap \mathcal{L}_\alpha$ forms an open and dense subset of $\mathcal{L}_\alpha$. Moreover, $\mathcal{L}_\alpha \setminus \mathcal{T}_\alpha$ has codimension infinity in $\mathcal{L}_\alpha$.*

Similarly to Lemma 6.4, Proposition 6.8 has the following consequence, which has appeared in [39]. For simplicity, we define

$$\mathcal{P}_\alpha = \{A \in C^\alpha(\Omega, \text{SL}(2, \mathbb{R})) : L(A, \mu) > 0\}. \quad (6.9)$$

Lemma 6.9 *We have $\mathcal{T}_\alpha \subseteq \mathcal{P}_\alpha$. In other words, $L(A, \mu) > 0$ for each A that is typical. Moreover, the set $\mathcal{P}_\alpha$ contains an open and dense subset of $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ and the complement of $\mathcal{P}_\alpha$ has codimension infinity.*

Proof If $A \notin \mathcal{L}_\alpha$, then $L(A, \mu) \geq \frac{\alpha}{4} > 0$. If $A \in \mathcal{L}_\alpha$ is typical, then we may apply the proof of Lemma 6.4 to get $L(A, \mu) > 0$. However, here we have to use the full strength of Sect. 5.1. Specifically, $\frac{\alpha}{2}$ -bunching of p and q and the proof of Lemma 4.2 guarantee the existence of the holonomies associated with p and q . Then Lemmas 4.4 and 4.9 and Corollary 4.15 can be used to guarantee the existence and holonomy-invariance of Z_p and Z_q . Once we have all these tools, the proof of $L(A, \mu) > 0$ is then identical to the proof of Lemma 6.4.

Next, we want to show that the set $\mathcal{P}_\alpha$ contains an open and dense set. To this end, we fix any $A \in C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$. If there is an open neighborhood $\mathcal{U}_A$ of A such that for each $B \in \mathcal{U}_A$, $L(B, \mu) \geq \frac{\alpha}{4}$, then there is nothing we need to say. Otherwise, in any open neighborhood $\mathcal{U}$ of A , we can find a $B \in \mathcal{L}_\alpha$. Then by Proposition 6.8 and the proof above, we can find an open set $\mathcal{V} \subseteq \mathcal{U} \cap \mathcal{T}_\alpha$, which implies that $L(B, \mu) > 0$ for each $B \in \mathcal{V}$.

Finally, it is clear that the complement of $\mathcal{P}_\alpha$ is contained in $\mathcal{L}_\alpha \setminus \mathcal{T}_\alpha$, which has codimension infinity in $\mathcal{L}_\alpha$. Hence, the complement of $\mathcal{P}_\alpha$ has codimension infinity in $C^\alpha(\Omega, \text{SL}(2, \mathbb{R}))$ as well. $\square$

Note that this is an improved version of Lemma 6.4, as here we remove the assumption of global bunching or local constancy of f .

Now we are ready to generically remove the discrete set that appeared in Theorem 1.1.

Proof of Theorem 1.4 By Remark 6.5 and via the arguments from the proof of Lemma 6.9 we can show that the set

$$\mathcal{Z}_\alpha := \{f \in C^\alpha(\Omega, \mathbb{R}) : L(A^{(f)}, \mu) = 0\} \quad (6.10)$$

has codimension infinity in $C^\alpha(\Omega, \mathbb{R})$. In other words, $\mathcal{Z}_\alpha$ is locally contained in finite unions of closed submanifolds with arbitrary codimension. More precisely, for each $k \in \mathbb{Z}_+$ and each $f \in \mathcal{Z}_\alpha$, we can find an open neighborhood $\mathcal{U}_f$ of f and submanifolds $\mathcal{M}_j$, $1 \leq j \leq m$, each with codimension k , so that

$$(\mathcal{Z}_\alpha \cap \mathcal{U}_f) \subseteq \bigcup_{j=1}^m \mathcal{M}_j. \quad (6.11)$$

On the other hand, if we define the set $\mathcal{B}_\alpha$ to be

$$\mathcal{B}_\alpha := \{g \in C^\alpha(\Omega, \mathbb{R}) : E - g \in \mathcal{Z}_\alpha \text{ for some } E \in \mathbb{R}\}, \quad (6.12)$$

then for each $g \in \mathcal{B}_\alpha$, we can find $f \in \mathcal{Z}_\alpha$ and $E \in \mathbb{R}$ so that $g = E - f$. Thus $\mathcal{B}_\alpha$ is locally contained in finite unions of submanifolds of arbitrary codimension as well. Indeed, for the g and f above, we may just assume that

f is the one in (6.11). Thus for a fixed $k \in \mathbb{Z}_+$, for each $\mathcal{M}_j$ in (6.11), the set

$$\mathcal{N}_j := \{h \in C^\alpha(\Omega, \mathbb{R}) : h - E \in \mathcal{M}_j \text{ for some } E \in \mathbb{R}\}$$

may be viewed as a submanifold of $C^\alpha(\Omega, \mathbb{R})$ with codimension $k - 1$ whose local charts can be obtained from those of $\mathcal{M}_j$ and $E \in \mathbb{R}$. In particular, it is nowhere dense if $k \geq 2$. On the other hand, (6.11) clearly implies that the open neighborhood

$$\mathcal{U}_g = \mathcal{U}_f + E := \{h \in C^\alpha(\Omega, \mathbb{R}) : h - E \in \mathcal{U}_f\}$$

of g satisfies

$$(\mathcal{B}_\alpha \cap \mathcal{U}_g) \subseteq \bigcup_{j=1}^m \mathcal{N}_j.$$

Since $g \in \mathcal{B}_\alpha$ and $k \in \mathbb{Z}_+$ can be arbitrarily chosen, we obtain that $\mathcal{B}_\alpha$ is nowhere dense. Equivalently, we may say that the complement $\mathcal{B}_\alpha^c$ of $\mathcal{B}_\alpha$ is residual in $C^\alpha(\Omega, \mathbb{R})$. By definition of $\mathcal{B}_\alpha$, we have for each $f \in \mathcal{B}_\alpha^c$ that

$$L(A^{(E-f)}, \mu) > 0 \text{ for all } E \in \mathbb{R},$$

concluding the proof. $\square$

Remark 6.10 Similarly to Remark 5.14, all the main results in this section can be applied to Hölder continuous sampling functions defined on (Ω^+, T_+, μ^+) , where the lift μ of μ^+ has a local product structure. Indeed, in this case, $C^\alpha(\Omega^+, \mathbb{R})$ can be considered as a closed subspace of $C^\alpha(\Omega, \mathbb{R})$ whose elements depend only on the future. All the perturbations can then be performed within this subspace.

7 Applications

All of the results of this paper may be applied to Hölder continuous cocycles defined over any transitive Anosov diffeomorphism (or transitive, uniformly expanding differentiable map), where μ is taken to be the equilibrium state of a Hölder continuous potential. By a standard technique one can reduce the cocycles in question to Hölder continuous cocycles over a subshift of finite type via a Markov partition; see, for example, [14, 31]. Although the applicability is much wider, we will focus on a particular case as follows. It is standard result that if an invariant measure μ of a C^2 transitive Anosov diffeomorphism (or a C^2 transitive, uniformly expanding map) is absolutely continuous with respect

to the volume measure, then it is an equilibrium state of a Hölder continuous potential; see, for example, [14].

To illustrate this, we choose three differential models that have been widely studied in both the dynamical systems and mathematical physics communities. The first type of model is given by linear expanding maps of the circle,

$$T : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}, \quad Tx = kx, \quad k \geq 2, \quad (7.1)$$

and the measure is taken to be the Lebesgue measure m on $\mathbb{R}/\mathbb{Z}$. One may find some existing results for this case in [6, 9, 12, 18, 22, 36, 40, 42, 44]. In particular, the case $k = 2$ corresponds to the doubling map, which is the most difficult map to study within this family of maps, as it is the least mixing among them. The second type is given by hyperbolic automorphisms of $\mathbb{R}^d/\mathbb{Z}^d$, where μ is taken to be the Lebesgue measure m on $\mathbb{R}^d/\mathbb{Z}^d$. The most intensively studied case is the famous Arnold cat map, where

$$T : \mathbb{R}^2/\mathbb{Z}^2 \rightarrow \mathbb{R}^2/\mathbb{Z}^2, \quad T = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad (7.2)$$

and μ is taken to be the Lebesgue measure m on $\mathbb{R}^2/\mathbb{Z}^2$. One may find earlier results for this case in [12, 18, 36, 42]. It is clear that both linear expanding maps of the circle and hyperbolic toral automorphisms meet all the conditions necessary to apply our main theorems in Sects. 5 and 6. In particular, they all have a fixed point.

Our theorems then yield the following results. To unify the statements, we let (Ω, T, μ) be any of the following: $(\mathbb{R}/\mathbb{Z}, T_k, m)$, where $T_k x = kx$ and $k \geq 2$ is an integer; $(\mathbb{R}^d/\mathbb{Z}^d, T_A, m)$ where $d \geq 2$ and T_A is the hyperbolic toral automorphism generated by some hyperbolic $A \in \mathrm{SL}(d, \mathbb{Z})$. Recall that for a sampling function f , we set $L(E) = L(A^{(E-f)}, \mu)$ and define

$$\mathcal{Z}_f := \{E : L(E) = 0\} \subseteq \mathbb{R}. \quad (7.3)$$

For $0 < \alpha \leq 1$ and $\lambda > 0$, we $C_\lambda^\alpha(\Omega, \mathbb{R}) = \{f \in C^\alpha(\Omega, \mathbb{R}) : \|f\|_\infty < \lambda\}$.

Theorem 7.1 *Let (Ω, T, μ) be as above and let $0 < \alpha \leq 1$. For all non-constant $f \in C^\alpha(\Omega, \mathbb{R})$, $\mathcal{Z}_f$ is a discrete set. Moreover, $\mathcal{Z}_f = \emptyset$ for f 's in a residual subset of $C^\alpha(\Omega, \mathbb{R})$. There is $\lambda_0 = \lambda_0(\alpha) > 0$ such that $\mathcal{Z}_f$ is a finite set for all non-constant $f \in C_{\lambda_0}^\alpha(\Omega, \mathbb{R})$. Finally, there is an open and dense subset $\mathcal{O}^\alpha$ of $C_{\lambda_0}^\alpha(\Omega, \mathbb{R})$ such that for all $f \in \mathcal{O}^\alpha$, $\inf_{E \in \mathbb{R}} L(E) > 0$.*

If we introduce a coupling constant λ , then we have the following immediate consequence of Theorem 7.1.

Corollary 7.2 *Let (Ω, T, μ) and α be as in Theorem 7.1. Fix a non-constant $f \in C^\alpha(\Omega, \mathbb{R})$. Then $\mathcal{Z}_{\lambda f}$ is a discrete set for all $\lambda > 0$. Moreover, there is a $\lambda_0 = \lambda_0(\|f\|_\infty, \alpha) > 0$ such that $\mathcal{Z}_{\lambda f}$ is finite for all $0 < \lambda < \lambda_0$.*

Remark 7.3 To the best of our knowledge, if we take T to be the doubling map for $d = 1$ or the Arnold cat map for $d \geq 2$, then the results we stated in Theorem 7.1 and Corollary 7.2 are the first global results that do away with smallness or largeness assumptions for the coupling constant. In the large coupling regime, Herman's subharmonicity trick [28] can be applied (for trigonometric polynomials), and in the (perturbatively!) small coupling regime, the perturbative analysis of Chulaevsky–Spencer [18] and Sadel–Schulz–Baldes [35, 36] can be applied. Other methods get around changing the coupling constant by changing the base dynamics instead, specifically to increase its hyperbolicity; compare Bourgain–Bourgain–Chang [9] and Bjerklöv [6].

Remark 7.4 Taking the doubling map as an example, we give two sample computations. First, we show how to reduce a Hölder continuous cocycle on $\mathbb{R}/\mathbb{Z} \times \mathbb{R}^2$ to one on $\Omega \times \mathbb{R}^2$, where Ω is the full shift, which is in particular a subshift of finite type. Let $\Omega^+ = \{0, 1\}^{\mathbb{N}}$ and (Ω^+, T_+, μ^+) be the one-sided Bernoulli shift. Here we choose $\mu^+ = \tilde{\mu}^{\mathbb{N}}$ where $\tilde{\mu}(0) = \tilde{\mu}(1) = \frac{1}{2}$. Then it is well known that the map

$$\pi : \Omega^+ \rightarrow \mathbb{R}/\mathbb{Z}, \quad \omega^+ \mapsto \sum_{n=0}^{\infty} \frac{\omega_n^+}{2^{n+1}}$$

codes the dynamics of doubling map $(\mathbb{R}/\mathbb{Z}, T_2, m)$ to that of (Ω^+, T_+, μ^+) since $T_2 \circ \pi = \pi \circ T_+$ and $\pi_* \mu^+ = m$. In particular, for any cocycle map $A : \mathbb{R}/\mathbb{Z} \rightarrow \mathrm{SL}(2, \mathbb{R})$, we set $A^+ : \Omega^+ \rightarrow \mathrm{SL}(2, \mathbb{R})$ where $A^+ = A \circ \pi$, and we then have by construction $L(A, m) = L(A^+, \mu^+)$. Now we consider the full shift space (Ω, T, μ) whose one-sided shift is (Ω^+, T_+, μ^+) , as described above. By setting $\bar{A}(\omega) = A^+(\pi^+\omega)$, we clearly have $L(T, \bar{A}) = L(T_+, A^+)$. It is clear that $\bar{A}$ is α -Hölder continuous as long as A^+ is, since $d(\pi^+\omega, \pi^+\tilde{\omega}) \leq d(\omega, \tilde{\omega})$. So we just need to show that the Hölder continuity can be carried over from A to A^+ . This in turn follows from the following straightforward estimate:

$$|\pi \omega^+ - \pi \tilde{\omega}^+| \leq d(\omega^+, \tilde{\omega}^+)^{\log 2}.$$

In particular, α -Hölder continuity of A implies $(\alpha \log 2)$ -Hölder continuity of A^+ since

$$\begin{aligned} \|A^+(\omega^+) - A^+(\tilde{\omega}^+)\| &= \|A(\pi \omega^+) - A(\pi \tilde{\omega}^+)\| \\ &\leq C |\pi \omega^+ - \pi \tilde{\omega}^+|^\alpha \end{aligned}$$

$$\leq Cd(\omega^+, \tilde{\omega}^+)^{\alpha \log 2}.$$

Next, we compute some explicit choices for the value of λ_0 appearing in Theorem 7.1 and Corollary 7.2, when the base dynamics in question are given by the doubling map. Clearly, the above process still works if we replace $A : \mathbb{R}/\mathbb{Z} \rightarrow \mathrm{SL}(2, \mathbb{R})$ by $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}$. Given $f \in C^\alpha(\mathbb{R}/\mathbb{Z}, \mathbb{R})$, we may instead consider the corresponding $\tilde{f} \in C^{\alpha \log 2}(\Omega, \mathbb{R})$. In particular, $\|f\|_\infty = \|\tilde{f}\|_\infty$. We want to find a λ_0 so that f is globally bunched if $\|f\|_\infty < \lambda_0$. In other words,

$$A^E(\omega) = \begin{pmatrix} E - \tilde{f}(\omega) & -1 \\ 1 & 0 \end{pmatrix}$$

is fiber bunched for all $E \in [-2 - \|f\|_\infty, 2 + \|f\|_\infty]$. To simplify the computation, we ensure that fiber bunching is satisfied with $n_0 = 1$. That is, we want for all $E \in [-2 - \|\tilde{f}\|_\infty, 2 + \|\tilde{f}\|_\infty]$ that

$$\|A^E(\cdot)\|_\infty < e^{\frac{\log 2}{2}\alpha} = 2^{\frac{\alpha}{2}}.$$

Recall the fiber bunching condition is only assumed to ensure the existence of stable and unstable holonomies. Thus, by the construction of the holonomies from the proof of Lemma 4.2, it is clear that we may reduce the condition above to the following condition. For each $E \in [-2 - \|\tilde{f}\|_\infty, 2 + \|\tilde{f}\|_\infty]$, there is a $P(E) \in \mathrm{SL}(2, \mathbb{R})$ so that

$$\|P(E)^{-1}A^E(\cdot)P(E)\|_\infty < 2^{\frac{\alpha}{2}}. \quad (7.4)$$

First, we take care of the E 's that are away from ± 2 . For each $E \in (-2, 2)$, a direct computation shows that

$$P(E)^{-1} \begin{pmatrix} E & -1 \\ 1 & 0 \end{pmatrix} P(E) \in \mathrm{SO}(2, \mathbb{R})$$

which has norm one and where

$$P(E) = \begin{pmatrix} \frac{\sqrt{2}}{(4-E^2)^{\frac{1}{4}}} & 0 \\ \frac{E}{\sqrt{2}(4-E^2)^{\frac{1}{4}}} & \frac{(4-E^2)^{\frac{1}{4}}}{\sqrt{2}} \end{pmatrix}.$$

If we choose λ_0 so that for all $E \in [-2 + \lambda_0, 2 - \lambda_0]$ and all $|\lambda| < \lambda_0$, we have

$$\left\| P(E)^{-1} \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix} P(E) \right\| < 2^{\frac{\alpha}{2}} - 1,$$

then we have (7.4) for any $\|\tilde{f}\|_\infty = \|f\|_\infty < \lambda_0$ and all $E \in [-2 + \lambda_0, 2 - \lambda_0]$. It is straightforward to see that

$$P(E)^{-1} \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix} P(E) = \begin{pmatrix} \lambda & 0 \\ -\frac{E\lambda}{\sqrt{4-E^2}} & 0 \end{pmatrix}.$$

Thus we have fiber bunching for all $E \in [-2 + \lambda_0, 2 - \lambda_0]$ if for all such E 's and for all $|\lambda| < \lambda_0$, we have

$$|\lambda| + \left| \frac{E\lambda}{\sqrt{4-E^2}} \right| < 2^{\frac{\alpha}{2}} - 1.$$

Since the supremum of the left hand is attained at $\lambda = \lambda_0$ and $E = 2 - \lambda_0$, one can check that it suffices to have

$$\lambda_0 + \frac{\lambda_0}{\sqrt{\lambda_0 - \lambda_0^2}} < 2^{\frac{\alpha}{2}} - 1,$$

which in turn can be guaranteed, for example, by the condition $3\sqrt{\lambda_0} \leq 2^{\frac{\alpha}{2}} - 1$. In particular, if we choose any

$$0 < \lambda_0 \leq \frac{(2^{\frac{\alpha}{2}} - 1)^2}{9}, \quad (7.5)$$

then we have fiber bunching for all $E \in [-2 + \lambda_0, 2 - \lambda_0]$ and for all $\|f\|_\infty < \lambda_0$.

Now we take care of the energies $E \in [-2 - \lambda_0, -2 + \lambda_0] \cup [2 - \lambda_0, 2 + \lambda_0]$. Take $E = 2$ for example. Then we have

$$G_a^{-1} \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} G_a = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix},$$

where $a > 0$ and

$$G_a = \begin{pmatrix} \frac{1}{\sqrt{a}} & -\sqrt{a} \\ \frac{1}{\sqrt{a}} & 0 \end{pmatrix}.$$

It is easy to see that we have

$$\left\| \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \right\| \leq 1 + |a|.$$

On the other hand, we can see via a straightforward computation that

$$G_a^{-1} \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix} G_a = \begin{pmatrix} 0 & 0 \\ -\frac{\lambda}{a} & \lambda \end{pmatrix}.$$

Thus it suffices to choose $a > 0$ and $\lambda_0 > 0$ so that for all $|\lambda| \leq \lambda_0$, we have

$$1 + a + \frac{|\lambda|}{a} + |\lambda| < 2^{\frac{\alpha}{2}},$$

which may be guaranteed by

$$a + \frac{\lambda_0}{a} + \lambda_0 < 2^{\frac{\alpha}{2}} - 1.$$

Clearly, we may choose $a = \frac{1}{2}(2^{\frac{\alpha}{2}} - 1)$. It is then easy to see that if we choose any λ_0 such that

$$0 < \lambda_0 \leq \frac{(2^{\frac{\alpha}{2}} - 1)^2}{8}, \quad (7.6)$$

then we have fiber bunching for all $E \in [2 - \lambda_0, 2 + \lambda_0]$ and for all f with $\|f\|_\infty < \lambda_0$. A similar computation shows that the λ_0 in (7.6) works for $E \in [-2 - \lambda_0, -2 + \lambda_0]$ as well. Combining (7.5) and (7.6), we see that in the statement of Theorem 7.1 and Corollary 7.2 for the doubling map, we may choose

$$\lambda_0 = \frac{(2^{\frac{\alpha}{2}} - 1)^2}{9}.$$

Remark 7.5 The computation of λ_0 in Remark 7.4 actually works for A^E defined on any subshift of finite type (Ω, T, μ) . Moreover, since we do not have the coding process as in Remark 7.4, we have that $f \in C^\alpha(\Omega, \mathbb{R})$ is globally bunched if

$$\|f\|_\infty \leq \lambda_0 = \frac{(e^{\frac{\alpha}{2}} - 1)^2}{9}. \quad (7.7)$$

In particular, this value of λ_0 works for Theorem 7.6 below.

Let us now apply our results to Markov chains. We consider the full shift $(\mathcal{A}^{\mathbb{Z}}, T)$, where $\mathcal{A} = \{1, \dots, \ell\}$. Let $P = (P_{ij})_{1 \leq i, j \leq \ell}$ be a stochastic matrix, in other words, $P_{ij} \geq 0$ and $\sum_{j=1}^{\ell} P_{ij} = 1$. Assume that P is irreducible, that is, for all $i, j \in \mathcal{A}$, there is $n \in \mathbb{Z}_+$ such that the (i, j) -entry of P^n is positive.

Then there is a unique probability vector $\underline{p} = (p_1, \dots, p_\ell)$ (i.e., $p_i > 0$ and $\sum_{i=1}^\ell p_i = 1$) such that $\sum_{i=1}^\ell p_i P_{ij} = p_j$. Now we define the measure μ on $\mathcal{A}^{\mathbb{Z}}$ via

$$\mu([0; k_0, \dots, k_n]) = p_{k_0} \prod_{i=0}^{n-1} P_{k_i k_{i+1}}. \quad (7.8)$$

Such a measure μ is called a Markov measure. By a standard result, the topological support of μ is a subshift of finite type Ω with the adjacency matrix $A = (a_{ij})$ such that $a_{ij} = 1$ whenever $p_{ij} > 0$ and $a_{ij} = 0$ otherwise. Thus we may instead consider the space (Ω, T, μ) . Moreover, μ is T -ergodic if and only if P is irreducible. Consider its associated one-sided space (Ω^+, T_+, μ^+) . It is a standard result that μ^+ is the unique equilibrium state of the potential $\phi(\omega^+) = -\log P_{\omega_0^+ \omega_1^+}$, which is locally constant; see, for example, [41]. Thus by Lemma 3.4, μ has the bounded distortion property, and hence a local product structure as well.

Theorem 7.6 *Let (Ω, T, μ) be a Markov chain as described above. Fix $0 < \alpha \leq 1$. Then we have the following statements:*

- (a) *There is a residual set $\mathcal{G}^\alpha \subseteq C^\alpha(\Omega, \mathbb{R})$ such that $\mathcal{Z}_f = \emptyset$ for all $f \in \mathcal{G}^\alpha$.*
- (b) *There are $\lambda_0 = \lambda_0(\alpha) > 0$ and an open dense subset $\mathcal{O}^\alpha \subseteq C_{\lambda_0}^\alpha(\Omega, \mathbb{R})$ such that for each $f \in \mathcal{O}^\alpha$, we have $\inf_{E \in \mathbb{R}} L(E) > 0$.
If in addition (Ω, T) has a fixed point (which happens if and only if $P_{ii} > 0$ for some $1 \leq i \leq \ell$), the following stronger statements hold true:*
- (c) *$\mathcal{Z}_f$ is a discrete set for all non-constant $f \in C^\alpha(\Omega, \mathbb{R})$ and it is a finite set for all non-constant $f \in C_{\lambda_0}^\alpha(\Omega, \mathbb{R})$ or for all non-constant f that are locally constant.*
- (d) *In particular, $\mathcal{Z}_{\lambda f}$ is discrete for all $\lambda > 0$ and finite for all $0 < \lambda < \lambda_0$ for all non-constant $f \in C^\alpha(\Omega, \mathbb{R})$. If f is locally constant and non-constant, then $\mathcal{Z}_{\lambda f}$ is a finite set for all $\lambda > 0$.*

Remark 7.7 Reiterating what we said in Remark 5.13, even if (Ω, T) does not have a fixed point (i.e., when $P_{ii} = 0$ for every $1 \leq i \leq \ell$), we can work with periodic spectra of higher periods and test for non-coincidence of two of them. In concrete cases this procedure is easy to implement and will in many cases lead to the desired result. For instance, we can apply it to the last example we present in the end of this section.

Note that the Anderson model is a special case of the Markov chains described above, provided that the single-site measure is supported on a finite set. Indeed, such models may be generated as follows. Let μ be a probability measure on the full shift space $\mathcal{A}^{\mathbb{Z}}$ that is generated by a single site measure $\bar{\mu}\{i\} = p_i$ where $\underline{p} = (p_1, \dots, p_\ell)$ is a probability vector. It is clearly a

Markov chain with the same probability vector and with the stochastic matrix $p_{ij} = p_j$. Thus, we have the following corollary of Theorem 7.6.

Corollary 7.8 *Consider the full shift space $(\mathcal{A}^{\mathbb{Z}}, T, \mu)$, where $\mu = \tilde{\mu}^{\mathbb{Z}}$ and $\tilde{\mu}$ is a probability measure on $\mathcal{A} = \{1, \dots, \ell\}$ that has full support. Then all the conclusions that we stated in Theorem 7.6 hold true. In particular, if f is locally constant and non-constant, then $\mathcal{Z}_{\lambda f}$ is a finite set for all $\lambda > 0$.*

In particular, the Anderson model is generated by a sampling function $f : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathbb{R}$ that depends only the 0th position. Note that such a function is in particular locally constant. Corollary 7.8 implies the finiteness of $\mathcal{Z}_{\lambda f}$ for all such f 's that are non-constant. Of course, in this case, the celebrated Furstenberg's Theorem yields uniform positivity of the Lyapunov exponent. However, the finiteness of $\mathcal{Z}_f$ for all non-constant locally constant $f : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathbb{R}$ already may not be directly obtained from Furstenberg's Theorem. Moreover, our result is basically sharp. Indeed, there are plenty of examples where $\mathcal{Z}_f$ is not empty for locally constant and non-constant $f : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathbb{R}$, see [15]. Nevertheless, the finiteness of $\mathcal{Z}_f$ can already be a starting point to prove full spectral localization.

For the reader's convenience, we provide an example with the property $\mathcal{Z}_f \neq \emptyset$, where f is a non-constant locally constant function defined over a Markov chain. To give such an example, let us show that the well-known random dimer model (cf., e.g., [5, 26]) is covered by our framework. The random dimer model arises from the standard Bernoulli–Anderson model by doubling up the sites. That is, with $\{\omega_n\}_{n \in \mathbb{Z}}$ i.i.d. random variables taking two different values, say 0 and λ with probability $0 < p < 1$ and $1 - p$, the potentials are given by $V_{\omega}(2n) = V_{\omega}(2n + 1) = \omega_n$. To realize these potentials in our framework, consider the subshift of finite type Ω over the alphabet $\{1, 2, 3, 4\}$ with the adjacency matrix

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}. \quad (7.9)$$

The measure μ is the Markov measure generated by the following probability vector and the stochastic matrix

$$\underline{p} = \left(\frac{p}{2}, \frac{1-p}{2}, \frac{p}{2}, \frac{1-p}{2} \right), \quad P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ p & 1-p & 0 & 0 \\ p & 1-p & 0 & 0 \end{pmatrix}. \quad (7.10)$$

The sampling function $f : \Omega \rightarrow \mathbb{R}$ is generated by $\bar{f} : \{1, 2, 3, 4\} \mapsto \{0, \lambda\}$, $\bar{f}(1) = \bar{f}(3) = 0$, $\bar{f}(2) = \bar{f}(4) = \lambda$ via $f(\omega) = \bar{f}(\omega_0)$ which is locally constant. It is readily checked that the resulting model is indeed the random dimer model. It is well known, and in fact easy to see, that for $-2 < \lambda < 2$, $A_n^{(E-f)}(\omega)$ is bounded for all n at energies 0 and λ . Thus $\{0, \lambda\} \subseteq \mathcal{Z}_f$. Although this system has no fixed point, we do have that f is constant on the orbit of $\omega \in \Omega$ where $\omega_{2n} = 1, \omega_{2n+1} = 3$. Note that in statement of Theorems 1.1 and 5.12, the fixed point is only there to produce a constant potential $V_\omega(n)$. Thus, Theorem 7.6 can still be applied to obtain the finiteness of $\mathcal{Z}_f$. However, for this model, we can provide more information. It actually follows from Furstenberg's Theorem that the Lyapunov exponent is positive away from these two energies $\{0, \lambda\}$. This shows that in this particular case $\mathcal{Z}_f = \{0, \lambda\}$.

Acknowledgements We are grateful to the anonymous referee for the careful reading of the manuscript and the numerous suggestions of how to improve the presentation and explanation.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Avila, A., Damanik, D., Zhang, Z.: Schrödinger operators with potentials generated by hyperbolic transformations: II. Large deviations and localization (in preparation)
2. Avila, A., Delecroix, V.: Weak mixing directions in non-arithmetic Veech surfaces. *J. Am. Math. Soc.* **29**, 1167–1208 (2016)
3. Avila, A., Viana, M.: Extremal Lyapunov exponents: an invariance principle and applications. *Invent. Math.* **181**, 115–189 (2010)
4. Backes, L., Brown, A., Butler, C.: Continuity of Lyapunov exponents for cocycles with invariant holonomies. *J. Mod. Dyn.* **12**, 223–260 (2018)
5. de Bièvre, S., Germinet, F.: Dynamical localization for the random dimer Schrödinger operator. *J. Stat. Phys.* **98**, 1135–1148 (2000)
6. Bjerklov, K.: Positive Lyapunov exponent for some Schrödinger cocycles over strongly expanding circle endomorphisms. *Commun. Math. Phys.* **379**, 353–360 (2020)
7. Bonatti, C., Gómez-Mont, X., Viana, M.: Généricité d'exposants de Lyapunov non-nuls pour des produits déterministes de matrices. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **20**, 579–624 (2003)
8. Bonatti, C., Viana, M.: Lyapunov exponents with multiplicity 1 for deterministic products of matrices. *Ergod. Theory Dyn. Syst.* **24**, 1295–1330 (2004)
9. Bourgain, J., Bourgain-Chang, E.: A note on Lyapunov exponents of deterministic strongly mixing potentials. *J. Spectr. Theory* **5**, 1–15 (2015)

10. Bourgain, J., Goldstein, M.: On nonperturbative localization with quasi-periodic potential. *Ann. Math.* **152**, 835–879 (2000)
11. Bourgain, J., Goldstein, M., Schlag, W.: Anderson localization for Schrödinger operators on $\mathbb{Z}$ with potentials given by the skew-shift. *Commun. Math. Phys.* **220**, 583–621 (2001)
12. Bourgain, J., Schlag, W.: Anderson localization for Schrödinger operators on $\mathbb{Z}$ with strongly mixing potentials. *Commun. Math. Phys.* **215**, 143–175 (2000)
13. Bowen, R.: Periodic points and measures for Axiom A diffeomorphisms. *Trans. Am. Math. Soc.* **154**, 377–397 (1971)
14. Bowen, R.: *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. Lecture Notes in Mathematics*, vol. 470. Springer, Berlin (1975)
15. Bucaj, V.: On the Kunz–Souillard approach to localization for the discrete one dimensional generalized Anderson model. Preprint [arXiv:1608.01379](https://arxiv.org/abs/1608.01379)
16. 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. Am. Math. Soc.* **372**, 3619–3667 (2019)
17. Butler, C.: Measurable rigidity of the cohomological equation for linear cocycles over hyperbolic systems. *Isr. J. Math.* **227**, 27–61 (2018)
18. Chulaevsky, V., Spencer, T.: Positive Lyapunov exponents for a class of deterministic potentials. *Commun. Math. Phys.* **168**, 455–466 (1995)
19. Damanik, D.: Schrödinger operators with dynamically defined potentials. *Ergod. Theory Dyn. Syst.* **37**, 1681–1764 (2017)
20. Damanik, D., Fillman, J.: One-dimensional ergodic Schrödinger operators, I. General theory, Monograph (in preparation)
21. Damanik, D., Fillman, J.: One-dimensional ergodic Schrödinger operators, II. Special classes, Monograph (in preparation)
22. Damanik, D., Killip, R.: Almost everywhere positivity of the Lyapunov exponent for the doubling map. *Commun. Math. Phys.* **257**, 287–290 (2005)
23. Damanik, D., Tcheremchantsev, S.: Power-law bounds on transfer matrices and quantum dynamics in one dimension. *Commun. Math. Phys.* **236**, 513–534 (2003)
24. Deift, P., Simon, B.: Almost periodic Schrödinger operators. III. The absolutely continuous spectrum in one dimension. *Commun. Math. Phys.* **90**, 389–411 (1983)
25. Douady, A., Earle, C.J.: Conformally natural extension of homeomorphisms of the circle. *Acta Math.* **157**, 23–48 (1986)
26. Dunlap, D., Wu, H.-L., Phillips, P.: Absence of localization in a random dimer model. *Phys. Rev. Lett.* **65**, 88–91 (1990)
27. Furstenberg, H.: Noncommuting random products. *Trans. Am. Math. Soc.* **108**, 377–428 (1963)
28. Herman, M.: Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser sur le tore de dimension 2. *Comment. Math. Helv.* **58**, 453–502 (1983)
29. Jitomirskaya, S., Schulz-Baldes, H., Stolz, G.: Delocalization in random polymer models. *Commun. Math. Phys.* **233**, 27–48 (2003)
30. Kalinin, B.: Livšic theorem for matrix cocycles. *Ann. Math.* **173**, 1025–1042 (2011)
31. Katok, A., Hasselblatt, B.: *Introduction to the Modern Theory of Dynamical Systems. Encyclopedia of Mathematics and its Applications*, vol. 54. Cambridge University Press, Cambridge (1995)
32. Pastur, L.: Spectral properties of disordered systems in the one-body approximation. *Commun. Math. Phys.* **75**, 179–196 (1980)
33. Ledrappier, F.: Positivity of the exponent for stationary sequences of matrices. *Lect. Notes Math.* **1186**, 56–73 (1986)
34. Leplaideur, R.: Local product structure for equilibrium states. *Trans. Am. Math. Soc.* **352**, 1889–1912 (2000)

35. Sadel, C., Schulz-Baldes, H.: Scaling diagram for the localization length at a band edge. *Annales Henri Poincaré* **8**, 1595–1621 (2007)
36. Sadel, C., Schulz-Baldes, H.: Positive Lyapunov exponents and localization bounds for strongly mixing potentials. *Adv. Theor. Math. Phys.* **12**, 1377–1400 (2008)
37. Sigmund, K.: On dynamical systems with the specification property. *Trans. Am. Math. Soc.* **190**, 285–299 (1974)
38. Simon, B.: Szegő's Theorem and its Descendants. *Spectral Theory for L^2 Perturbations of Orthogonal Polynomials*, M. B. Porter Lectures, Princeton University Press, Princeton, NJ, 2011
39. Viana, M.: Almost all cocycles over any hyperbolic system have non-vanishing Lyapunov exponents. *Ann. Math.* **167**, 643–680 (2008)
40. Viana, M., Yang, J.: Continuity of Lyapunov exponents in the C^0 topology. *Isr. J. Math.* **229**, 461–485 (2019)
41. Young, L.-S.: Large deviations in dynamical systems. *Trans. Am. Math. Soc.* **318**, 525–543 (1990)
42. Young, L.-S.: Some open sets of nonuniformly hyperbolic cocycles. *Ergod. Theory Dyn. Syst.* **13**, 409–415 (1993)
43. Zhang, Z.: Uniform hyperbolicity and its relation with spectral analysis of 1D discrete Schrödinger operators. *J. Spectr. Theory* **10**, 1471–1517 (2020)
44. Zhang, Z.: Uniform positivity of the Lyapunov exponent for monotone potentials generated by the doubling map. Preprint [arXiv:1610.02137](https://arxiv.org/abs/1610.02137)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.