Ergod. Th. & Dynam. Sys. (2020), **40**, 2788–2853
doi:10.1017/etds.2019.19

Effective multi-scale approach to the Schrödinger cocycle over a skew-shift base

R. HAN†, M. LEMM‡ and W. SCHLAG§

† *School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA*
(e-mail: rui.han@math.gatech.edu)
‡ *Harvard University, Department of Mathematics, 1 Oxford Street, Cambridge, MA 02138, USA*
(e-mail: mlemm@math.harvard.edu)
§ *Yale University, Department of Mathematics, 10 Hillhouse Ave, New Haven, CT 06511, USA*
(e-mail: willsg69@gmail.com)

(*Received* 13 *April* 2018 *and accepted in revised form* 3 *January* 2019)

Abstract. We prove a conditional theorem on the positivity of the Lyapunov exponent for a Schrödinger cocycle over a skew-shift base with a cosine potential and the golden ratio as frequency. For coupling below 1, which is the threshold for Herman's subharmonicity trick, we formulate three conditions on the Lyapunov exponent in a finite but large volume and on the associated large-deviation estimates at that scale. Our main results demonstrate that these finite-size conditions imply the positivity of the infinite-volume Lyapunov exponent. This paper shows that it is possible to make the techniques developed for the study of Schrödinger operators with deterministic potentials, based on large-deviation estimates and the avalanche principle, effective.

Key words: Hamiltonian dynamics, random dynamics, Lyapunov exponent, Skew-shift
2010 Mathematics Subject Classification: 47B36, 35J10 (Primary)

1. *Introduction*

The study of Lyapunov exponents occupies a central role in ergodic theory and dynamical systems. They arise in a multitude of distinct settings, such as diffeomorphisms on a manifold, chaotic dynamics in nonlinear systems as exhibited by the standard map, cocycles defined over some base, and the theory of localization. Perhaps the most fundamental question about Lyapunov exponents relates to their simplicity; or, more quantitatively, to the gaps between them. In the case of $\mathrm{SL}_2(\mathbb{R})$ cocycles this amounts to the question of positivity of the top Lyapunov exponent. Another much studied property of these exponents concerns their continuity relative to external parameters. For a beautiful introduction to this field see the textbook [**Via**].

This paper studies Schrödinger cocycles

$$(x, v) \in X \times \mathbb{R}^2 \mapsto (Tx, A_\lambda(x, E)v)$$
$$A_\lambda(x, E) = \begin{bmatrix} \lambda f(x) - E & -1 \\ 1 & 0 \end{bmatrix} \in \mathrm{SL}_2(\mathbb{R})$$

where (X, μ, T) is some ergodic system, $\lambda, E \in \mathbb{R}$ and $f : X \to \mathbb{R}$ is measurable. These cocycles arise in the spectral analysis of the operators

$$(H_{\lambda,x}\psi)_n = \psi_{n+1} + \psi_{n-1} + \lambda f(T^n x)\psi_n, \quad n \in \mathbb{Z}.$$

Indeed, solutions of $H_{\lambda,x}\psi = E\psi$ are given by

$$\begin{pmatrix} \psi_{n+1} \\ \psi_n \end{pmatrix} = M_n(x; \lambda, E)\begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix}, \tag{1.1}$$
$$M_n(x; \lambda, E) = \prod_{j=n}^{1} A_\lambda(T^j x, E), \quad n \geq 1.$$

The growth of solutions to (1.1) μ-almost everywhere in x is governed by the Lyapunov exponent

$$L(\lambda, E) = \lim_{n\to\infty} n^{-1} \int_X \log \|M_n(x; \lambda, E)\| \, \mu(dx)$$

which always exists by subadditivity. By unimodularity of the matrices, $L(\lambda, E) \geq 0$. The main issue is then to determine strict positivity. We remark that by classical ergodic theory (Fürstenberg–Kesten theorem, Kingman's subadditive ergodic theorem [**Via**]),

$$n^{-1} \log \|M_n(x; \lambda, E)\| \to L(\lambda, E) \quad \mu\text{-almost surely}$$

as $n \to \infty$. Fürstenberg's theorem [**Fur**], shows that $L > 0$ for all λ, E for T the Bernoulli shift and μ a non-trivial probability distribution. Herman's subharmonicity argument [**Her**], which is recalled in §6, shows that $L(\lambda, E) \geq \log \lambda > 0$ if $\lambda > 1$, $X = \mathbb{T}$, $f(x) = 2\cos(2\pi x)$, and $Tx = x + \omega$ a rotation (for general analytic f and large λ; see [**SorSpe**]). On the other hand, one has $L(\lambda, E) = 0$ for all $0 < \lambda < 1$ and $E \in \mathrm{spec}(H_{\lambda,x})$. The latter is the spectrum of the Harper or almost Mathieu operator

$$(H_{\lambda,x}\psi)_n = \psi_{n+1} + \psi_{n-1} + 2\lambda\cos(2\pi(x + n\omega))\psi_n, \quad n \in \mathbb{Z},$$

which does not depend on x (assuming ω irrational). In particular, $L(\lambda, 0) = 0$; cf. [**BelSim**, **Dam**].

In contrast to the Harper operator, its analog over the skew-shift base is conjectured to exhibit positive Lyapunov exponents for all $\lambda > 0$ and E. To be specific, let $X = \mathbb{T}^2$, $T(x, y) = (x + y, y + \omega)$, where ω is irrational (or Diophantine). Iterating T yields

$$M_n(x, y; \lambda, E) = \prod_{j=n}^{1} \begin{bmatrix} 2\lambda f(x + jy + j(j-1)\omega/2, y + j\omega) - E & -1 \\ 1 & 0 \end{bmatrix}. \tag{1.2}$$

The presence of $j^2\omega/2$ in these matrices appears to be the origin of the conjectured exponential growth of the norm of these matrices for all E (assuming $\partial_x f(x, y) \not\equiv 0$, with f analytic). In fact, the distribution of the fractional parts of $\{j^2\omega\}_{j=1}^N$ is known to be 'random' in some sense as $N \to \infty$ for generic ω; see the Poissonian conjecture in [**RudSarZah**], as well as [**MarStr**, **Hea**]. Note that this is in stark contrast to the distribution of $\{j\omega\}_{j=1}^N$.

However, not only is this randomness property in and of itself delicate (see some negative results to this effect in [**RudSarZah**]), but also how to use it in the context of (1.2) is entirely unclear. As far as rigorous results are concerned, Bourgain [**Bou2**] proved that for all $\lambda > 0$ there exists a set of $\omega \in \mathbb{T}$ with positive measure (which decreases to 0 as $\lambda \to 0$), so that the operator

$$(H\psi)_n = \psi_{n+1} + \psi_{n-1} + \lambda \cos(2\pi n(n-1)\omega/2)\psi_n$$

exhibits point spectrum whose closure has positive measure. This was the first result of its kind which showed that for small λ the skew shift leads to completely different behavior than the shift, that is, potentials $\cos(n\omega)$. Bourgain [**Bou3**] also showed that for small $\lambda > 0$, and most energies, the Lyapunov exponent is positive if $0 < \omega < \omega_0(\lambda)$ (which decreases to 0 as $\lambda \to 0$). A quantitative version of Bourgain's result was obtained later by Krüger in [**Kru2**]. These two results are proved by viewing the skew-shift model as a perturbation of Harper's model, hence require the smallness of ω. For any irrational ω, Krüger [**Kru1**] proved positivity of the Lyapunov exponent for skew shift on $\mathbb{T}^d$ with d sufficiently large, for small $\lambda > 0$, and most energies.

In this paper we present an effective multi-scale machinery aiming at positivity of the Lyapunov exponent for the matrices

$$M_n(x, y; \lambda, E) = \prod_{j=n}^{1} \begin{bmatrix} 2\lambda \cos(2\pi(x + jy + j(j-1)\omega/2)) - E & -1 \\ 1 & 0 \end{bmatrix} \tag{1.3}$$

uniformly in E, and in the range $0 < \lambda \le 1$. We fix ω to be the golden ratio. By the aforementioned estimate by Herman, one has $L(\lambda, E) \ge \log \lambda > 0$ for $\lambda > 1$. So only $\lambda \le 1$ is of interest here. The basis of our analysis is the inductive argument from [**BouGolSch**], which established Anderson localization for large λ for the skew-shift model, at the expense of removing a small set (in measure) of frequencies ω and phases (x, y) (the largeness of λ depended on the smallness of the measure of excluded parameters). The proof in [**BouGolSch**] is not effective, and it was not possible to explicitly determine the size of admissible λ in relation to the other parameters.

To formulate our main results, recall the finite-volume Lyapunov exponents

$$L_N(\lambda, E) := \int_{\mathbb{T}^2} \frac{1}{N} \log \|M_N(x, y; \lambda, E)\| \, dx\, dy,$$

and their limits $L = \lim_{N\to\infty} L_N$. We quantify the failure of the Fürstenberg–Kesten theorem via the level sets

$$\mathcal{B}_N := \left\{(x, y) \in \mathbb{T}^2 : \left|\frac{1}{N} \log \|M_N(x, y; \lambda, E)\| - L_N(\lambda, E)\right| > \frac{1}{10} L_N(\lambda, E)\right\}.$$

The machinery developed in this paper establishes a method for checking the positivity of the Lyapunov exponent $L(\lambda, E)$ by verifying information on a finite, initial scale.

We could have formulated a very general 'finite-size criterion' which establishes $L(\lambda, E) > 0$ under appropriate assumptions on the initial scale and for appropriate values of various other parameters. Instead, we have opted to present three representative theorems that can be obtained from the machinery developed in this paper by making specific choices.

These representative theorems differ by the precise assumptions (i)–(iii) made at the initial scale. We comment on this further after the first theorem. Moreover, the various other parameters appearing in our proof are identical in all three cases. These parameters are only chosen in the final part of the proof, §9, so they can easily be modified.

For a Borel set, $|\cdot|$ denotes the Lebesgue measure.

THEOREM 1.1. *Consider the skew-shift cocycle given by (1.3). Let ω be the golden ratio and let $\lambda \in [1/2, 1]$. Let $N_0 := 2 \times 10^{37}$. Assume that for some energy $E \in [-2-2\lambda, 2+2\lambda]$ the following inequalities hold:*

(i) $L_{N_0}(\lambda, E) \geq 2 \times 10^{-4}$;

(ii) $L_{N_0}(\lambda, E) - L_{2N_0}(\lambda, E) \leq L_{N_0}(\lambda, E)/8$;

(iii) $\max(|\mathcal{B}_{N_0}|, |\mathcal{B}_{2N_0}|) \leq N_0^{-21}$.

Then we have

$$L(\lambda, E) \geq \tfrac{1}{2} L_{N_0}(\lambda, E) > 0.$$

Before we give the two alternative theorems, we comment on conditions (i)–(iii).

Remark 1.2.

(i) First, one might expect that $L(\lambda, E) > c\lambda^2$ holds for small λ, by analogy with the Figotin–Pastur asymptotics in the random case. Numerical experimentation suggests that is indeed the case for our model with $c > 10^{-2}$ (with a generous margin of error). Therefore, we would expect to have $L_{N_0}(\lambda, E) \geq 2 \times 10^{-3}$. Condition (i) was chosen to allow for an even wider margin. We remark that we can lower the number 2×10^{-4} to basically any positive constant, at the expense of increasing N_0.

(ii) Condition (ii) is known to hold if the Lyapunov exponent is positive and N_0 is large enough. Indeed, it follows from the methods in [**GolSch**] that

$$L_{N_0}(\lambda, E) - L_{2N_0}(\lambda, E) \leq c L_{N_0}(\lambda, E)/N_0$$

with some absolute constant $c \sim 1$; see §8 below for the details. Given the size of N_0, condition (ii) is indeed asking for very little.

(iii) Finally, condition (iii) is some weak form of a *large-deviation estimate* as in [**BouGol**, **GolSch**, **BouGolSch**]. In fact, analogy with these references suggests that a bound of the form $|\mathcal{B}_N| < \exp(-N^{1/10})$ should hold for large N (and perhaps a much stronger bound, say with $N^{1/2}$ or larger). For (iii) to hold in this case would then require $N > 8 \cdot 10^{31}$, which is within our range. It is important to note that condition (iii) differs strongly from (i) and (ii). Indeed, while the latter conditions are intimately related to the $L > 0$, (iii) is not. For the rotation with Diophantine frequency dynamics it is known that the large-deviation estimates hold *a priori*, that is, without any reference to the positivity of the Lyapunov exponent; see [**BouGol**, **GolSch**, **Bou1**]. For the skew shift, as well as for the Bernoulli shift,

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

however, such *a priori* derivations are currently not known. Rather, we rely on an inductive procedure that uses lower bounds on L_n, the Lyapunov exponents in finite volume.

We now state two further representative theorems. These alternative finite-size criteria both involve much smaller initial scales N_0, at the price of having a more restrictive assumption (iii) on the measure of the set $\mathcal{B}_{N_0}$.

THEOREM 1.3. *Consider the skew-shift cocycle given by (1.3). Let ω be the golden ratio and let $\lambda \in [1/2, 1]$. Let $N_0 := 3 \times 10^5$. Assume that for some energy $E \in [-2-2\lambda, 2+2\lambda]$ the following inequalities hold:*

(i) $L_{N_0}(\lambda, E) \geq 2 \times 10^{-4}$;

(ii) $L_{N_0}(\lambda, E) - L_{2N_0}(\lambda, E) \leq L_{N_0}(\lambda, E)/8$;

(iii) $\max(|\mathcal{B}_{N_0}|, |\mathcal{B}_{2N_0}|) \leq N_0^{-141}$.

Then we have

$$L(\lambda, E) \geq \tfrac{1}{2} L_{N_0}(\lambda, E) > 0.$$

The upper bound in assumption (iii) is more restrictive than in Theorem 1.3. Importantly, it is still polynomial in nature. Hence, in view of Remark 1.2(iii), it may hold depending on the precise kind of exponential decay that is presumably exhibited by the true $|\mathcal{B}_N|$.

In the next representative result, we strengthen assumption (i) somewhat (in a way that is compatible with the numerics described in Remark 1.2(i) above). This allows us to reduce the initial scale even further, to the value $N_0 = 3 \times 10^4$, which may be amenable to numerical investigation.

THEOREM 1.4. *Consider the skew-shift cocycle given by (1.3). Let ω be the golden ratio and let $\lambda \in [1/2, 1]$. Let $N_0 := 3 \times 10^4$. Assume that for some energy $E \in [-2-2\lambda, 2+2\lambda]$ the following inequalities hold:*

(i) $L_{N_0}(\lambda, E) \geq 2 \times 10^{-3}$;

(ii) $L_{N_0}(\lambda, E) - L_{2N_0}(\lambda, E) \leq L_{N_0}(\lambda, E)/8$;

(iii) $\max(|\mathcal{B}_{N_0}|, |\mathcal{B}_{2N_0}|) \leq N_0^{-165}$.

Then we have

$$L(\lambda, E) \geq \tfrac{1}{2} L_{N_0}(\lambda, E) > 0.$$

Regarding assumption (iii), the comment made after Theorem 1.3 still applies. In particular, the relatively small value of N_0 in this result demonstrates the importance of the problem of finding an *analytical proof* of (iii). Indeed, (i) and (ii) are accessible numerically by a Figotin–Pastur expansion, for example, but it seems completely unreasonable to ask for a computer-assisted proof of (iii).

The restriction $\lambda \in [1/2, 1]$ was chosen for convenience. In fact, our methods apply to any given interval of the form $[\lambda_0, 1]$, $\lambda_0 > 0$, albeit with increasing N_0 as $\lambda_0 \to 0$. Similarly, the golden ratio was chosen for simplicity. One can replace it by a class of Diophantine frequencies obeying an explicit Diophantine condition.

It remains to be seen what the true range of applicability of our methods is, and to what extent they can also be refined. It may be possible to verify assumptions (i) and (ii) of Theorem 1.4 numerically. However, the measure estimates (iii) would seem the most delicate to check reliably.

The methods in this paper are an adaptation of those in [**GolSch**, **BouGolSch**, **Bou1**]. One of our motivations was to obtain an effective rendition of the techniques based on harmonic analysis (subharmonic functions, Riesz representation theorem, John–Nirenberg type estimates for bounded mean oscillation functions) in combination with linear algebra and the geometry of matrix products ('avalanche principle' [**GolSch**]). This had never been attempted before, but we show that it is possible to do so.

2. *Effective Riesz representation*

It is of fundamental importance to the entire method to make the underlying potential theory effective. To this end it is most convenient to remain on the disk since other geometries will lead to complicated Green functions. The disk will suffice for our purposes, thanks to a variant of Herman's regularization [**Her**], which we present in §6.

Definition 2.1. Given $R > 0$, we write D_R for the open disk of radius R around the origin in $\mathbb{C}$. Let $z = re(\phi)$. We write $P_z(\theta)$ for the Poisson kernel

$$P_z(\theta) := \frac{1 - r^2}{1 - 2r\cos(2\pi(\phi - \theta)) + r^2}.$$

The following constants will be used throughout, with $1 < R_2 < R_1 < R$:

$$B_0(R, R_1, R_2) := \frac{1}{2\log(R/R_1)}\left(\frac{R_1 + R_2}{R_1 - R_2}\right) \times \begin{cases} \log R & \text{if } R^2 - R_1^2 > R, \\ \log\left(\dfrac{R^2}{R^2 - R_1^2}\right) & \text{if } R^2 - R_1^2 < R. \end{cases} \tag{2.1}$$

$$B_1(R, R_1, R_2) := B_0(R, R_1, R_2)\frac{8R_2}{R_2^2 - 1},$$

$$B_2(R, R_1, R_2) := B_0(R, R_1, R_2) \times \frac{16\pi(R_2^2 - 1)\sqrt{16R_2^2 - \left(\sqrt{R_2^4 + 34R_2^2 + 1} - 1 - R_2^2\right)^2}}{(3R_2^2 + 3 - \sqrt{R_2^4 + 34R_2^2 + 1})^2}$$

as well as

$$B_3(R, R_1, R_2) := \sqrt{5B_2(R, R_1, R_2) + \frac{10\pi}{\log R/R_1}}. \tag{2.2}$$

The main result of this section is the following Riesz representation theorem for subharmonic functions. The essential feature here are the explicitly computable constants. Recall that a subharmonic function in some domain $\Omega \subset \mathbb{C}$ is an upper semicontinuous function $u : \Omega \to \mathbb{R} \cup \{-\infty\}$ which satisfies the sub-mean value property in Ω.

THEOREM 2.2. *Let $1 < R_2 < R_1 < R$ and let $v : \overline{D_R} \to \mathbb{R} \cup \{-\infty\}$ be a subharmonic function satisfying*

$$v(z) \le B, \quad v(0) = m. \tag{2.3}$$

Then, for all $w \in D_{R_1}$, we have the Riesz representation

$$v(w) = \int_{D_{R_1}} \log|z - w|\mu(dz) + h(w), \tag{2.4}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

where:

(i) *μ is a positive measure satisfying the bound*

$$\mu(D_{R_1}) \leq \frac{B-m}{\log(R/R_1)}; \tag{2.5}$$

(ii) *h is harmonic on D_{R_1} and satisfies the following bounds*

$$\min_{c\in\mathbb{R}} \max_{|w|\leq R_2} |h(w)-c| \leq B_0(R, R_1, R_2)(B-m),$$

$$\left|\frac{d^k}{d\varphi^k} h(e(\varphi))\right| \leq B_k(R, R_1, R_2)(B-m), \quad k = 1, 2. \tag{2.6}$$

The proof of this theorem will occupy the rest of this section.

Proof of Theorem 2.2. The basic idea of the proof is that the equation $\mu = (1/2\pi)\Delta v$ holds in the distributional sense, with μ a positive measure. Without loss of generality, we may assume that v is smooth. If this is not the case, we convolve v with a radial non-negative mollifier. The sub-mean property then guarantees monotone convergence. We skip this technical detail, and commence with some basic potential theory for smooth functions.

2.1. *Riesz representation.* Rescaling the unit disk yields the Green function on any disk.

LEMMA 2.3. (Green's function for the disk) *The function $G : D_R \times D_R \to \mathbb{R}$ given by*

$$G(z, w) := \frac{1}{2\pi} \log\left|\frac{R(z-w)}{R^2 - z\overline{w}}\right|$$

satisfies $\Delta_z G(z, w) = \delta_w$ and $G(z, w) = 0$ when $|z| = R$.

Proof. To see this, notice that $G(z, w) = G_1(z/R, w/R)$ where

$$G_1(z, w) := \frac{1}{2\pi} \log\left|\frac{z-w}{1-z\overline{w}}\right|$$

is the Green function of the unit disk. □

Let $w \in D_{R_1}$. By Green's second identity for the domain D_R, we have

$$v(w) - \int_{D_R} G(z, w)\Delta v(z) \operatorname{Vol}(dz) = \int_{\partial D_R} v(z) \frac{\partial G}{\partial n_z}(z, w)\sigma(dz),$$

where Vol is the standard volume measure and σ is the (unnormalized) arclength measure on the circle ∂D_R. Since v is smooth and subharmonic, Δv is a non-negative, continuous function; call it $2\pi\mu$. Therefore

$$v(w) = \int_{D_R} 2\pi G(z, w)\mu(dz) + h_0(w), \tag{2.7}$$

where

$$h_0(w) := \int_{\partial D_R} v(z) \frac{\partial G}{\partial n_z}(z, w)\sigma(dz). \tag{2.8}$$

By Lemma 2.3, we then have Riesz representation with the functions

$$v(w) = \int_{D_{R_1}} \log|z-w|\mu(dz) + h(w), \tag{2.9}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

where

$$h(w) := \int_{D_R\setminus D_{R_1}} \log \left|\frac{R(z-w)}{R^2 - z\overline{w}}\right| \mu(dz) + \int_{D_{R_1}} \log\left|\frac{R}{R^2 - z\overline{w}}\right| \mu(dz) + h_0(w). \quad (2.10)$$

LEMMA 2.4. *$h(w)$ is harmonic in D_{R_1}.*

Proof. Write $w = re^{2\pi i\varphi}$. The first and second terms in (2.10) are harmonic because they are real parts of analytic functions on D_{R_1}. For the third term, recall that $(\partial G/\partial n_z)(z, w)$ is the Poisson kernel, whence

$$h_0(w) = \int_0^1 v(Re(\theta))P_{r/R}(\varphi - \theta)\, d\theta. \quad (2.11)$$

The Poisson kernel is harmonic in all of D_R and this proves the lemma. □

2.2. *Control of the Riesz mass.*

LEMMA 2.5. *We have the following bound on the Riesz mass:*

$$\mu(D_{R_1}) \le \frac{B - m}{\log(R/R_1)}. \quad (2.12)$$

Proof. Taking $w = 0$ in (2.7), we see that

$$(\log R/R_1)\mu(D_{R_1}) \le \int_{D_R} \log \frac{R}{|z|}\mu(dz) = h_0(0) - v(0) \le B - m,$$

in which we used

$$h_0(0) \le B, \quad (2.13)$$

which comes from the maximum principle and the fact that $h_0(w)$ is the harmonic function on D_R with boundary values $v(\partial D_R)$ by (2.11). □

2.3. *Control of the harmonic part.*

We have the following estimate for the harmonic part.

LEMMA 2.6. *Let $1 < R_2 < R_1 < R$. Then*

$$\min_{c\in\mathbb{R}} \max_{|w|\le R_2} |h(w) - c| \le B_0(R, R_1, R_2)(B - m),$$

with constant $B_0(R, R_1, R_2)$ given by (2.1).

Proof. We will first prove an upper bound and then use Harnack's inequality to conclude a lower bound. From (2.10), and $G(z, w) \le 0$ on $D_R \times D_R$,

$$h(w) \le \int_{D_{R_1}} \log \left|\frac{R}{R^2 - z\overline{w}}\right| \mu(dz) + h_0(w).$$

From (2.13), we infer that for all $w \in D_{R_1}$,

$$h(w) \le \log \left|\frac{R}{R^2 - R_1^2}\right| \mu(D_{R_1}) + B.$$

Now we distinguish cases. On the one hand, if $R < R^2 - R_1^2$, then the logarithm is negative and (2.14) implies

$$h(w) \le B.$$

On the other hand, if $R > R^2 - R_1^2$, then we use Lemma 2.5 to obtain

$$\begin{aligned} h(w) &\leq \frac{1}{\log R/R_1}\left(B\log\frac{R^2}{R_1(R^2-R_1^2)} - m\log\frac{R}{R^2-R_1^2}\right) \\ &= \frac{m\log R - B\log R_1}{\log R/R_1} + \frac{\log R^2/(R^2-R_1^2)}{\log R/R_1}(B-m). \end{aligned}$$

Combining the two cases, we arrive at the upper bound

$$h(w) \leq \alpha, \quad \alpha := \begin{cases} B & \text{if } R^2 - R_1^2 > R, \\ \dfrac{m\log R - B\log R_1}{\log R/R_1} & \\ \quad + \dfrac{\log R^2/(R^2-R_1^2)}{\log R/R_1}(B-m) & \text{if } R^2 - R_1^2 < R. \end{cases} \tag{2.14}$$

Consider the non-negative harmonic function $\alpha - h(w)$ on D_{R_1}. By Harnack's inequality,

$$\alpha - h(w) \leq \frac{R_1 + |w|}{R_1 - |w|}(\alpha - h(0)),$$

which implies the lower bound

$$h(w) \geq \frac{R_1 + |w|}{R_1 - |w|}h(0) - \frac{2|w|}{R_1 - |w|}\alpha. \tag{2.15}$$

By (2.9) with $w = 0$ and (2.5), we have

$$h(0) = v(0) - \int_{D_{R_1}} \log|z|\mu(dz) \geq m - \frac{\log R_1}{\log R/R_1}(B-m).$$

Together with (2.15), this yields

$$h(w) \geq \frac{R_1 + |w|}{R_1 - |w|}\frac{m\log R - B\log R_1}{\log R/R_1} - \frac{2|w|}{R_1 - |w|}\alpha. \tag{2.16}$$

Based on (2.14) and (2.16), we obtain

$$\begin{aligned} &\min_{c\in\mathbb{R}}\max_{|w|\leq R_2}|h(w) - c| \\ &\leq \frac{1}{2}\left(\alpha - \min_{|w|\leq R_2}\left(\frac{R_1+|w|}{R_1-|w|}\frac{m\log R - B\log R_1}{\log R/R_1} - \frac{2|w|}{R_1-|w|}\alpha\right)\right) \\ &= \frac{1}{2}\max_{|w|\leq R_2}\left(\frac{R_1+|w|}{R_1-|w|}\right)\left(\alpha - \frac{m\log R - B\log R_1}{\log R/R_1}\right) \\ &= \frac{1}{2}\left(\frac{R_1+R_2}{R_1-R_2}\right)\left(\alpha - \frac{m\log R - B\log R_1}{\log R/R_1}\right) \\ &= B_0(R, R_1, R_2)(B-m). \end{aligned}$$

This proves Lemma 2.6. □

LEMMA 2.7. *For $k = 1, 2$, we have*

$$\left|\frac{d^k}{d\varphi^k}h(e^{2\pi i\varphi})\right| \leq B_k(R, R_1, R_2)(B-m)$$

with constants B_1, B_2 given by (2.1).

Proof. Since h is harmonic in D_{R_1}, we have that for any constant c,

$$h(e^{2\pi i\varphi}) - c = \int_0^1 (h(R_2 e^{2\pi i\theta}) - c) P_{1/R_2}(\varphi - \theta)\, d\theta.$$

We take a derivative in φ and estimate h using Lemma 2.6. This gives

$$\begin{aligned}\left|\frac{d}{d\varphi} h(e^{2\pi i\varphi})\right| &\le \Big(\min_{c\in\mathbb{R}} \max_{|w|\le R_2} |h(w) - c|\Big) \int_0^1 \left|\frac{\partial}{\partial\varphi} P_{1/R_2}(\varphi - \theta)\right| d\theta \\ &\le B_0(R, R_1, R_2)(B - m) \int_0^1 \left|\frac{\partial}{\partial\theta} P_{1/R_2}(\theta)\right| d\theta.\end{aligned}$$

We recall that

$$P_{1/R_2}(\theta) = \frac{R_2^2 - 1}{R_2^2 - 2R_2\cos(2\pi\theta) + 1}$$

and therefore

$$\frac{\partial}{\partial\theta} P_{1/R_2}(\theta) = -\frac{4\pi R_2(R_2^2 - 1)\sin(2\pi\theta)}{(R_2^2 - 2R_2\cos(2\pi\theta) + 1)^2}.$$

Since $\sin(2\pi\theta)$ changes sign at $\theta = 1/2$, we conclude that

$$\begin{aligned}\int_0^1 \left|\frac{\partial}{\partial\theta} P_{1/R_2}(\theta - \varphi)\right| d\theta &= -\int_0^{1/2} \frac{\partial}{\partial\theta} P_{1/R_2}(\theta)\, d\theta + \int_{1/2}^1 \frac{\partial}{\partial\theta} P_{1/R_2}(\theta)\, d\theta \\ &= 2P_{1/R_2}(0) - 2P_{1/R_2}(1/2) \\ &= \frac{8R_2}{R_2^2 - 1}.\end{aligned}$$

This proves the claim for $k = 1$.

For the second derivative, we argue similarly. We have

$$\left|\frac{d^2}{d\varphi^2} h(e^{2\pi i\varphi})\right| \le B_0(R, R_1, R_2)(B - m) \int_0^1 \left|\frac{\partial^2}{\partial\theta^2} P_{1/R_2}(\theta)\right| d\theta,$$

where

$$\frac{\partial^2}{\partial\theta^2} P_{1/R_2}(\theta) = \frac{-8\pi^2 R_2(R_2^2 - 1)(2R_2\cos^2(2\pi\theta) + (R_2^2 + 1)\cos(2\pi\theta) - 4R_2)}{(R_2^2 - 2R_2\cos(2\pi\theta) + 1)^3}.$$

By symmetry, we may restrict our attention to $\theta \in [0, 1/2]$ from now on. On that interval, the function $(\partial^2/\partial\theta^2) P_{1/R_2}$ has exactly one zero. Its location (call it $\theta_0 \in [0, 1/2]$) is given by

$$\theta_0 = \frac{1}{2\pi} \arccos\left(\frac{\sqrt{R_2^4 + 34R_2^2 + 1} - (R_2^2 + 1)}{4R_2}\right). \tag{2.17}$$

It is easy to see that $(\partial^2/\partial\theta^2) P_{1/R_2}$ is negative on $[0, \theta_0)$, and hence positive on $(\theta_0, 1/2]$. Therefore

$$\begin{aligned}\int_0^1 \left|\frac{\partial^2}{\partial\theta^2} P_{1/R_2}(\theta)\right| d\theta &= -2\int_0^{\theta_0} \frac{\partial^2}{\partial\theta^2} P_{1/R_2}(\theta)\, d\theta + 2\int_{\theta_0}^{1/2} \frac{\partial^2}{\partial\theta^2} P_{1/R_2}(\theta)\, d\theta \\ &= 2\frac{\partial}{\partial\theta} P_{1/R_2}(0) - 4\frac{\partial}{\partial\theta} P_{1/R_2}(\theta_0) + 2\frac{\partial}{\partial\theta} P_{1/R_2}(1/2) \\ &= \frac{16\pi R_2(R_2^2 - 1)\sin(2\pi\theta_0)}{(R_2^2 - 2R_2\cos(2\pi\theta_0) + 1)^2},\end{aligned}$$

since $(\partial/\partial\theta)P_{1/R_2}(0) = (\partial/\partial\theta)P_{1/R_2}(1/2) = 0$. When we evaluate the last expression using the definition (2.17) of θ_0, we obtain the quantity

$$\frac{16\pi(R_2^2-1)\sqrt{16R_2^2-(1+R_2^2-\sqrt{R_2^4+34R_2^2+1})^2}}{(\sqrt{R_2^4+34R_2^2+1}-3R_2^2-3)^2}.$$

This proves the claim for $k=2$. □

3. $\mathbb{T}^1$ *splitting lemma*

For any $f \in L^1(\mathbb{T})$, let $\langle f\rangle = \int_{\mathbb{T}} f(x)\,dx$. For a function f on $\mathbb{C}$, let us denote $f(e(x))$ by $f(x)$ for simplicity. For a Borel set U, let $|U|$ be its Lebesgue measure.

LEMMA 3.1. *Let v be as in Theorem 2.2. Assume that for some constant c,*

$$v(x) = v_1(x) + v_0(x) + c \tag{3.1}$$

with $\|v_1\|_{L^1(\mathbb{T})} < \varepsilon_1$ and $\|v_0\|_{L^\infty(\mathbb{T})} < \varepsilon_0$. Then we have

$$\begin{gathered}\int_{\mathbb{T}} \exp\bigg(\frac{\pi}{4}\delta_0^{-1}|v(x)-c|\bigg)\,dx \le C_0,\\ \delta_0 := \tfrac{9}{2}\,\varepsilon_0 + 2B_3(R, R_1, R_2)\sqrt{\varepsilon_1(B-m)},\\ C_0 := 2\sqrt{2}\exp\bigg(\pi\bigg[\frac{17}{144}+\frac{B_1}{16B_3^2}\bigg]\bigg),\end{gathered} \tag{3.2}$$

with constants given by Definition 2.1.

As a corollary of the exponential integrability, we have the following estimate on the level sets from Markov's inequality.

COROLLARY 3.2. *For any $\varepsilon_2 > 0$, we have*

$$|\{x\in\mathbb{T} : |v(x)| > \varepsilon_2\}| \le 2\sqrt{2}\exp\bigg(\frac{\pi}{4}\bigg[\frac{17}{144}+\frac{B_1}{16B_3^2}-\varepsilon_2\delta_0^{-1}\bigg]\bigg)$$

with δ_0 as in (3.2).

Note that this level set estimate is only useful if $\varepsilon_2 \gg \varepsilon_0$ and $\varepsilon_2^2 \gg \varepsilon_1(B-m)$.

Proof of Lemma 3.1. For simplicity, we will denote $B_3(R, R_1, R_2)$ by B_3 throughout the proof. We will first show the following special form of the Riesz representation, valid only on the unit circle. The idea is simply to reflect the part of the disk outside the circle back inside it.

LEMMA 3.3. *Let $1 < R_2 < R_1 < R$ and v be defined as in Theorem 2.2. Then there exist a positive measure $\tilde{\mu}$ and a harmonic function $\tilde{h}$ on D_R such that*

$$v(e(\varphi)) = \int_{\overline{D_1}} \log|z-e(\varphi)|\tilde{\mu}(dz) + \tilde{h}(e(\varphi)), \tag{3.3}$$

with the estimate

$$\tilde{\mu}(\overline{D_1}) \le \frac{B-m}{\log R/R_1}, \tag{3.4}$$

and $\tilde{h}$ satisfies the bound in Lemma 2.6 on the circle as well as

$$\left|\frac{d^k}{d\varphi^k}\tilde{h}(e(\varphi))\right| \le B_k(R, R_1, R_2)(B-m), \qquad k = 1, 2, \tag{3.5}$$

where B_1, B_2 are the same constants as those in Theorem 2.2.

Proof. By Theorem 2.2, we have

$$v(w) = \int_{D_{R_1}} \log|z-w|\mu(dz) + h(w), \tag{3.6}$$

with $\mu(D_{R_1}) \le (B-m)/(\log R/R_1)$, and $|(d^k/d\varphi^k)h(e(\varphi))| \le B_k(R, R_1, R_2)(B-m)$, $k = 1, 2$.

Let us define μ^* by reflection, that is,

$$\mu^*(E) = \mu(E^*), \tag{3.7}$$

where

$$E^* = \{\overline{z^{-1}} : z \in E\}$$

for any measurable set $E \subset \mathbb{C}$. Then for any $|w| = 1$,

$$\begin{aligned}
&\int_{D_{R_1}} \log|z-w|\,\mu(dz)\\
&= \int_{\overline{D_1}} \log|z-w|\mu(dz) + \int_{D_{R_1}\setminus\overline{D_1}} \log|z-w|\mu(dz)\\
&= \int_{\overline{D_1}} \log|z-w|\mu(dz) + \int_{D_{R_1}\setminus\overline{D_1}} \log|w-\overline{z^{-1}}|\mu(dz) + \int_{D_{R_1}\setminus\overline{D_1}} \log|z|\mu(dz)\\
&= \int_{\overline{D_1}} \log|z-w|\mu(dz) + \int_{D_1\setminus\overline{D_{1/R_1}}} \log|w-z|\mu^*(dz) + \int_{D_{R_1}\setminus\overline{D_1}} \log|z|\mu(dz)\\
&= \int_{\overline{D_1}} \log|z-w|\tilde{\mu}(dz) + \int_{D_{R_1}\setminus\overline{D_1}} \log|z|\mu(dz),
\end{aligned} \tag{3.8}$$

where, for any $E \subset \overline{D_1}$,

$$\tilde{\mu}(E) = \mu(E) + \mu^*(E \cap (D_1\setminus\overline{D_{1/R_1}})) = \mu(E) + \mu(E^* \cap (D_{R_1}\setminus\overline{D_1})). \tag{3.9}$$

By (3.9), it is clear that we have the following estimate for $\tilde{\mu}$:

$$\tilde{\mu}(\overline{D_1}) = \mu(D_{R_1}) \le \frac{B-m}{\log(R/R_1)}. \tag{3.10}$$

By (3.6) and (3.8), we have for $|w| = 1$,

$$v(w) = \int_{\overline{D_1}} \log|z-w|\tilde{\mu}(dz) + h(w) + \int_{D_{R_1}\setminus\overline{D_1}} \log|z|\mu(dz), \tag{3.11}$$

in which the third term is a constant. Let us take $\tilde{h} = h + \int_{D_{R_1}\setminus\overline{D_1}} \log|z|\,\mu(dz)$. Since $\tilde{h}$ only differs from h by a constant, the estimates on the derivatives still hold. □

The Riesz representation (3.3) allows us to give upper bounds on the parameters ε_0 and ε_1 in Lemma 3.1 in terms of $B - m$. This will be relevant in the proof of that lemma.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

COROLLARY 3.4. *We may always assume in Lemma 3.1 that*

$$\varepsilon_0 \le B_0(R, R_1, R_2)(B - m), \qquad \varepsilon_1 \le \frac{13}{20}\frac{B - m}{\log(R/R_1)}. \tag{3.12}$$

Alternatively, we can assume that $\varepsilon_0 = 0$ *and*

$$\varepsilon_1 = \left(B_0(R, R_1, R_2) + \frac{13}{20\log(R/R_1)}\right)(B - m). \tag{3.13}$$

Proof. In view of (3.3) we set

$$v_0(\varphi) := \tilde{h}(e(\varphi)), \qquad v_1(\varphi) = \int_{\overline{D_1}} \log|z - e(\varphi)|\tilde{\mu}(dz).$$

Then ε_0 is the constant from Lemma 2.6 and we claim that

$$\varepsilon_1 := \|\log|1 - e(\varphi)|\|_{L^1_\varphi}\|\tilde{\mu}\|$$

is an admissible choice. Indeed,

$$\begin{aligned}\|v_1\|_{L^1(\mathbb{T})} &\le \int_{\overline{D_1}} \|\log|z - e(\varphi)|\|_{L^1(\mathbb{T}_\varphi)}\tilde{\mu}(dz)\\ &= \int_{\overline{D_1}} \|\log||z| - e(\varphi)|\|_{L^1(\mathbb{T}_\varphi)}\tilde{\mu}(dz)\\ &\le \max_{0\le r\le 1} \|\log|r - e(\varphi)|\|_{L^1(\mathbb{T}_\varphi)}\|\tilde{\mu}\|.\end{aligned}$$

Set $h(r) := \|\log|r - e(\varphi)|\|_{L^1(\mathbb{T}_\varphi)}$ with $0 \le r \le 1$. In order to establish the claim, it suffices to verify that $h(r)$ is non-decreasing. First,

$$\int_0^1 \log|r - e(\varphi)|\,d\varphi = \int_0^1 \log|1 - re(\varphi)|\,d\varphi = 0$$

since $\log|1 - r\zeta|$ is harmonic in ζ for $|\zeta| < 1$ and any fixed $0 \le r \le 1$. Therefore, if $0 < r < 1$ and $0 < \varphi_0(r) < \frac{1}{2}$ is the unique solution of $|r - e(\varphi_0)| = 1$, then

$$h(r) = 2\int_{\varphi_0(r)}^{1-\varphi_0(r)} \log|r - e(\varphi)|\,d\varphi = \int_{\varphi_0(r)}^{1-\varphi_0(r)} \log(1 + r^2 - 2r\cos(2\pi\varphi))\,d\varphi.$$

Consequently,

$$\begin{aligned}h'(r) &= \int_{\varphi_0(r)}^{1-\varphi_0(r)} \frac{2(r - \cos(2\pi\varphi))}{1 + r^2 - 2r\cos(2\pi\varphi)}\,d\varphi\\ &\ge \int_{\varphi_0(r)}^{1-\varphi_0(r)} \frac{r}{1 + r^2 - 2r\cos(2\pi\varphi)}\,d\varphi \ge 0.\end{aligned}$$

In the second line we used that on the domain of integration

$$|r - e(\varphi)|^2 = 1 + r^2 - 2r\cos(2\pi\varphi) \ge 1,$$

whence $2r - 2\cos(2\pi\varphi) \ge r$. Therefore, indeed $h(r) \le h(1)$, justifying our choice of ε_1 above. Finally,

$$\begin{aligned}h(1) &= \|\log|1 - e(\varphi)|\|_{L^1_\varphi} = -2\int_0^1 \min(\log|1 - e(\varphi)|, 0)\,d\varphi\\ &= -2\int_{-1/6}^{1/6} \log|1 - e(\varphi)|\,d\varphi = -4\int_0^{1/6} \log(2\sin(\pi\varphi))\,d\varphi < 13/20\end{aligned}$$

and $\|\tilde{\mu}\|$ is controlled by (3.4). $\square$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Definition 3.5. Henceforth we impose the condition that

$$289\left(B_0(R, R_1, R_2) + \frac{13}{20\log R/R_1}\right) < B_3^2(R, R_1, R_2), \tag{3.14}$$

where the constants are those from Definition 2.1.

Returning to the proof of Lemma 3.1, we denote the first term in (3.3) by u, namely,

$$u(x) = \int_{\overline{D_1}} \log|z - e(x)|\, \tilde{\mu}(dz). \tag{3.15}$$

Then $v = u + \tilde{h}$. For any $f \in L^1(\mathbb{T})$ with $\langle f\rangle = 0$, the anti-derivative $D^{-1}f$ is uniquely defined as the absolutely continuous function

$$(D^{-1}f)(t) = \int_0^t f(x)\,dx + m(f), \quad \langle D^{-1}f\rangle = 0, \tag{3.16}$$

for arbitrary $t \in \mathbb{T}$. The constant $m(f)$ is chosen to ensure the vanishing mean. For $e_n(x) = \exp(2\pi i n x)$, one has $D^{-1}(e_n) = (2\pi i n)^{-1}e_n$ for all $n \neq 0$, whereas $D^{-1}e_0 = 0$. In the distributional sense, D^{-1} also applies to (complex) measures. For example, with δ_0 now being the Dirac delta,

$$D^{-1}(\delta_0 - 1)(x) = -(x + \tfrac{1}{2})\mathbb{1}_{[-1/2<x<0]}(x) + (\tfrac{1}{2} - x)\mathbb{1}_{[0<x<1/2]}(x).$$

For any $z = |z|e(y) \in D_1$, one has the elementary relation

$$\begin{aligned}\frac{d}{dx}\log|e(x) - z| &= \frac{2\pi|z|\sin(2\pi(x-y))}{1 - 2|z|\cos(2\pi(x-y)) + |z|^2}\\ &= \pi Q_z(x)\\ &= \pi(\mathcal{H}[P_z])(x),\end{aligned} \tag{3.17}$$

where $\mathcal{H}$ denotes the Hilbert transform and Q_z is the standard notation for the conjugate function of the Poisson kernel. In particular,

$$\log|e(x) - z| = \pi(D^{-1}\mathcal{H}[P_z])(x) \tag{3.18}$$

holds for any $z \in D_1$. We thus have

$$u(x) = (D^{-1}\mathcal{H}[\nu])(x), \tag{3.19}$$

where

$$\frac{d\nu}{dx}(x) = \pi\int_{\overline{D_1}} P_z(x)\tilde{\mu}(dz)$$

is a positive measure, with

$$\nu(\mathbb{T}) = \pi\tilde{\mu}(\overline{D_1}). \tag{3.20}$$

Set

$$\epsilon := \frac{1}{B_3}\sqrt{\frac{\varepsilon_1}{B - m}}, \tag{3.21}$$

and define $J_\epsilon(x) := (1/2\epsilon)\mathbb{1}_{[-\epsilon,\epsilon]}(x)$ to be the box kernel. Because of the upper bound in (3.13) on ε_1 and (3.14), one has

$$\varepsilon_1 \leq \left(B_0(R, R_1, R_2) + \frac{13}{20\log R/R_1}\right)(B - m) < \frac{B_3^2}{289}(B - m),$$

which ensures that $\epsilon < \frac{1}{17}$. We will use this smallness property for the remainder of the proof. For example, it guarantees that J_ϵ is in fact well defined on the circle (less than $\frac{1}{2}$ is enough here, but below we will need this sharper bound). Then

$$\begin{aligned} v &= v - J_\epsilon * v + J_\epsilon * v_1 + J_\epsilon * v_0 \\ &= (u - J_\epsilon * u) + (\tilde{h} - J_\epsilon * \tilde{h}) + J_\epsilon * v_1 + J_\epsilon * v_0. \end{aligned}$$

The last three terms have small L^∞ norms, in the sense that

$$\begin{cases} \|J_\epsilon * v_1\|_{L^\infty(\mathbb{T})} \le \|J_\epsilon\|_{L^\infty(\mathbb{T})}\|v_1\|_{L^1(\mathbb{T})} < \dfrac{\varepsilon_1}{2\epsilon} = \dfrac{B_3}{2}\sqrt{\varepsilon_1(B-m)}, \\ \|J_\epsilon * v_0\|_{L^\infty(\mathbb{T})} \le \|J_\epsilon\|_{L^1(\mathbb{T})}\|v_0\|_{L^\infty(\mathbb{T})} < \varepsilon_0, \\ \|\tilde{h} - J_\epsilon * \tilde{h}\|_{L^\infty(\mathbb{T})} \le \dfrac{\epsilon}{2}\|\tilde{h}'\|_{L^\infty(\mathbb{T})} \le \dfrac{1}{2}\epsilon B_1(B-m) = \dfrac{B_1}{2B_3}\sqrt{\varepsilon_1(B-m)}. \end{cases}$$

Hence

$$|(u - J_\epsilon * u)(x)| \ge |v(x) - C| - \varepsilon_0 - \left(\frac{B_3}{2} + \frac{B_1}{2B_3}\right)\sqrt{\varepsilon_1(B-m)}. \tag{3.22}$$

By (3.19), we have

$$\begin{aligned} (u - J_\epsilon * u)(x) &= (D^{-1}\mathcal{H}[\nu - J_\epsilon * \nu])(x) \\ &= \mathcal{H}[D^{-1}(\nu - J_\epsilon * \nu)](x). \end{aligned} \tag{3.23}$$

Next, we control the pointwise size of the term in brackets in (3.23). Since the Hilbert transform eliminates constants, the integration constant in (3.16) drops out.

LEMMA 3.6. *Modulo additive constants, the function $D^{-1}(\nu - J_\epsilon * \nu)$ satisfies*

$$\|D^{-1}(\nu - J_\epsilon * \nu)\|_{L^\infty(\mathbb{T})} \le \tfrac{9}{2}\varepsilon_0 + 2B_3\sqrt{\varepsilon_1(B-m)}. \tag{3.24}$$

Proof. We begin with the observation that (recall ν is a positive measure)

$$\begin{aligned} |(\nu - J_\epsilon * \nu)([a,b])| &= \left|\int_{\mathbb{T}} (\mathbb{1}_{[a,b]} - \mathbb{1}_{[a,b]} * J_\epsilon)(x)\nu(dx)\right| \\ &\le \sup_{\theta\in\mathbb{T}} \nu([\theta-\epsilon, \theta+\epsilon]), \end{aligned} \tag{3.25}$$

uniformly in $[a, b] \subset \mathbb{T}$. If $b - a \ge 2\epsilon$, then on the one hand,

$$|(\mathbb{1}_{[a,b]} - \mathbb{1}_{[a,b]} * J_\epsilon)(x)| \le \tfrac{1}{2}(\mathbb{1}_{[a-\epsilon,a+\epsilon]} + \mathbb{1}_{[b-\epsilon,b+\epsilon]})(x).$$

On the other hand, if $b - a < 2\epsilon$ then by translation invariance it suffices to consider the symmetric expression

$$f(x) := \mathbb{1}_{[-d,d]} - \mathbb{1}_{[-d,d]} * J_\epsilon, \quad 2d = b - a,\ d < \epsilon.$$

For any $0 < d < \epsilon/2$ this function satisfies

$$\begin{aligned} |f| &\le \frac{d}{\epsilon}\mathbb{1}_{(-\epsilon-d,\epsilon+d)} + \left(1 - \frac{2d}{\epsilon}\right)\mathbb{1}_{[-d,d]} \\ &\le \frac{d}{\epsilon}\mathbb{1}_{(-\epsilon-d,\epsilon-d]} + \frac{d}{\epsilon}\mathbb{1}_{(\epsilon-d,2\epsilon-d)} + \left(1 - \frac{2d}{\epsilon}\right)\mathbb{1}_{(-\epsilon,\epsilon)} \end{aligned}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

whereas for $\epsilon/2 < d < \epsilon$ one has

$$|f| \leq \tfrac{1}{2}\mathbb{1}_{(-\epsilon-d,\epsilon+d)} \leq \tfrac{1}{2}\mathbb{1}_{(-\epsilon-d,\epsilon-d]} + \tfrac{1}{2}\mathbb{1}_{(\epsilon-d,3\epsilon-d)}.$$

In either case (3.25) holds.

It therefore suffices to estimate $\sup_{\theta\in\mathbb{T}} \nu([\theta-\epsilon, \theta+\epsilon])$. Next, we define an atom τ' in the Hardy space $H^1(\mathbb{T})$ as follows:

$$\tau'(x) = \begin{cases} (x-(a-3\epsilon))/\epsilon^2, & a-3\epsilon \leq x \leq a-2\epsilon, \\ ((a-\epsilon)-x)/\epsilon^2, & a-2\epsilon \leq x \leq a-\epsilon, \\ ((a+\epsilon)-x)/\epsilon^2, & a+\epsilon \leq x \leq a+2\epsilon, \\ (x-(a+3\epsilon))/\epsilon^2, & a+2\epsilon \leq x \leq a+3\epsilon, \\ 0 & \text{otherwise.} \end{cases} \tag{3.26}$$

Note that this is well defined on the circle since $\epsilon < \frac{1}{6}$. By construction, $\langle\tau'\rangle = 0$. Set $\tau(x) := \int_{a-1/2}^{x} \tau'(t)\,dt$. Moreover, $\tau \geq 0$, $\langle\tau\rangle = 4\epsilon$, and $\tau(x) = 1$ on $[a-\epsilon, a+\epsilon]$. Thus

$$\nu([a-\epsilon, a+\epsilon]) \leq \int_{\mathbb{T}} \tau(x)\nu(dx) = \langle\tau, \nu\rangle. \tag{3.27}$$

Let us consider

$$\begin{aligned} &|(\tau - \langle\tau\rangle, \nu)| \\ &\quad= \left|\left(\frac{d}{dx}\mathcal{H}[\tau], D^{-1}\mathcal{H}[\nu]\right)\right| = |(\mathcal{H}[\tau'], u)| \\ &\quad= |(\mathcal{H}[\tau'], v - \tilde{h})| = |(\mathcal{H}[\tau'], v_0 + v_1 - \tilde{h})| \\ &\quad\leq |(\mathcal{H}[\tau'], v_0)| + |(\mathcal{H}[\tau'], v_1)| + |(\tau, \mathcal{H}[\tilde{h}'])| \\ &\quad\leq \|\mathcal{H}[\tau']\|_{L^1(\mathbb{T})}\|v_0\|_{L^\infty(\mathbb{T})} + \|\mathcal{H}[\tau']\|_{L^\infty(\mathbb{T})}\|v_1\|_{L^1(\mathbb{T})} + \|\tau\|_{L^1(\mathbb{T})}\|\mathcal{H}[\tilde{h}']\|_{L^\infty(\mathbb{T})} \\ &\quad\leq \varepsilon_0\|\mathcal{H}[\tau']\|_{L^1(\mathbb{T})} + \varepsilon_1\|\mathcal{H}[\tau']\|_{L^\infty(\mathbb{T})} + 2\epsilon\left\|\frac{d^2}{dx^2}\tilde{h}\right\|_{L^\infty(\mathbb{T})}. \end{aligned} \tag{3.28}$$

In the last line, we used the following lemma on the third term.

LEMMA 3.7. *For any $f \in C^1(\mathbb{T})$ one has $\|\mathcal{H}[f]\|_\infty \leq \frac{1}{2}\|f'\|_\infty$.*

Proof. Since $\sin(\pi x) \geq 2x$ for all $0 < x \leq \frac{1}{2}$, one has

$$\|\mathcal{H}[f]\|_\infty = \sup_{y\in\mathbb{T}}\left|\int_{\mathbb{T}} \frac{f(x)-f(y)}{\sin(\pi(x-y))}\cos(\pi(x-y))\,dx\right| \leq \frac{1}{2}\|f'\|_\infty$$

as claimed. □

In order to bound the the other terms in the last line of (3.28) we prove two lemmas.

LEMMA 3.8. *Let τ' be defined by (3.26) and assume that $0 < \epsilon \leq \frac{1}{17}$. Then we have*

$$\|\mathcal{H}[\tau']\|_{L^1(\mathbb{T})} \leq \tfrac{9}{2}. \tag{3.29}$$

Remark 3.9. The upper bound $\frac{1}{17}$ is a particular choice which we have found to be convenient in the last section of the paper. A more restrictive assumption on ϵ will slightly improve the bound; for example, assuming $\epsilon < \frac{1}{36}$ yields the value 4.2 for $\frac{9}{2}$. Such improvements are mainly due to the lower bound in (3.40) approaching π as $\epsilon \to 0$.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Proof. By translation symmetry, we may assume $a = 0$. We let $0 < \epsilon \leq \frac{1}{17}$ and $I = [-r\epsilon, r\epsilon]$ with $r = 3.6 > 3$. Notice that $r\epsilon < \frac{1}{2}$. We decompose $\|\mathcal{H}[\tau']\|_{L^1(\mathbb{T})}$ into the two parts

$$\|\mathcal{H}[\tau']\|_{L^1(\mathbb{T})} = \|\mathcal{H}[\tau']\|_{L^1(I)} + \|\mathcal{H}[\tau']\|_{L^1(I^c)}.$$

By Cauchy–Schwarz and the fact that the Hilbert transform is an isometry on $L^2(\mathbb{T})$, we have

$$\|\mathcal{H}[\tau']\|_{L^1(I)} \leq \sqrt{|I|}\|\mathcal{H}[\tau']\|_{L^2(I)} \leq \sqrt{|I|}\|\tau'\|_{L^2(I)} = \frac{2\sqrt{2}}{\sqrt{3}}\sqrt{r}.$$

In the remainder of the proof, we bound $\|\mathcal{H}[\tau']\|_{L^1(I^c)}$. By symmetry, we have $\|\mathcal{H}[\tau']\|_{L^1(I^c)} = 2\|\mathcal{H}[\tau']\|_{L^1([r\epsilon,1/2])}$. Hence, it suffices to consider the interval $[r\epsilon, \frac{1}{2}]$. For all $x \in [r\epsilon, 1/2]$, we have

$$\begin{aligned}
|\mathcal{H}[\tau'](x)| &= \left|\int_{\operatorname{supp}\tau'} \tau'(y)\cot(\pi(x-y))\,dy\right| \\
&= \left|\int_{\operatorname{supp}\tau'} \tau'(y)(\cot(\pi(x-y)) - \cot(\pi x))\,dy\right| \\
&= \left|\int_{\operatorname{supp}\tau'} \tau'(y)\frac{\sin(\pi y)}{\sin(\pi(x-y))\sin(\pi x)}\,dy\right| \\
&\leq \int_{\operatorname{supp}\tau'} |\tau'(y)|\frac{|\sin(\pi y)|}{\sin(\pi(x-y))\sin(\pi x)}\,dy.
\end{aligned} \tag{3.30}$$

Here we used that $x - y \in [(r-3)\epsilon, \frac{1}{2} + 3\epsilon] \subset [0, 1]$. To estimate this expression further, we decompose $\operatorname{supp}\tau'$ into the four intervals

$$I_1 := [2\epsilon, 3\epsilon], \quad I_2 := [\epsilon, 2\epsilon], \quad I_3 := [-2\epsilon, -\epsilon], \quad I_4 := [-3\epsilon, -2\epsilon].$$

We write $\rho_j(x) \in I_j$ for the point in I_j that is nearest to x in the toroidal distance, that is, $\|x - \rho_j(x)\|_{\mathbb{T}} = \operatorname{dist}_{\mathbb{T}}(x, I_j)$. (That point is not unique if x is 'antipodal' to the center of I_j; in this case we define ρ_j as the right endpoint of I_j for definiteness.) Notice that $\rho_j(x)$ is constant for $j \in \{1, 2\}$ with

$$\rho_1(x) = \rho_{1,1} := 3\epsilon, \quad \rho_2(x) = \rho_{2,1} := 2\epsilon, \tag{3.31}$$

and piecewise constant for $j \in \{3, 4\}$, that is,

$$\rho_j(x) = \begin{cases} \rho_{j,1} & \text{if } x \in [0, t_j], \\ \rho_{j,2} & \text{if } x \in (t_j, \frac{1}{2}]. \end{cases}$$

Specifically, we have

$$\begin{aligned}
\rho_3(x) &= \begin{cases} -\epsilon & \text{if } x \in \left[0, \dfrac{1}{2} - \dfrac{3\epsilon}{2}\right], \\ -2\epsilon & \text{if } x \in \left(\dfrac{1}{2} - \dfrac{3\epsilon}{2}, \dfrac{1}{2}\right], \end{cases} \\
\rho_4(x) &= \begin{cases} -2\epsilon & \text{if } x \in \left[0, \dfrac{1}{2} - \dfrac{5\epsilon}{2}\right], \\ -3\epsilon & \text{if } x \in \left(\dfrac{1}{2} - \dfrac{5\epsilon}{2}, \dfrac{1}{2}\right]. \end{cases}
\end{aligned} \tag{3.32}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

From (3.30) and $|\sin(\pi y)| \leq \pi |y|$, we obtain

$$|\mathcal{H}[\tau'](x)| \leq \pi \sum_{j=1}^{4} \frac{1}{\sin(\pi(x-\rho_j(x)))\sin(\pi x)} \int_{I_j} |\tau'(y)||y|\,dy.$$

We integrate both sides over $x \in [r\epsilon, \frac{1}{2}]$ and find

$$\|\mathcal{H}[\tau']\|_{L^1([r\epsilon,1/2])} \leq \pi \sum_{j=1}^{4} f_j(\epsilon) \int_{I_j} |\tau'(y)||y|\,dy, \tag{3.33}$$

where we introduced the notation

$$f_j(\epsilon) := \int_{r\epsilon}^{1/2} \frac{1}{\sin(\pi(x-\rho_j(x)))\sin(\pi x)}\,dx,$$

for $j \in \{1, 2, 3, 4\}$. The following lemma gives a bound on these integrals.

LEMMA 3.10. *For $j \in \{1, 2\}$, we have*

$$f_j(\epsilon) \leq \frac{1}{2.98\pi\rho_{j,1}} \log\biggl(1 + \frac{\pi}{2.98}\frac{\rho_{j,1}}{r\epsilon - \rho_{j,1}}\biggr), \tag{3.34}$$

and for $j \in \{3, 4\}$, we have

$$f_j(\epsilon) \leq \frac{2}{5\pi|\rho_{j,1}|} \log\biggl(1 + \frac{2\pi}{5}\frac{|\rho_{j,1}|}{r\epsilon}\biggr) + 0.21. \tag{3.35}$$

We postpone the proof of Lemma 3.10 for now. To continue the proof of Lemma 3.8, recall (3.33). We perform the integration in y and find

$$\begin{aligned}\int_{I_1} |\tau'(y)||y|\,dy &= \int_{I_4} |\tau'(y)||y|\,dy = \frac{7\epsilon}{6},\\ \int_{I_2} |\tau'(y)||y|\,dy &= \int_{I_3} |\tau'(y)||y|\,dy = \frac{5\epsilon}{6}.\end{aligned} \tag{3.36}$$

Then we apply Lemma 3.10 and recall Definitions (3.31) and (3.32) of $\rho_{j,1}$. This gives

$$\begin{aligned}\|\mathcal{H}[\tau']\|_{L^1([r\epsilon,1/2])} &\leq \pi \sum_{j=1}^{4} f_j(\epsilon)|\tau'(y)||y|\,dy\\ &\leq \frac{1}{17.88}\biggl(\frac{7}{3}\log\biggl(1+\frac{\pi}{2.98}\frac{3}{r-3}\biggr) + \frac{5}{2}\log\biggl(1+\frac{\pi}{2.98}\frac{2}{r-2}\biggr)\biggr)\\ &\quad + \frac{1}{15}\biggl(5\log\biggl(1+\frac{2\pi}{5}\frac{1}{r}\biggr) + \frac{7}{2}\log\biggl(1+\frac{2\pi}{5}\frac{2}{r}\biggr)\biggr) + (0.42)\pi\epsilon.\end{aligned} \tag{3.37}$$

Notice that $(0.42)\pi\epsilon < 0.08$. We write $\nu(r)$ for the expression in the last line. Altogether, we have shown that

$$\|\mathcal{H}[\tau']\|_{L^1(\mathbb{T})} \leq \frac{2\sqrt{2}}{\sqrt{3}}\sqrt{r} + 2\nu(r) < 4.5.$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

In the second step, we evaluated the expression at $r = 3.6$. This proves the main claim of Lemma 3.8. □

It remains to give the proof of Lemma 3.10.

Proof of Lemma 3.10. We begin by observing that

$$\frac{d}{dx}\left(\frac{1}{\sin x'}\log\left(\frac{\sin(x-x')}{\sin(x)}\right)\right) = \frac{1}{\sin x \, \sin(x-x')}, \tag{3.38}$$

whenever x, x' are such that the logarithm is well defined.

Let $j \in \{1, 2\}$, which implies that $\rho_j(x) = \rho_{j,1} > 0$. By (3.38) and sin, cos ≤ 1, we have

$$\begin{aligned} f_j(\epsilon) &= \frac{1}{\pi \, \sin(\pi\rho_{j,1})}\left[\log\left(\frac{\sin(\pi(x-\rho_{j,1}))}{\sin(\pi x)}\right)\right]_{x=r\epsilon}^{x=1/2} \\ &\leq \frac{1}{\pi \, \sin(\pi\rho_{j,1})}\log\left(\frac{\sin(\pi r\epsilon)}{\sin(\pi(r\epsilon-\rho_{j,1}))}\right) \\ &\leq \frac{1}{\pi \, \sin(\pi\rho_{j,1})}\log\left(1+\frac{\sin(\pi\rho_{j,1})}{\sin(\pi(r\epsilon-\rho_{j,1}))}\right). \end{aligned} \tag{3.39}$$

Next, we estimate the sines by linear functions. While the upper bound $\sin(\pi x) \leq \pi x$ is valid for all x (and is sharp for small x), a linear lower bound on $\sin(\pi x)$ depends directly on the allowed range of x values. This is conveniently expressed via the quotient

$$\inf_{x\in[0,3\epsilon]}\frac{\sin(\pi x)}{x} = \frac{\sin(3\pi\epsilon)}{3\epsilon} > 2.98.$$

In the last step, we used that $\epsilon < \frac{1}{17}$. We may verify that all the arguments of $\sin(\pi\,\cdot)$ in the last line of (3.39) are located in the interval $[0, 3\epsilon]$. Therefore

$$\begin{aligned} f_j(\epsilon) &\leq \frac{1}{\pi \, \sin(\pi\rho_{j,1})}\log\left(1+\frac{\sin(\pi\rho_{j,1})}{\sin(\pi(r\epsilon-\rho_{j,1}))}\right) \\ &\leq \frac{1}{2.98\pi\rho_{j,1}}\log\left(1+\frac{\pi}{2.98}\frac{\rho_{j,1}}{r\epsilon-\rho_{j,1}}\right). \end{aligned}$$

This proves (3.34).

Next, let $j \in \{3, 4\}$, so that $\rho_{j,1}, \rho_{j,2} < 0$. We have $r\epsilon < t_j$ by our assumptions on r, ϵ, and (3.38) yields

$$\begin{aligned} f_j(\epsilon) \leq{} & \frac{1}{\pi \, \sin(\pi\rho_{j,1})}\left[\log\left(\frac{\sin(\pi(x-\rho_{j,1}))}{\sin(\pi x)}\right)\right]_{x=r\epsilon}^{x=t_j} \\ &+ \left(\frac{1}{2}-t_j\right)\max_{x\in(t_j,1/2]}\frac{1}{\sin(\pi(x-\rho_{j,2}))\,\sin(\pi x)}. \end{aligned}$$

The second term is an error term (it vanishes as $\epsilon \to 0$). Indeed, recalling the definition of t_j and $\rho_{j,2}$ from (3.32), we see that for $j \in \{3, 4\}$,

$$\left(\frac{1}{2}-t_j\right)\frac{1}{\sin(\pi(\frac{1}{2}-\rho_{j,2}))\,\sin(\pi t_j)} \leq \frac{5\epsilon}{2\cos^2(3\pi\epsilon)} \leq \frac{5\epsilon}{2(1-\frac{9}{2}\pi^2\epsilon^2)^2} \leq 0.21,$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

where the last estimate used that $\epsilon \leq \frac{1}{17}$. Therefore, we have

$$\begin{aligned}
&\int_{r\epsilon}^{1/2} \frac{1}{\sin(\pi(x-\rho_j(x)))\sin(\pi x)}\,dx \\
&\quad \leq \frac{1}{\pi \sin(\pi\rho_{j,1})}\left[\log\left(\frac{\sin(\pi(x-\rho_{j,1}))}{\sin(\pi x)}\right)\right]_{x=r\epsilon}^{x=t_j} + 0.21 \\
&\quad \leq \frac{1}{\pi \sin(\pi|\rho_{j,1}|)} \log\left(\frac{\sin(\pi(r\epsilon-\rho_{j,1}))}{\sin(\pi r\epsilon)}\right) + 0.21 \\
&\quad \leq \frac{2}{5\pi|\rho_{j,1}|} \log\left(1+\frac{2\pi}{5}\frac{|\rho_{j,1}|}{r\epsilon}\right) + 0.21.
\end{aligned}$$

In the second step, we used that $0 \leq t_j < t_j - \rho_{j,1} \leq \frac{1}{2}$ and monotonicity properties of sin. In the last step, we used $\cos \leq 1$ and

$$\inf_{x\in[0,(r+2)\epsilon]} \frac{\sin(\pi x)}{x} = \frac{\sin(\pi(r+2)\epsilon)}{(r+2)\epsilon} > \frac{5}{2}. \tag{3.40}$$

This bound holds because $(r+2)\epsilon < \frac{1}{3}$, which may be verified from $r = 3.6$ and $\epsilon \leq \frac{1}{17}$. This shows (3.35) and concludes the proof of Lemma 3.10, and hence also of Lemma 3.8. □

Next, we control the pointwise size of $\mathcal{H}[\tau']$.

LEMMA 3.11. *We have*

$$\|\mathcal{H}[\tau']\|_{L^\infty(\mathbb{T})} \leq \frac{5}{2\epsilon}.$$

Proof. By translation invariance, we can set $a = 0$. Let us first consider

$$\int_{\mathrm{supp}(\tau')} \tau'(x)\cot\pi(x-y)\,dy.$$

If $x \notin \mathrm{supp}(\tau')$, then $\int_{\mathrm{supp}(\tau')} \tau'(x)\cot\pi(x-y)\,dy = 0$. Thus, we can assume without loss of generality that $x \in [-3\epsilon, -\epsilon]$. On the one hand,

$$\begin{aligned}
\left|\int_{[\epsilon,3\epsilon]} \tau'(x)\cot\pi(x-y)\,dy\right| &\leq 2\epsilon|\tau'(x)| \sup_{y\in[\epsilon,3\epsilon]} |\cot\pi(x-y)| \\
&\leq 2\sup_{u\in[2\epsilon,6\epsilon]} |\cot(\pi u)| \leq 2\cot(2\pi\epsilon) \leq \frac{1}{2\epsilon},
\end{aligned}$$

since $6\epsilon\pi \leq \pi - 2\pi\epsilon$ and $\sin(2\pi\epsilon) \geq 4\epsilon$ (recall that $\epsilon < \frac{1}{8}$). On the other hand, for the negative support of τ', we can further assume by symmetry that $x \in [-3\epsilon, -2\epsilon]$. Thus,

$$\begin{aligned}
&\left|\int_{[-3\epsilon,-\epsilon]} \tau'(x)\cot\pi(x-y)\,dy\right| \\
&\quad = \left|\tau'(x)\int_{2x+3\epsilon}^{-\epsilon} \cot\pi(y-x)\,dy\right| = \left|\tau'(x)\int_{x+3\epsilon}^{-\epsilon-x} \cot(\pi y)\,dy\right| \\
&\quad = \frac{1}{\epsilon^2}(x+3\epsilon)\int_{x+3\epsilon}^{-\epsilon-x} \frac{1}{2y}\,dy = \frac{1}{\epsilon}\frac{x+3\epsilon}{2\epsilon}\log\left(\frac{2\epsilon}{x+3\epsilon}-1\right) \\
&\leq \frac{1}{\epsilon}\sup_{t\in[0,1/2]} t\log\left(\frac{1}{t}-1\right) < \frac{1}{4\epsilon}.
\end{aligned} \tag{3.41}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Hence, overall we have

$$\left|\int_{\operatorname{supp}(\tau')} \tau'(x) \cot \pi(x-y)\, dy\right| < \frac{3}{4\epsilon}. \tag{3.42}$$

Now let us consider, with $x \in [-3\epsilon, -\epsilon]$,

$$\begin{aligned}
&\left|\mathcal{H}[\tau'](x) - \int_{\operatorname{supp}(\tau')} \tau'(x) \cot \pi(x-y)\, dy\right| \\
&\quad = \left|\int_{\operatorname{supp}(\tau')} (\tau'(y)-\tau'(x)) \cot \pi(x-y)\, dy\right| \\
&\quad \le \int_{-3\epsilon}^{-\epsilon} \left|\frac{\tau'(y)-\tau'(x)}{\sin \pi(y-x)}\right| dy + \int_{\epsilon}^{3\epsilon} \frac{|\tau'(y)|+|\tau'(x)|}{|\sin \pi(y-x)|}\, dy \\
&\quad \le \int_{-3\epsilon}^{-\epsilon} \frac{|x-y|\epsilon^{-2}}{2|x-y|}\, dy + \left(\frac{1}{4}+\frac{1}{2}\log 2\right)\epsilon^{-1} \le \frac{1.6}{\epsilon}.
\end{aligned} \tag{3.43}$$

In summary,

$$\|\mathcal{H}[\tau']\|_{L^\infty(\mathbb{T})} < \frac{5}{2\epsilon} \tag{3.44}$$

as claimed. □

Combining (3.20), (3.27), (3.28) with Lemmas 2.5, 3.8 and 3.11, we have

$$\begin{aligned}
(\tau, \nu) &\le \frac{9}{2}\varepsilon_0 + 5\frac{\varepsilon_1}{2\epsilon} + 2\epsilon B_2(B-m) + \langle\tau\rangle\nu(\mathbb{T}) \\
&\le \frac{9}{2}\varepsilon_0 + 5\frac{\varepsilon_1}{2\epsilon} + \epsilon\left(2B_2 + \frac{4\pi}{\log R/R_1}\right)(B-m) \\
&= \frac{9}{2}\varepsilon_0 + 2B_3\sqrt{\varepsilon_1(B-m)}.
\end{aligned}$$

Note that our choice (3.21) of ϵ minimizes the contribution of ε_1. Finally, (3.25) concludes the proof of Lemma 3.6. □

In order to prove the exponential integrability of $v - c$, and thus complete the proof of Lemma 3.1, we invoke the following classical result about the Hilbert transform of bounded functions on the circle; see, for example, [**Kat**].

LEMMA 3.12. *Let f be a real-valued function on $\mathbb{T}$ such that $|f| \le 1$. Then for any $0 \le \alpha < \frac{1}{2}\pi$,*

$$\int_{\mathbb{T}} e^{\alpha|\mathcal{H}[f](x)|}\, dx \le \frac{2}{\cos\alpha} = 2\sec\alpha.$$

Applying Lemma 3.12 to $f = (D^{-1}(\nu - J_\epsilon * \nu))/\|(D^{-1}(\nu - J_\epsilon * \nu))\|_{L^\infty(\mathbb{T})}$, by (3.23), we have

$$\int_{\mathbb{T}} \exp(\beta|(u - J_\epsilon * u)(x)|)\, dx \le 2\sec\alpha, \tag{3.45}$$

where $\alpha = \beta(\frac{9}{2}\varepsilon_0 + 2B_3\sqrt{\varepsilon_1(B-m)}) < \pi/2$. Taking $\alpha = \pi/4$ in (3.45), then (3.22) yields (absorbing the constant c into v for simpicity)

$$\begin{aligned}\int_{\mathbb{T}} \exp\bigg(\frac{\pi|v(x)|}{18\varepsilon_0 + 8B_3\sqrt{\varepsilon_1(B-m)}}\bigg)\,dx \\ &\leq 2\sqrt{2}\exp\bigg(\frac{\pi\varepsilon_0 + \pi(B_3 + B_1/B_3)\sqrt{\varepsilon_1(B-m)}/2}{18\varepsilon_0 + 8B_3\sqrt{\varepsilon_1(B-m)}}\bigg) \\ &\leq 2\sqrt{2}\exp\bigg(\pi\bigg[\frac{17}{144} + \frac{B_1}{16B_3^2}\bigg]\bigg), \end{aligned} \tag{3.46}$$

which concludes the proof of Lemma 3.1. □

We conclude this section with an important decay estimate on the Fourier coefficients of the subharmonic function v. This lemma will be used in §7.

LEMMA 3.13. *Let v be as in Theorem 2.2. Then the Fourier coefficients of v satisfy*

$$|\hat{v}(k)| \leq \frac{C(R, R_1, R_2)}{|k|}(B-m) \quad \textit{for any } k \neq 0,$$

in which

$$C(R, R_1, R_2) = \frac{1}{2\log R/R_1} + \frac{1}{2\pi}B_1(R, R_1, R_2). \tag{3.47}$$

Proof. For any $k \neq 0$, we have

$$|\hat{v}(k)| = \bigg|\int_{\mathbb{T}} (u(x) + \tilde{h}(x))e^{-2\pi ikx}\,dx\bigg| \leq \frac{1}{2\pi|k|}(|\widehat{u'}(k)| + |\widehat{\tilde{h}'}(k)|). \tag{3.48}$$

By Lemma 3.3, we have $|\tilde{h}'(x)| \leq B_1(R, R_1, R_2)(B-m)$, hence

$$|\widehat{\tilde{h}'}(k)| \leq B_1(R, R_1, R_2)(B-m). \tag{3.49}$$

By (3.19), (3.20) and Theorem 2.2, we have

$$|\widehat{u'}(k)| = |\widehat{\mathcal{H}[v]}(k)| = |\hat{v}(k)| \leq \pi\tilde{\mu}(\overline{D_1}) \leq \frac{\pi(B-m)}{\log R/R_1}. \tag{3.50}$$

In view of (3.48), (3.49) and (3.50) we infer that

$$|\hat{v}(k)| \leq \frac{1}{2\pi|k|}\bigg(B_1(R, R_1, R_2)(B-m) + \frac{\pi(B-m)}{\log R/R_1}\bigg),$$

as claimed. □

4. $\mathbb{T}^2$ *splitting lemma*

Our applications to the skew-shift dynamics on $\mathbb{T}^2$ require a version of the splitting lemma in two variables. First, we formalize the class of plurisubharmonic functions that we will be working with.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Definition 4.1. Let $v(z, w)$ be a continuous plurisubharmonic function on $D_R \times D_R$, satisfying the following estimates for $R > 1$:

$$\begin{cases} v(z, w) \leq B_4(R) \text{ for all } (z, w) \in D_R \times \partial D_1, \text{ and } v(0, e(y)) \geq m_4 \text{ for all } y \in \mathbb{T}, \\ v(z, w) \leq B_5(R) \text{ for all } (z, w) \in \partial D_1 \times D_R, \text{ and } v(e(x), 0) \geq m_5 \text{ for all } x \in \mathbb{T}, \\ |v(e(x), e(y))| \leq B_6 \text{ for all } (x, y) \in \mathbb{T}^2. \end{cases} \tag{4.1}$$

For a function f defined on a polydisk in $\mathbb{C}^2$ which contains $\mathbb{T}^2$, let us denote $f(e(x), e(y))$ by $f(x, y)$ for simplicity. In particular, we will write $v(x, y)$ on $\mathbb{T}^2$. The average is denoted by $\langle f \rangle_{\mathbb{T}^2} := \int_{\mathbb{T}^2} f(x, y)\, dx\, dy$.

Below, we will analyze a particular Schrödinger cocycle over a skew-shift base, and we will specify the constants in Definitions 2.1 and 4.1. But for now we develop more analytical machinery with these constants as parameters. Recall that for a Borel set U, the Lebesgue measure will be written as $|U|$.

LEMMA 4.2. *Let v be as in Definition 4.1, and assume (3.14). Let $0 < r < 1$, $0 < \varepsilon_1$ and $0 < \varepsilon_0 < \varepsilon_3 < \varepsilon_2$. Assume*

$$v(x, y) = v_0(x, y) + v_1(x, y) + \langle v \rangle_{\mathbb{T}^2},$$

where $\|v_0\|_{L^\infty(\mathbb{T}^2)} < \varepsilon_0$ and $\|v_1\|_{L^1(\mathbb{T}^2)} < \varepsilon_1$. Then

$$\begin{aligned} |\{(x, y) \in \mathbb{T}^2 : |v(x, y) - \langle v \rangle_{\mathbb{T}^2}| > \varepsilon_2\}| &< 2(2C_0)^{1/2} \exp\left(-\frac{\pi}{8\delta_0^{(1)}} \varepsilon_3\right) \\ &\quad + C_0 \exp\left(-\frac{\pi}{4\delta_0^{(2)}} \varepsilon_2\right) \end{aligned}$$

in which $B_3 = B_3(R, R_1, R_2)$ are defined as in (2.2), C_0 in (3.2), and

$$\begin{aligned} \delta_0^{(1)} &:= \tfrac{9}{2}\varepsilon_0 + 2B_3\sqrt{\varepsilon_1^r(B_4 - m_4)}, \\ \delta_0^{(2)} &:= \tfrac{9}{2}\varepsilon_3 + 4B_3\sqrt{B_6}\sqrt{\varepsilon_1^{1-r}(B_5 - m_5)}. \end{aligned} \tag{4.2}$$

Proof. Fix $0 < r < 1$. Let

$$A_1 := \left\{ y \in \mathbb{T} : \int_{\mathbb{T}} |v_1(x, y)|\, dx < \varepsilon_1^r \right\}. \tag{4.3}$$

By Markov's inequality, we have

$$|A_1^c| < \varepsilon_1^{1-r}. \tag{4.4}$$

For any fixed $y \in A_1$, we have $\|v_0(\cdot, y)\|_{L^\infty(\mathbb{T})} < \varepsilon_0$ and $\|v_1(\cdot, y)\|_{L^1(\mathbb{T})} < \varepsilon_1^r$. Applying Lemma 3.1 in the x variable, we then have

$$\int_{\mathbb{T}} \exp\left(\frac{\pi}{4\delta_0^{(1)}} |v(x, y) - \langle v \rangle_{\mathbb{T}^2}|\right) dx \leq C_0.$$

Integrating over $y \in A_1$ and interchanging the integrations yields

$$\int_{\mathbb{T}} \int_{A_1} \exp\left(\frac{\pi}{4\delta_0^{(1)}} |v(x, y) - \langle v \rangle_{\mathbb{T}^2}|\right) dy\, dx \leq C_0. \tag{4.5}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

For $\gamma > 0$, let us define

$$A_2 := \left\{ x \in \mathbb{T} : \int_{A_1} \exp\left(\frac{\pi}{4\,\delta_0^{(1)}} |v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}| \right) dy \le C_0\, \gamma^{-1} \right\}. \tag{4.6}$$

By Markov's inequality,

$$|A_2^c| < \gamma. \tag{4.7}$$

For $x \in A_2$ and $\varepsilon_3 > \varepsilon_0$, let us define

$$A_3 := \{ y \in A_1 : |v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}| < \varepsilon_3 \}. \tag{4.8}$$

Again, by Markov's inequality,

$$|A_3^c| < C_0 \gamma^{-1} \exp\left(-\frac{\pi\, \varepsilon_3}{4\,\delta_0^{(1)}} \right). \tag{4.9}$$

Thus for $x \in A_2$, $\{ y \in \mathbb{T} : |v(x,\, y) - \langle v \rangle| > \varepsilon_3 \} \subseteq A_1^c \cup A_3^c$, with the measure estimate

$$|\{ y \in \mathbb{T} : |v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}| > \varepsilon_3 \}| \le \varepsilon_1^{1-r} + C_0 \gamma^{-1} \exp\left(-\frac{\pi\, \varepsilon_3}{4\delta_0^{(1)}} \right). \tag{4.10}$$

Here we divide the argument into two different cases, depending on which term on the right-hand side dominates.

Case 1: $\varepsilon_1^{1-r} < C_0 \gamma^{-1} \exp(-\pi\, \varepsilon_3 / 4\,\delta_0^{(1)})$. Then (4.10) directly implies that for any $x \in A_2$,

$$|\{ y \in \mathbb{T} : |v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}| > \varepsilon_3 \}| \le 2C_0 \gamma^{-1} \exp\left(-\frac{\pi\, \varepsilon_3}{4\delta_0^{(1)}} \right). \tag{4.11}$$

Together with (4.7), we conclude that

$$|\{ (x,\, y) \in \mathbb{T}^2 : |v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}| > \varepsilon_3 \}| \le \gamma + 2C_0 \gamma^{-1} \exp\left(-\frac{\pi\, \varepsilon_3}{4\delta_0^{(1)}} \right). \tag{4.12}$$

Case 2: $\varepsilon_1^{1-r} \ge C_0 \gamma^{-1} \exp(-\pi\, \varepsilon_3 / 4\,\delta_0^{(1)})$. Then for any $x \in A_2$,

$$|\{ y \in \mathbb{T} : |v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}| > \varepsilon_3 \}| \le 2\varepsilon_1^{1-r}. \tag{4.13}$$

For $x \in A_2$, let

$$\begin{aligned} \tilde{v}_{x,0}(y) &= (v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}) \mathbb{1}_{\{ y \in \mathbb{T} : |v(x,y) - \langle v \rangle_{\mathbb{T}^2}| \le \varepsilon_3 \}}, \\ \tilde{v}_{x,1}(y) &= (v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}) \mathbb{1}_{\{ y \in \mathbb{T} : |v(x,y) - \langle v \rangle_{\mathbb{T}^2}| > \varepsilon_3 \}}. \end{aligned} \tag{4.14}$$

Then (4.13) implies, assuming $x \in A_2$,

$$\begin{gathered} v(x,\, y) = \tilde{v}_{x,0}(y) + \tilde{v}_{x,1}(y) + \langle v \rangle_{\mathbb{T}^2}, \\ \|\tilde{v}_{x,0}(\cdot)\|_{L^\infty(\mathbb{T})} \le \varepsilon_3, \\ \|\tilde{v}_{x,1}(\cdot)\|_{L^1(\mathbb{T})} \le 2\varepsilon_1^{1-r} \|v(x,\, \cdot) - \langle v \rangle_{\mathbb{T}^2}\|_{L^\infty(\mathbb{T})} \le 4B_6\, \varepsilon_1^{1-r}. \end{gathered} \tag{4.15}$$

Applying Corollary 3.2 in the y variable, we obtain that for any $x \in A_2$ and any $\varepsilon_2 > \varepsilon_3$,

$$|\{ y \in \mathbb{T} : |v(x,\, y) - \langle v \rangle_{\mathbb{T}^2}| > \varepsilon_2 \}| \le C_0 \exp\left(-\frac{\pi\, \varepsilon_2}{4\delta_0^{(2)}} \right). \tag{4.16}$$

Together with (4.7), we then get

$$|\{(x, y) \in \mathbb{T}^2 : |v(x, y) - \langle v\rangle_{\mathbb{T}^2}| > \varepsilon_2\}| \le \gamma + C_0 \exp\biggl(-\frac{\pi\varepsilon_2}{4\delta_0^{(2)}}\biggr). \tag{4.17}$$

Finally, we choose γ to equalize the terms in (4.12):

$$\gamma = (2C_0)^{1/2} \exp\biggl(-\frac{\pi\varepsilon_3}{8\,\delta_0^{(1)}}\biggr).$$

Then the estimate of case 1, namely (4.12), yields

$$\begin{aligned} &|\{(x, y) \in \mathbb{T}^2 : |v(x, y) - \langle v\rangle_{\mathbb{T}^2}| > \varepsilon_2\}| \\ &\quad \le |\{(x, y) \in \mathbb{T}^2 : |v(x, y) - \langle v\rangle_{\mathbb{T}^2}| > \varepsilon_3\}| \\ &\quad \le 2(2C_0)^{1/2} \exp\biggl(-\frac{\pi\varepsilon_3}{8\delta_0^{(1)}}\biggr). \end{aligned} \tag{4.18}$$

The estimate of case 2, namely (4.17), becomes

$$\begin{aligned} &|\{(x, y) \in \mathbb{T}^2 : |v(x, y) - \langle v\rangle_{\mathbb{T}^2}| > \varepsilon_2\}| \\ &\quad \le (2C_0)^{1/2} \exp\biggl(-\frac{\pi\varepsilon_3}{8\delta_0^{(1)}}\biggr) + C_0 \exp\biggl(-\frac{\pi\varepsilon_2}{4\delta_0^{(2)}}\biggr). \end{aligned} \tag{4.19}$$

Combining (4.18) with (4.19), we conclude that

$$\begin{aligned} &|\{(x, y) \in \mathbb{T}^2 : |v(x, y) - \langle v\rangle_{\mathbb{T}^2}| > \varepsilon_2\}| \\ &\quad \le 2(2C_0)^{1/2} \exp\biggl(-\frac{\pi\varepsilon_3}{8\delta_0^{(1)}}\biggr) + C_0 \exp\biggl(-\frac{\pi\varepsilon_2}{4\delta_0^{(2)}}\biggr), \end{aligned}$$

as claimed. □

5. *Avalanche principle*

The avalanche principle (AP) is a device to compare the logarithm of the norm of a long product $A_n A_{n-1} \cdots A_2 A_1$ of matrices to the sum of the logarithms of the norms of shorter sections of the product. In the original formulation from [**GolSch**] for $\mathrm{SL}_2(\mathbb{R})$ matrices the length of the chain was limited depending on the norms of the individual matrices A_j. The same restriction applied to the extension of the AP to $\mathrm{SL}_d(\mathbb{R})$ matrices in [**Sch**]. Later, Duarte and Klein [**DuaKle**] found a different proof of the AP which does not impose any restriction on the length of the chain. Even though the older version [**GolSch**] would suffice for our purposes, we present the argument from [**DuaKle**] with explicit constants. (These are not provided in [**DuaKle**].)

Thus, this section is devoted to making the constants in [**DuaKle2**, Ch. 2] effective (we mostly follow [**DuaKle2**] instead of [**DuaKle**] for the sake of simplicity). We use the same notation as [**DuaKle2**], which we first recall. Although we only need the results for $\mathrm{SL}_2(\mathbb{R})$ matrices in this paper, we aim to prove more general results which are of independent interest.

Let $\mathrm{GL}_d(\mathbb{R})$ be the general linear group of real $d \times d$ matrices.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Definition 5.1. Given matrices $g_0, g_1, \ldots, g_{n-1} \in \mathrm{GL}_d(\mathbb{R})$, the expansion rift is the ratio

$$\rho(g_0, g_1, \ldots, g_{n-1}) := \frac{\|g_{n-1} \cdots g_1 g_0\|}{\|g_{n-1}\| \cdots \|g_1\| \|g_0\|} \in (0, 1].$$

Given $g \in \mathrm{GL}_d(\mathbb{R})$, let

$$s_1(g) \geq s_2(g) \geq \cdots \geq s_d(g) > 0$$

denote the sorted singular values of g. The first singular value, $s_1(g)$, is the operator norm

$$s_1(g) = \max_{x \in \mathbb{R}^d \setminus \{0\}} \frac{\|gx\|}{\|x\|} := \|g\|.$$

The last singular value of g is the least expansion factor of g, regarded as a linear transformation, and it can be characterized by

$$s_d(g) = \min_{x \in \mathbb{R}^d \setminus \{0\}} \frac{\|gx\|}{\|x\|} = \|g^{-1}\|^{-1}.$$

Definition 5.2. The gap (or the singular gap) of $g \in \mathrm{GL}_d(\mathbb{R})$ is the ratio between its first and second singular values.

$$\mathrm{gr}(g) := \frac{s_1(g)}{s_2(g)}.$$

Remark 5.3. If $g \in \mathrm{SL}_2(\mathbb{R})$, then $\mathrm{gr}(g) = \|g\|^2$.

Let $\mathbb{P}(\mathbb{R}^d)$ denote the projective space. Points in $\mathbb{P}(\mathbb{R}^d)$ are equivalence classes $\hat{x}$ of non-zero vectors $x \in \mathbb{R}^d$. We consider the projective distance $\delta : \mathbb{P}(\mathbb{R}^d) \times \mathbb{P}(\mathbb{R}^d) \to [0, 1]$,

$$\delta(\hat{x}, \hat{y}) := \sin(\angle(x, y)),$$

where $\angle$ is the length of the arc connecting x and y.

Definition 5.4. Given $g \in \mathrm{GL}_d(\mathbb{R})$ such that $\mathrm{gr}(g) > 1$, the most expanding direction of g is the singular direction $\hat{\mathfrak{v}} \in \mathbb{P}(\mathbb{R}^d)$ associated with the first singular value $s_1(g)$ of g. Let $\mathfrak{v}(g)$ be any of the two unit vector representatives of the projective point $\hat{\mathfrak{v}}(g)$. We set $\hat{\mathfrak{v}}^*(g) := \hat{\mathfrak{v}}(g^*)$ and $\mathfrak{v}^*(g) := \mathfrak{v}(g^*)$.

Any matrix $g \in \mathrm{GL}_d(\mathbb{R})$ maps the most expanding direction of g to the most expanding direction of g^*, multiplying vectors by the factor $s_1(g) = \|g\|$:

$$g\mathfrak{v}(g) = \pm s_1(g)\mathfrak{v}^*(g).$$

The matrix g also induces a projective map $\hat{g} : \mathbb{P}(\mathbb{R}^d) \to \mathbb{P}(\mathbb{R}^d)$, $\hat{g}(\hat{x}) := \widehat{gx}$, for which one has

$$\hat{g}(\hat{\mathfrak{v}}(g)) = \hat{\mathfrak{v}}^*(g) \quad \text{and} \quad \widehat{g^*}(\hat{\mathfrak{v}}^*(g)) = \hat{\mathfrak{v}}(g).$$

THEOREM 5.5. *Let $n \geq 1$ and $0 < \epsilon \leq \frac{1}{10}$. Given $0 < \kappa \leq \frac{1}{10}\epsilon^2$ and $g_0, g_1, \ldots, g_{n-1} \in \mathrm{GL}_d(\mathbb{R})$, if*

(G) $\mathrm{gr}(g_i) \geq \kappa^{-1}$ *for* $j = 0, 1, \ldots, n-1$,

(A) $\rho(g_{j-1}, g_j) \geq \epsilon$ *for* $j = 1, 2, \ldots, n-1$,

then, writing $g^j := g_{j-1} \cdots g_1 g_0$, we have:

(i) $\max\{\delta(\hat{\mathfrak{v}}(g^n), \hat{\mathfrak{v}}(g_0)), \delta(\hat{\mathfrak{v}}^*(g^n), \hat{\mathfrak{v}}^*(g_{n-1}))\} \leq 3\kappa\epsilon^{-1}$;

(ii) $e^{-5n\kappa/\epsilon^2} \leq (\rho(g_0, g_1, \ldots, g_{n-1}))/(\rho(g_0, g_1) \cdots \rho(g_{n-2}, g_{n-1})) \leq e^{11n\kappa/\epsilon^2}$.

The proof follows the general line of argumentation in [**DuaKle2**], keeping track of the effective constants throughout.

Staging the proof. The projective distance $\delta : \mathbb{P}(\mathbb{R}^d) \times \mathbb{P}(\mathbb{R}^d) \to [0, 1]$ determines a complementary angle function $\alpha : \mathbb{P}(\mathbb{R}^d) \times \mathbb{P}(\mathbb{R}^d) \to [0, 1]$, defined by

$$\alpha(\hat{x}, \hat{y}) := |\cos(\angle(x, y))|.$$

Let us also introduce the algebraic operation

$$a \oplus b := a + b - ab.$$

For properties of $a \oplus b$, one may refer to [**DuaKle2**, Proposition 2.1].

LEMMA 5.6. *Given $g \in \mathrm{GL}_d(\mathbb{R})$ with $\mathrm{gr}(g) > 1$, $\hat{x} \in \mathbb{P}(\mathbb{R}^d)$ and a unit vector $x \in \hat{x}$, writing $\alpha = \alpha(\hat{x}, \hat{\mathfrak{v}}(g))$ we have the following statements.*

(a) $\alpha \leq \|gx\|/\|x\| \leq \sqrt{\alpha^2 \oplus \mathrm{gr}(g)^{-2}}$.

(b) $\delta(\hat{g}(\hat{x}), \hat{\mathfrak{v}}^*(g)) \leq \alpha^{-1}\, \mathrm{gr}(g)^{-1} \delta(\hat{x}, \hat{\mathfrak{v}}(g))$.

(c) *The restriction of the map $\hat{g} : \mathbb{P}(\mathbb{R}^d) \to \mathbb{P}(\mathbb{R}^d)$ to the disk*

$$\{\hat{x} \in \mathbb{P}(\mathbb{R}^d) : \delta(\hat{x}, \hat{\mathfrak{v}}(g)) \leq r\}$$

has Lipschitz constant at most $(\pi(r + \sqrt{1 - r^2}))/(2\, \mathrm{gr}(g)(1 - r^2))$ with respect to the δ-metric.

Proof. The factor $\pi/2$ in the Lipschitz constant is already explicit in [**DuaKle2**, proof of Lemma 2.2]. □

COROLLARY 5.7. *Given $g \in \mathrm{GL}_d(\mathbb{R})$ such that $\mathrm{gr}(g) \geq \kappa^{-1}$, define*

$$\Sigma_\epsilon := \{\hat{x} \in \mathbb{P}(\mathbb{R}^d) : \alpha(\hat{x}, \hat{\mathfrak{v}}(g)) \geq \epsilon\} = B(\hat{\mathfrak{v}}(g), \sqrt{1 - \epsilon^2}).$$

Given a point $\hat{x} \in \Sigma_\epsilon$:

(a) $\delta(\hat{g}(\hat{x}), \hat{g}(\hat{\mathfrak{v}}(g))) \leq \kappa\epsilon^{-1}\delta(\hat{x}, \hat{\mathfrak{v}}(g))$*;*

(b) *the map $\hat{g}|_{\Sigma_\epsilon} \to \mathbb{P}(\mathbb{R}^d)$ has Lipschitz constant at most $(\sqrt{2}\pi/2)\kappa\epsilon^{-2}$.*

Proof. Inequality (a) follows directly from (b) of Lemma 5.6. Statement (b) follows from (c) of Lemma 5.6 and the fact that $\epsilon + \sqrt{1 - \epsilon^2} \leq \sqrt{2}$. □

Definition 5.8. Given $g, g' \in \mathrm{GL}_d(\mathbb{R})$ with $\mathrm{gr}(g)$, $\mathrm{g}(g') > 1$, we define their lower angle as

$$\alpha(g, g') := \alpha(\hat{\mathfrak{v}}^*(g), \hat{\mathfrak{v}}(g')).$$

The upper angle between g and g' is

$$\beta(g, g') := \sqrt{\mathrm{gr}(g)^{-2} \oplus \alpha(g, g')^2 \oplus \mathrm{gr}(g')^{-2}}.$$

LEMMA 5.9. *Given $g, g' \in \mathrm{GL}_d(\mathbb{R})$, if $\mathrm{gr}(g), \mathrm{gr}(g') > 1$, then*

$$\alpha(g, g') \leq \rho(g, g') \leq \beta(g, g').$$

This lemma has the following immediate corollary. It shows how assumptions (G) and (A) in Theorem 5.5 will be used.

COROLLARY 5.10. *Given $g, g' \in \mathrm{GL}_d(\mathbb{R})$, if $\mathrm{gr}(g), \mathrm{gr}(g') \geq \kappa^{-1}$ and $\rho(g, g') \geq \epsilon$, then*

$$\delta(\hat{\mathfrak{v}}^*(g), \hat{\mathfrak{v}}(g')) \leq \sqrt{1 - \frac{\epsilon^2}{1 + 2(\kappa^2/\epsilon^2)}}.$$

We recall that $g^j := g_{j-1} \cdots g_1 g_0$.

LEMMA 5.11. *If $\mathrm{gr}(g_j) > 1$ for $j = 0, 1, \ldots, n-1$, and $\mathrm{gr}(g^j) > 1$ for $j = 1, 2, \ldots, n$, then*

$$\prod_{j=1}^{n-1} \alpha(g^j, g_j) \leq \rho(g_0, g_1, \ldots, g_{n-1}) \leq \prod_{j=1}^{n-1} \beta(g^j, g_j).$$

Proof of Theorem 5.5. To simplify the notation, we will write $c_0 = \frac{1}{10}$, $\hat{\mathfrak{v}}_j := \hat{\mathfrak{v}}(g_j)$ and $\hat{\mathfrak{v}}_j^* := \hat{\mathfrak{v}}^*(g_j)$ for $j = 0, 1, \ldots, n-1$. We also let

$$g_j = g_{2n-1-j}^*, \quad \hat{\mathfrak{v}}_j = \hat{\mathfrak{v}}_{2n-1-j}^* \quad \text{and} \quad \hat{\mathfrak{v}}_j^* = \hat{\mathfrak{v}}_{2n-1-j} \quad \text{for } j = n, n+1, \ldots, 2n-1.$$

For each $i = 0, 1, \ldots, 2n-1$ and $j = 0, 1, \ldots, 2n-i$, set

$$\hat{\mathfrak{v}}_i^j := \hat{g}_{i+j-1} \cdots \hat{g}_{i+1} \hat{g}_i \hat{\mathfrak{v}}_i.$$

In terms of the notation above, we have $\widehat{(g^n)^* g^n}(\hat{\mathfrak{v}}(g_0)) = \hat{\mathfrak{v}}_0^{2n}$ and $\hat{\mathfrak{v}}_0 = \hat{\mathfrak{v}}_{2n-1}^1$.

By assumption (A), we have $\rho(g_{j-1}, g_j) \geq \epsilon$ for $1 \leq j \leq n-1$. Hence for $n+1 \leq j \leq 2n-1$,

$$\rho(g_{j-1}, g_j) = \rho(g_{2n-j}^*, g_{2n-j-1}^*) = \rho(g_{2n-j-1}, g_{2n-j}) \geq \epsilon. \tag{5.1}$$

Clearly, we also have

$$\rho(g_{n-1}, g_n) = \rho(g_{n-1}, g_{n-1}^*) = \frac{\|g_{n-1}^* g_{n-1}\|}{\|g_{n-1}\|^2} = 1. \tag{5.2}$$

Therefore combining assumption (A) with (5.1) and (5.2), we have

$$\rho(g_{j-1}, g_j) \geq \epsilon, \quad \text{for } j = 1, 2, \ldots, 2n-1. \tag{5.3}$$

We begin with the proof of statement (i). We will prove $\delta(\hat{\mathfrak{v}}(g^n), \hat{\mathfrak{v}}(g_0)) \leq 3\kappa\epsilon^{-1}$. The other bound can be proved in exactly the same way.

First, we will show that for $\tilde{\epsilon} = t\epsilon$, $t = 2/3$, we have the following lemma.

LEMMA 5.12. *For any $1 \leq j \leq 2n-1$,*

$$\hat{g}_{j-1}(B(\hat{\mathfrak{v}}_{j-1}, \sqrt{1-\tilde{\epsilon}^2})) \subseteq B(\hat{\mathfrak{v}}_j, \sqrt{1-\tilde{\epsilon}^2}).$$

Proof. Taking any $\hat{x} \in B(\hat{\mathfrak{v}}_{j-1}, \sqrt{1-\tilde{\epsilon}^2})$, we have

$$\delta(\hat{x}, \hat{\mathfrak{v}}_{j-1}) = \sin(\angle(\hat{x}, \hat{\mathfrak{v}}_{j-1})) \leq \sqrt{1-\tilde{\epsilon}^2}.$$

By (a) of Corollary 5.7,

$$\begin{aligned}\delta(\hat{g}_{j-1}\hat{x}, \hat{g}_{j-1}\hat{\mathfrak{v}}_{j-1}) = \delta(\hat{g}_{j-1}\hat{x}, \hat{\mathfrak{v}}_{j-1}^1) &= \sin(\angle(\hat{g}_{j-1}\hat{x}, \hat{\mathfrak{v}}_{j-1}^1)) \\ &\leq \kappa\tilde{\epsilon}^{-1}\sqrt{1-\tilde{\epsilon}^2} \leq \frac{c_0\epsilon}{t}\sqrt{1-t^2\epsilon^2}.\end{aligned} \tag{5.4}$$

By (5.3) and Corollary 5.10, we have

$$\delta(\hat{\mathfrak{v}}^*(g_{j-1}), \hat{\mathfrak{v}}(g_j)) = \delta(\hat{\mathfrak{v}}^1_{j-1}, \hat{\mathfrak{v}}_j) = \sin(\angle(\hat{\mathfrak{v}}^1_{j-1}, \hat{\mathfrak{v}}_j)) \le \sqrt{1 - \frac{\epsilon^2}{1 + 2(\kappa^2/\epsilon^2)}} \le \sqrt{1 - \frac{\epsilon^2}{1 + 2c_0^2\epsilon^2}}. \tag{5.5}$$

Let $\theta_1 = \angle(\hat{g}_{j-1}\hat{x}, \hat{\mathfrak{v}}^1_{j-1})$ and $\theta_2 = \angle(\hat{\mathfrak{v}}^1_{j-1}, \hat{\mathfrak{v}}_j)$. Then

$$\delta(\hat{g}_{j-1}\hat{x}, \hat{\mathfrak{v}}_j) \le |\cos\theta_1|\sin\theta_2 + \sin\theta_1|\cos\theta_2| = \sqrt{1 - \sin^2\theta_1}\,\sin\theta_2 + \sin\theta_1\sqrt{1 - \sin^2\theta_2} := f(\sin\theta_1, \sin\theta_2). \tag{5.6}$$

With $f(x, y) = y\sqrt{1 - x^2} + x\sqrt{1 - y^2}$, it is easy to see that both $\partial f/\partial x$ and $\partial f/\partial y$ have the same sign as $\sqrt{1 - x^2}\sqrt{1 - y^2} - xy$. Thus both $\partial f/\partial x$ and $\partial f/\partial y$ are positive if $x^2 + y^2 < 1$.

By (5.4) and (5.5), we have

$$\sin^2\theta_1 + \sin^2\theta_2 \le \frac{c_0^2\epsilon^2}{t^2}(1 - t^2\epsilon^2) + 1 - \frac{\epsilon^2}{1 + 2c_0^2\epsilon^2} < 1. \tag{5.7}$$

Here it is enough to have that for $\tilde{\epsilon} = t\epsilon$,

$$\frac{c_0^2}{t^2} + c_0^2\epsilon^2 < 1.$$

Then (5.6) implies

$$\begin{aligned}\delta(\hat{g}_{j-1}\hat{x}, \hat{\mathfrak{v}}_j) &\le f\left(\frac{c_0\epsilon}{t}\sqrt{1 - t^2\epsilon^2}, \sqrt{1 - \frac{\epsilon^2}{1 + 2c_0^2\epsilon^2}}\right)\\ &< \left(1 - \frac{c_0^2\epsilon^2}{2t^2}(1 - t^2\epsilon^2)\right)\left(1 - \frac{\epsilon^2}{2 + 4c_0^2\epsilon^2}\right) + \frac{c_0\epsilon^2}{t}\left(1 - \frac{c_0^2\epsilon^2}{2}\right)\\ &< \sqrt{1 - t^2\epsilon^2} = \sqrt{1 - \tilde{\epsilon}^2}.\end{aligned} \tag{5.8}$$

(By our choice of $c_0 = \frac{1}{10}$ and $t = \frac{2}{3}$, the ϵ^2 coefficients of (5.8) correspond to $-\frac{9}{800} - \frac{1}{2} + \frac{3}{20} < -\frac{2}{9}$.) □

This lemma has the following intermediate corollary.

COROLLARY 5.13. *For any* $1 \le j \le 2n - 1$ *and* $1 \le m \le 2n - j - 1$, *we have*

$$\hat{g}_{j+m-1} \cdots \hat{g}_j \hat{g}_{j-1} B(\hat{\mathfrak{v}}_{j-1}, \sqrt{1 - \tilde{\epsilon}^2}) \subseteq B(\hat{\mathfrak{v}}_{j+m}, \sqrt{1 - \tilde{\epsilon}^2}).$$

Next, let us show the following lemma.

LEMMA 5.14. *For any* $0 \le j \le 2n - 1$, *for any* $\hat{x} \in B(\hat{\mathfrak{v}}_j, \sqrt{1 - \tilde{\epsilon}^2})$,

$$\delta(\hat{g}_{2n-1} \cdots \hat{g}_{j+2}\hat{g}_j\hat{x}, \hat{\mathfrak{v}}_j^{2n-j}) \le \kappa\tilde{\epsilon}^{-1}\left(\frac{\sqrt{2}\pi}{2}\kappa\tilde{\epsilon}^{-2}\right)^{2n-j-1}.$$

Proof. By Corollary 5.13, for any $0 \le m \le 2n - j - 1$, we have that both the elements $\hat{g}_{j+m-1} \cdots \hat{g}_{j+1}\hat{g}_j\hat{x}$ and $\hat{\mathfrak{v}}_j^m$ belong to $B(\hat{\mathfrak{v}}_{j+m}, \sqrt{1-\tilde{\epsilon}^2})$. Hence by (a) of Corollary 5.7, we have that for $m = 0$,

$$\delta(\hat{g}_j\hat{x}, \hat{\mathfrak{v}}_j^1) = \delta(\hat{g}_j\hat{x}, \hat{g}_j\hat{\mathfrak{v}}_j) \le \kappa\tilde{\epsilon}^{-1}\delta(\hat{x}, \hat{\mathfrak{v}}_j) < \kappa\tilde{\epsilon}^{-1}. \tag{5.9}$$

For $1 \le m \le 2n - j - 1$, by (b) of Corollary 5.7, we have

$$\begin{aligned}\delta(\hat{g}_{j+m}\hat{g}_{j+m-1} \cdots \hat{g}_{j+1}\hat{g}_j\hat{x}, \hat{\mathfrak{v}}_j^{m+1}) &= \delta(\hat{g}_{j+m}\hat{g}_{j+m-1} \cdots \hat{g}_{j+1}\hat{g}_j\hat{x}, \hat{g}_{j+m}\hat{\mathfrak{v}}_j^m) \\ &\le \frac{\sqrt{2}\pi}{2}\kappa\tilde{\epsilon}^{-2}\delta(\hat{g}_{j+m-1} \cdots \hat{g}_{j+1}\hat{g}_j\hat{x}, \hat{\mathfrak{v}}_j^m). \end{aligned}\tag{5.10}$$

Inequalities (5.9) and (5.10) imply that

$$\delta(\hat{g}_{2n-1} \cdots \hat{g}_{j+2}\hat{g}_j\hat{x}, \hat{\mathfrak{v}}_j^{2n-j}) \le \kappa\tilde{\epsilon}^{-1}\left(\frac{\sqrt{2}\pi}{2}\kappa\tilde{\epsilon}^{-2}\right)^{2n-j-1}, \tag{5.11}$$

as desired. □

In particular, combining Corollary 5.13 with Lemma 5.14, we have the following corollary.

COROLLARY 5.15. *For any* $1 \le j \le 2n - 1$,

$$\delta(\hat{\mathfrak{v}}_{j-1}^{2n-j+1}, \hat{\mathfrak{v}}_j^{2n-j}) \le \kappa\tilde{\epsilon}^{-1}\left(\frac{\sqrt{2}\pi}{2}\kappa\tilde{\epsilon}^{-2}\right)^{2n-j-1}.$$

Next, we will show the following lemma.

LEMMA 5.16. *For any* $\hat{x} \in B(\hat{\mathfrak{v}}_0, \sqrt{1-\tilde{\epsilon}^2})$*, we have*

$$\delta(\widehat{(g^n)^*g^n}\hat{x}, \hat{\mathfrak{v}}_0) \le 3\kappa\epsilon^{-1}, \quad \delta(\hat{\mathfrak{v}}_0^{2n}, \hat{\mathfrak{v}}_{2n-1}^1) \le 3\kappa\epsilon^{-1}.$$

Proof. By Corollary 5.15, we have

$$\delta(\hat{\mathfrak{v}}_0^{2n}, \hat{\mathfrak{v}}_{2n-1}^1) \le \sum_{j=1}^{2n-1} \delta(\hat{\mathfrak{v}}_{j-1}^{2n-j+1}, \hat{\mathfrak{v}}_j^{2n-j}) \le \kappa\tilde{\epsilon}^{-1}\sum_{j=0}^{2n-2}\left(\frac{\sqrt{2}\pi}{2}\kappa\tilde{\epsilon}^{-2}\right)^j. \tag{5.12}$$

By Lemma 5.14,

$$\delta(\hat{g}_{2n-1} \cdots \hat{g}_1\hat{g}_0\hat{x}, \hat{\mathfrak{v}}_0^{2n}) \le \kappa\tilde{\epsilon}^{-1}\left(\frac{\sqrt{2}\pi}{2}\kappa\tilde{\epsilon}^{-2}\right)^{2n-1}. \tag{5.13}$$

Hence, combining (5.12) with (5.13), we conclude that

$$\begin{aligned}\delta(\widehat{(g^n)^*g^n}\hat{x}, \hat{\mathfrak{v}}(g_0)) &\le \kappa\tilde{\epsilon}^{-1}\sum_{j=0}^{2n-1}\left(\frac{\sqrt{2}\pi}{2}\kappa\tilde{\epsilon}^{-2}\right)^j \\ &\le \frac{\kappa\tilde{\epsilon}^{-1}}{1-(\sqrt{2}\pi/2)\kappa\tilde{\epsilon}^{-2}} \le \frac{1}{t-(\pi c_0/\sqrt{2t})}\kappa\epsilon^{-1} < 3\kappa\epsilon^{-1}.\end{aligned}$$ □

We are now ready to give the proof of Theorem 5.5.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Proof of Theorem 5.5(i). Lemma 5.16 shows that $\widehat{(g^n)^*g^n}$ maps the ball $B(\hat{\mathfrak{v}}_0, \sqrt{1-\tilde{\epsilon}^2})$ into itself. By Corollary 5.7, it has contracting Lipschitz factor less than or equal to $((\sqrt{2}\pi/2)\kappa\tilde{\epsilon}^{-2})^{2n} \ll 1$. Therefore the map $\widehat{(g^n)^*g^n}$ has a unique fixed point in $B(\hat{\mathfrak{v}}_0, \sqrt{1-\tilde{\epsilon}^2})$; call it x_*. Lemma 5.16 implies that

$$\delta(x_*, \hat{\mathfrak{v}}(g_0)) < 3\kappa\epsilon^{-1}. \tag{5.14}$$

The claim will follow once we prove that $x_* = \hat{\mathfrak{v}}(g^n)$. Since $\hat{\mathfrak{v}}(g^n)$ is a fixed point of $\widehat{(g^n)^*g^n}$, it suffices to prove that $\hat{\mathfrak{v}}(g^n) \in B(\hat{\mathfrak{v}}_0, \sqrt{1-\tilde{\epsilon}^2})$. (Notice that $\widehat{(g^n)^*g^n}$ has several fixed points, one for every eigenvalue of $(g^n)^*g^n$.)

Let $\delta_* := \delta(\hat{\mathfrak{v}}(g^n), x_*)$. We will show that $\delta_* = 0$. For any unit vector v, we have

$$|\langle \mathfrak{v}(g^n), v\rangle| = \frac{1}{(s_1(g^n))^2}|\langle (g^n)^*g^n\mathfrak{v}(g^n), v\rangle| \le \left|\left\langle \mathfrak{v}(g^n), \frac{(g^n)^*g^n v}{|(g^n)^*g^n v|}\right\rangle\right|,$$

where we used that $|(g^n)^*g^n v| \le (s_1(g^n))^2$. This lifts to a relation on projective space:

$$\delta(\hat{\mathfrak{v}}(g^n), \widehat{(g^n)^*g^n}\hat{v}) \le \delta(\hat{\mathfrak{v}}(g^n), \hat{v}). \tag{5.15}$$

We apply this with $\hat{v} = \hat{v}_*$ as the 'halfway point' between $\hat{\mathfrak{v}}(g^n)$ and x_*, that is, $\hat{v}_*$ satisfies

$$\delta(\hat{\mathfrak{v}}(g^n), \hat{v}_*) = \delta(x_*, \hat{v}_*) = \frac{\delta_*}{2}.$$

(This $\hat{v}_*$ can be constructed by following the arc that connects $\hat{\mathfrak{v}}(g^n)$ with x_*, assuming that $\mathfrak{v}(g^n) \neq x_*$.)

Notice that $\hat{v}_* \in B(\hat{\mathfrak{v}}_0, \sqrt{1-\tilde{\epsilon}^2})$ because (5.14) gives

$$\delta(\hat{\mathfrak{v}}_0, \hat{v}_*) \le \delta(\hat{\mathfrak{v}}_0, x_*) + \frac{\delta_*}{2} \le \frac{3\kappa}{\epsilon} + \frac{\delta_*}{2} \le \frac{3\epsilon}{10} + \frac{1}{2} < \sqrt{1-\tilde{\epsilon}^2}.$$

Recall that $\widehat{(g^n)^*g^n}$ maps the ball $B(\hat{\mathfrak{v}}_0, \sqrt{1-\tilde{\epsilon}^2})$ into itself with Lipschitz factor less than or equal to $L_0 := ((\sqrt{2}\pi/2)\kappa\tilde{\epsilon}^{-2})^{2n} \ll 1$. Since $\widehat{(g^n)^*g^n}x_* = x_*$, we have

$$\delta(\widehat{(g^n)^*g^n}\hat{v}_*, x_*) \le L_0\delta(\hat{v}_*, x_*) = L_0\frac{\delta_*}{2}.$$

We combine this bound and (5.15) with $\hat{v} = \hat{v}_*$, to conclude that

$$\delta_* = \delta(\hat{\mathfrak{v}}(g^n), x_*) \le \delta(\hat{\mathfrak{v}}(g^n), \widehat{(g^n)^*g^n}\hat{v}_*) + \delta(\widehat{(g^n)^*g^n}\hat{v}_*, x_*) \le \left(1 + \frac{L_0}{2}\right)\delta_*.$$

Since $L_0 < 1$, this implies $\delta_* = 0$, that is, $x_* = \hat{\mathfrak{v}}(g^n)$. Consequently, (5.14) reads $\delta(\hat{\mathfrak{v}}(g^n), \hat{\mathfrak{v}}(g_0)) \le 3\kappa\epsilon^{-1}$ as claimed in (i) of Theorem 5.5. The other bound in (i) can be proved in exactly the same way. □

Proof of Theorem 5.5(ii). By Lemma 5.11, we have

$$\prod_{j=1}^{n-1}\frac{\alpha(g^j, g_j)}{\beta(g_{j-1}, g_j)} \le \frac{\rho(g_0, \ldots, g_{n-1})}{\prod_{j=1}^{n-1}\rho(g_{j-1}, g_j)} \le \prod_{j=1}^{n-1}\frac{\beta(g^j, g_j)}{\alpha(g_{j-1}, g_j)}. \tag{5.16}$$

We will show that the factors

$$\frac{\alpha(g^j, g_j)}{\beta(g_{j-1}, g_j)} \quad \text{and} \quad \frac{\beta(g^j, g_j)}{\alpha(g_{j-1}, g_j)}$$

are very close to 1, with logarithms of order $\kappa\epsilon^{-2}$. From conclusion (i), applied to the sequence of matrices $g_0, g_1, \ldots, g_{j-1}$, we have

$$\max\{\delta(\hat{\mathfrak{v}}^*(g^j), \hat{\mathfrak{v}}^*(g_{j-1})), \delta(\hat{\mathfrak{v}}(g^j), \hat{\mathfrak{v}}(g_0))\} \le 3\kappa\epsilon^{-1} \tag{5.17}$$

for all $1 \le j \le n$. From (5.17), we deduce that

$$\begin{aligned}\left|\log \frac{\alpha(g^j, g_j)}{\alpha(g_{j-1}, g_j)}\right| &\le \frac{|\alpha(g^j, g_j) - \alpha(g_{j-1}, g_j)|}{\min\{\alpha(g^j, g_j), \alpha(g_{j-1}, g_j)\}} \\ &\le \frac{\delta(\hat{\mathfrak{v}}^*(g^j), \hat{\mathfrak{v}}^*(g_{j-1}))}{\min\{\alpha(g^j, g_j), \alpha(g_{j-1}, g_j)\}} \le \frac{3\kappa\epsilon^{-1}}{\min\{\alpha(g^j, g_j), \alpha(g_{j-1}, g_j)\}}.\end{aligned} \tag{5.18}$$

We estimate the minimum as follows, using (5.17) and Corollary 5.10:

$$\begin{aligned}\min\{\alpha(g^j, g_j), \alpha(g_{j-1}, g_j)\} &\ge \alpha(g_{j-1}, g_j) - |\alpha(g^j, g_j) - \alpha(g_{j-1}, g_j)| \\ &\ge \alpha(g_{j-1}, g_j) - 3\kappa\epsilon^{-1} \\ &\ge \frac{\epsilon}{\sqrt{1 + 2\frac{\kappa^2}{\epsilon^2}}} - 3\frac{\kappa}{\epsilon} \\ &\ge \frac{2\epsilon}{3}.\end{aligned} \tag{5.19}$$

In the last step, we used $\kappa/\epsilon \le \epsilon/10$. Returning to (5.18), we have shown

$$\left|\log \frac{\alpha(g^j, g_j)}{\alpha(g_{j-1}, g_j)}\right| \le \frac{9}{2}\frac{\kappa}{\epsilon^2}. \tag{5.20}$$

From the definition of the upper angle β and Corollary 5.10, we also have

$$\left|\log \frac{\beta(g_{j-1}, g_j)}{\alpha(g_{j-1}, g_j)}\right| \le \log\sqrt{1 + 2\frac{\kappa^2}{\alpha(g_{j-1}, g_j)}} \le \log\sqrt{1 + 2\frac{\kappa^2}{\epsilon^2}} \le \frac{\kappa^2}{\epsilon^2}. \tag{5.21}$$

Hence (5.20) and (5.21) yield

$$\left|\log \frac{\alpha(g^j, g_j)}{\beta(g_{j-1}, g_j)}\right| \le \frac{9}{2}\frac{\kappa}{\epsilon^2} + \frac{\kappa^2}{\epsilon^2} < 5\frac{\kappa}{\epsilon^2}.$$

Together with (5.16), this implies the lower bound in (ii), that is,

$$e^{-5n\kappa/\epsilon^2} \le \frac{\rho(g_0, g_1, \ldots, g_{n-1})}{\rho(g_0, g_1)\cdots\rho(g_{n-2}, g_{n-1})}.$$

For the upper bound, we argue similarly. The only difference occurs in the analog of (5.21), that is, the estimate of

$$\left|\log \frac{\beta(g^j, g_j)}{\alpha(g^j, g_j)}\right|.$$

To bound this quantity, we need to control the gap ratio $(\mathrm{gr}(g^j))^{-1}$. This control is provided by the following lemma.

LEMMA 5.17. *We have* $(\mathrm{gr}(g^j))^{-1} \le \kappa' := 20(\kappa/\epsilon)$.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

We postpone the proof of the lemma for now. It gives

$$\left|\log \frac{\beta(g^j, g_j)}{\alpha(g^j, g_j)}\right| \le \log \sqrt{1 + 2\frac{(\kappa')^2}{\alpha(g^j, g_j)}} \le 600\frac{\kappa^2}{\epsilon^3}.$$

In the second step, we used that $\alpha(g^j, g_j) \ge 2\epsilon/3$ by (5.19). Recalling (5.20), one has

$$\left|\log \frac{\beta(g^j, g_j)}{\alpha(g_{j-1}, g_j)}\right| \le \frac{9}{2}\frac{\kappa}{\epsilon^2} + 600\frac{\kappa^2}{\epsilon^3} < 11\frac{\kappa}{\epsilon^2}.$$

In the last estimate, we used that $600\kappa/\epsilon \le 60\epsilon \le 6$.

By (5.16), this proves the upper bound in (ii), that is,

$$\frac{\rho(g_0, g_1, \ldots, g_{n-1})}{\rho(g_0, g_1)\cdots\rho(g_{n-2}, g_{n-1})} \le e^{11n\kappa/\epsilon^2}.$$

It remains to prove Lemma 5.17. For this part, we follow [**DuaKle**, p. 71] and make the constants precise. From [**DuaKle**, Proposition 2.28], we see that

$$(\mathrm{gr}(g^j))^{-1} = \|(D\hat{g^j})_{\hat{\mathfrak{v}}(g^j)}\|.$$

Where $(D\hat{g^j})_{\hat{\mathfrak{v}}(g^j)}$ is the derivative of $\hat{g^j} : \mathbb{P}(\mathbb{R}^d) \to \mathbb{P}(\mathbb{R}^d)$, evaluated at $\hat{\mathfrak{v}}(g^j)$. The norm of this derivative is bounded by the Lipschitz constant in a neighborhood of $\hat{\mathfrak{v}}(g^j)$. Since the Lipschitz constant is at most L^j, with $L := (\sqrt{2}\pi/2)\kappa\tilde{\epsilon}^{-2}$, everywhere on the ball $B(\hat{\mathfrak{v}}_0, \sqrt{1-\tilde{\epsilon}^2})$, we immediately obtain the bound $(\mathrm{gr}(g^j))^{-1} \le L^j$. However, this is not good enough for our purposes (note that L is an order-one quantity in general).

We may improve the estimate as follows. Applying statement (i) of the theorem with $n = j$, we obtain that

$$\delta(\hat{\mathfrak{v}}(g^j), \hat{\mathfrak{v}}(g_0)) < 3\frac{\kappa}{\epsilon}.$$

Now, a calculation based on [**DuaKle**, Proposition 2.28] shows that

$$\|(D\hat{g}_0)_{\hat{\mathfrak{v}}(g^j)} - (D\hat{g}_0)_{\hat{\mathfrak{v}}(g_0)}\| \le 12\pi\frac{\kappa}{\epsilon}$$

and therefore

$$\begin{aligned}\|(D\hat{g}_0)_{\hat{\mathfrak{v}}(g^j)}\| &\le \|(D\hat{g}_0)_{\hat{\mathfrak{v}}(g_0)}\| + 12\pi\frac{\kappa}{\epsilon} = \mathrm{gr}(g_0)^{-1} + 12\pi\frac{\kappa}{\epsilon} \le \kappa + 12\pi\frac{\kappa}{\epsilon}\\ &\le \left(12\pi + \frac{1}{10}\right)\frac{\kappa}{\epsilon}.\end{aligned}$$

Finally, we apply the chain rule and estimate the derivative of the product $g_{j-1}\ldots g_1$ by its Lipschitz constant L^{j-1}, which satisfies $L^{j-1} \le L \le 9\pi/40\sqrt{2}$ for $j \ge 2$. Therefore,

$$\begin{aligned}(\mathrm{gr}(g^j))^{-1} &= \|(D\hat{g^j})_{\hat{\mathfrak{v}}(g^j)}\| \le \|(D\hat{g}_{j-1}\ldots\hat{g}_1)_{\hat{\mathfrak{v}}(g^j)g}\|\|(D\hat{g}_0)_{\hat{\mathfrak{v}}(g^j)}\|\\ &\le L\left(12\pi + \frac{1}{10}\right)\frac{\kappa}{\epsilon} < 20\frac{\kappa}{\epsilon}.\end{aligned}$$

This proves Lemma 5.17 and hence completes the proof of Theorem 5.5. □

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

6. *Herman's regularization*

6.1. *Monodromy matrices.* One has $T_\omega^n(x, y) = (x + ny + (n(n-1)/2)\omega, y + n\omega)$ for any positive integer n, where T is the skew shift with frequency ω. Denote the projection of $\mathbb{T}^2$ onto the first coordinate by $\mathcal{P}$, that is, $\mathcal{P}(x, y) = x$.

We consider the Schrödinger operator

$$(H_{\lambda,\omega,x,y}\psi)_n = \psi_{n+1} + \psi_{n-1} + 2\lambda \cos(2\pi\mathcal{P}(T_\omega^n(x, y)))\psi_n$$

with $\lambda > 0$. This equation has the cocycle reformulation

$$\begin{aligned}\begin{pmatrix}\psi_{n+1}\\ \psi_n\end{pmatrix} &= \begin{pmatrix}E - 2\lambda\cos(2\pi\mathcal{P}(T_\omega^n(x, y))) & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\psi_n\\ \psi_{n-1}\end{pmatrix}\\ &=: M(\lambda, E; T_\omega^n(x, y))\begin{pmatrix}\psi_n\\ \psi_{n-1}\end{pmatrix}.\end{aligned} \tag{6.1}$$

Define the transfer matrices $M_n(\lambda, E; x, y)$ to be

$$M_n(\lambda, E; x, y) = \begin{cases}\displaystyle\prod_{j=n}^{1} M(\lambda, E; T_\omega^j(x, y)), & n \geq 1,\\ \mathrm{Id}, & n = 0,\\ (M_{-n}(\lambda, E; T_\omega^{n+1}(x, y)))^{-1}, & n < 0.\end{cases} \tag{6.2}$$

Then

$$\begin{pmatrix}\psi_{n+1}\\ \psi_n\end{pmatrix} = M_n(E; x, y)\begin{pmatrix}\psi_1\\ \psi_0\end{pmatrix}.$$

The following function on $\mathbb{T}^2$ plays a fundamental role in our analysis:

$$\begin{aligned}u_n(\lambda, E; x, y) &:= \frac{1}{n}\log\|M_n(\lambda, E; x, y)\|\\ &= \frac{1}{n}\log\left\|\prod_{j=n}^{1}\begin{pmatrix}E - \lambda e^{2\pi i(x+jy+(j(j-1)/2)\omega)} - \lambda e^{-2\pi i(x+jy+(j(j-1)/2)\omega)} & -1\\ 1 & 0\end{pmatrix}\right\|.\end{aligned} \tag{6.3}$$

Let $z = e^{2\pi ix}$, $w = e^{2\pi iy}$, $a = e^{\pi i\omega}$, as well as

$$A_\lambda(\lambda, E, z, w, a) := \begin{pmatrix}Ezw - \lambda z^2w^2a - \lambda\bar{a} & -zw\\ zw & 0\end{pmatrix}. \tag{6.4}$$

Then for $(z, w) \in \partial D_1 \times \partial D_1$,

$$\begin{aligned}u_n(\lambda, E; x, y) &= \frac{1}{n}\log\left\|\prod_{j=n}^{1}\begin{pmatrix}E - \lambda zw^j a^{j(j-1)} - \lambda z^{-1}w^{-j}\,\bar{a}^{j(j-1)} & -1\\ 1 & 0\end{pmatrix}\right\|\\ &= \frac{1}{n}\log\left\|\prod_{j=n}^{1}\begin{pmatrix}Ezw^j - \lambda z^2w^{2j}a^{j(j-1)} - \lambda\bar{a}^{j(j-1)} & -zw^j\\ zw^j & 0\end{pmatrix}\right\|\\ &= \frac{1}{n}\log\left\|\prod_{j=n}^{1}A_\lambda(E, z, w^j, a^{j(j-1)})\right\|\\ &=: v_n(\lambda, E; z, w).\end{aligned} \tag{6.5}$$

Note that $v_n(\lambda, E; z, w)$ is a plurisubharmonic function on $\mathbb{C}^2$. *Herman's regularization* refers to the transition from the first to the second line in (6.5), which removes the singularities z^{-1} and w^{-1}. Note that

$$v_n(\lambda, E; 0, w) = v_n(\lambda, E; z, 0) = \log \lambda. \tag{6.6}$$

For simplicity, we will write A instead of A_λ, since λ will be a fixed parameter within some range. As a general rule for the arguments of the matrix function A, the complex variables z, w will belong to some disk D_R, whereas $|a| = 1$ and E will be real-valued within some range. We will also keep $0 < \lambda < 1$.

6.2. *Explicit bounds on the monodromy matrices.* As a first step towards obtaining the explicit constants in Definitions 2.1 and 4.1 we prove the following bounds on v_n.

LEMMA 6.1. *Let* $0 < \lambda < 1$ *and* $R_3 \geq 1$. *Define*

$$U(\lambda, R_3) := \frac{1}{2} \log\left(\left(\lambda\left(1 + \frac{1}{R_3^2}\right)^2 + \frac{2}{R_3}\right)^2 + \frac{2}{R_3^2}\right). \tag{6.7}$$

Then for any E *with* $|E| \leq 2 + 2\lambda$, $v_n(\lambda, E, z, w)$ *from (6.5) satisfies the following estimates:*

- *for any* $w \in \partial D_1$,
$$v_n(\lambda, E; z, w) \leq 2 \log R_3 + U(\lambda, R_3) \quad \textit{for all } z \in \overline{D_{R_3}}, \quad \textit{and}$$
$$v_n(\lambda, E; 0, w) = \log \lambda; \tag{6.8}$$
- *for any* $(z, w) \in \partial D_1 \times \overline{D_{R_3}}$, *we have upper bound*
$$v_n(\lambda, E; z, w) \leq (n + 1) \log R_3 + U(\lambda, R_3), \tag{6.9}$$
and we also have $v_n(\lambda, E; z, 0) = \log \lambda$ *for any* $z \in \partial D_1$;
- *for any* $(z, w) \in \partial D_1 \times \partial D_1$,
$$|v_n(\lambda, E; z, w)| \leq U(\lambda, 1). \tag{6.10}$$

Remark 6.2. Let us note that

$$4U(\lambda, 1) \geq 2 \log 6 > 1. \tag{6.11}$$

Proof. Clearly (6.10) follows from (6.8) with $R_3 = 1$. We will use that for any complex-valued matrix

$$\|A\|^2 = \|A^* A\| \leq \mathrm{Tr}(A^* A).$$

For A as in (6.4) this means that

$$\begin{aligned}\|A(E, z, w, a)\|^2 &\leq |Ezw - \lambda z^2 w^2 a - \lambda \bar{a}|^2 + 2|z|^2|w|^2 \\ &\leq (|E||zw| + \lambda |zw|^2 + \lambda)^2 + 2|z|^2|w|^2,\end{aligned}$$

whence

$$\begin{aligned}|v_n(E; z, w)| &\leq \frac{1}{n} \sum_{j=1}^{2n} \log \|A(E, z, w^j, a^{j(j-1)})\| \\ &\leq \frac{1}{2n} \sum_{j=1}^{n} \log((\lambda(|z|^2|w|^{2j} + 1) + |E||z||w|^j)^2 + 2|z|^2|w|^{2j}).\end{aligned} \tag{6.12}$$

For $w \in \partial D_1$ and $|z| \le R_3$, (6.12) yields

$$
\begin{aligned}
|v_n(E; z, w)| &\le \frac{1}{2n} \sum_{j=1}^{n} \log((\lambda(R_3^2 + 1) + |E|R_3)^2 + 2R_3^2) \\
&\le 2 \log R_3 + \frac{1}{2} \log\bigg(\bigg(\lambda\bigg(1 + \frac{1}{R_3^2}\bigg) + \frac{2 + 2\lambda}{R_3}\bigg)^2 + \frac{2}{R_3^2}\bigg) \\
&= 2 \log R_3 + U(\lambda, R_3).
\end{aligned}
\tag{6.13}
$$

This proves (6.8).

Next, we turn to (6.9). For $z \in \partial D_1$ and $|w| \le R_3$, (6.12) yields

$$
\begin{aligned}
|v_n(E; z, w)| &\le \frac{1}{2n} \sum_{j=1}^{n} \log((\lambda(R_3^{2j} + 1) + |E|R_3^j)^2 + 2R_3^{2j}) \\
&\le (n + 1) \log R_3 + \frac{1}{2n} \sum_{j=1}^{n} \log\bigg(\bigg(\lambda\bigg(1 + \frac{1}{R_3^{2j}}\bigg) + (2 + 2\lambda)\frac{1}{R_3^j}\bigg)^2 + \frac{2}{R_3^{2j}}\bigg).
\end{aligned}
$$

Note that the summands are maximized at $j = 1$, which gives us the constant $2U(\lambda, R_3)$. Hence, in total

$$|v_n(E; z, w)| \le (n + 1) \log R_3 + U(\lambda, R_3),$$

as claimed. □

7. *Long sums of skew-shift functions*

In this section we establish a key large-deviation estimate on the ergodic averages of a plurisubharmonic function, as defined above, over a long skew-shift orbit. The argument is based on [**BouGolSch**, Lemma 2.6], but deviates from that reference in ways which are essential for our purposes. The precise dependence on all parameters is made explicit and effective. This leads to a somewhat cumbersome formulation which is, however, absolutely necessary for the main application. We wish to point out one technical feature of our version of this argument, namely that we only use a trivial bound on the number-theoretic divisor function; see the constant C^* below. We have found this to lead to the best constants. We also remark that significant gains in the following proposition would lead to dramatic improvements in the inductive machinery that we use to control the Lyapunov exponent; cf. the next two sections. At this point, however, it is not clear how to obtain such gains.

Recall that the constant $B_1(R, R_1, R_2)$ is as in (2.1), $B_3(R, R_1, R_2)$ is as in (2.2), and $B_4(R)$, $B_5(R)$, m_4, m_5 are as in (8.8). In the following we will write B_1, B_3, B_4, B_5 and omit the dependence on the radii.

PROPOSITION 7.1. *Let $\omega = (\sqrt{5} - 1)/2$ be the golden ratio. Let v be defined as in the beginning of §4, let $C = C(R, R_1, R_2)$ be the constant as in (3.47), and impose Definition 3.5. Let $\delta \in (0, 1/2)$ and $\delta_2, \delta_3 > 0$ be constants. Assume:*

(i) $C(B_5 - m_5) \le K^{\delta}$*;*

(ii) $K \ge 38$*;*

(iii) $\exp(4(\log K)^{\delta_2}) \ge K + 1$*;*

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

(iv) $21K^{-(9/10)+(9/5)\delta}(\log K)^{9/10+(9/5)\delta_2}+4C(B_4-m_4)\le K^{\delta}(\log K)^{\delta_2}$.

Then for any positive parameter $C_2>0$, we have

$$\left|\left\{(x,y)\in\mathbb{T}^2:\left|\frac{1}{K}\sum_{k=1}^{K}v\circ T_\omega^k(x,y)-\langle v\rangle\right|>\varepsilon_4\right\}\right|$$
$$\le 2\sqrt{2}\exp\left(\frac{\pi}{4}\left[\frac{17}{36}+\frac{B_1}{4B_3^2}-\varepsilon_5^{-1}\right]\right)$$
$$+\sqrt{2}(C(B_4-m_4))^{-1}K^{1/5-2\delta/5}(\log K)^{-1/5-2\delta_2/5}\exp(-2(\log K)^{\delta_2}),$$

where

$$\begin{aligned}\varepsilon_4&=C_2K^{-1/10+\delta/5}(\log K)^{1/10+\delta_2/5+\delta_3},\\ \varepsilon_5&=C_2^{-1}(472.5+3.2B_3(B_4-m_4)\sqrt{C})(\log K)^{-\delta_3}.\end{aligned}\tag{7.1}$$

Before proving Proposition 7.1, we will review some background of continued fractions.

7.1. *Continued fractions.* Each $\omega\in[0,1)$ has the unique expansion

$$\omega=\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+\frac{1}{\cdots}}}},\tag{7.2}$$

where $a_i\in\mathbb{N}_+$. We will denote this expansion by $\omega=[a_1,a_2,\ldots]$. If $\omega\in\mathbb{Q}$, the expansion is finite, while it is infinite for irrational ω.

Let $\omega\in[0,1)\backslash\mathbb{Q}$, and let

$$\frac{p_n}{q_n}=\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{\cdots+\frac{1}{a_n}}}}\tag{7.3}$$

be the continued fraction approximants of ω. These approximants satisfy the following three properties:

$$q_{n+1}=a_{n+1}q_n+q_{n-1}\quad\text{with } q_0:=1;\tag{7.4}$$
$$\|k\omega\|_{\mathbb{T}}\ge\|q_n\omega\|_{\mathbb{T}}\quad\text{for any } q_n\le|k|<q_{n+1};\tag{7.5}$$
$$\frac{1}{q_{n+1}+q_n}\le\|q_n\omega\|_{\mathbb{T}}\le\frac{1}{q_{n+1}}.\tag{7.6}$$

We are primarily interested in is the golden ratio $\omega=(\sqrt{5}-1)/2$. It is well known that ω has continued fraction expansion with $a_i\equiv1$ for any $i\ge1$. Then by (7.4), we have $q_{n+1}=q_n+q_{n-1}$ with $q_0=q_1=1$. Hence for any $n\ge0$, $q_{n+1}\le2q_n$. Then by (7.5) and (7.6), we have the following property of the golden ratio.

PROPOSITION 7.2. *The golden ratio satisfies*

$$\|k\omega\|_{\mathbb{T}}\ge\frac{1}{3|k|}\quad\textit{for any } k\ne0.\tag{7.7}$$

The optimal bound here is $((\sqrt{5}+1)/2)+$, but the constant 3 is sufficient. We will use the following corollary of (7.7) repeatedly.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

COROLLARY 7.3. *The golden ratio satisfies the following two properties.*

- *If ℓ is a positive integer such that $\|\ell\omega\|_{\mathbb{T}} \leq \sigma$, then $\ell \geq 1/(3\sigma)$.*
- *If $\ell, \tilde{\ell}$ are two distinct positive integers such that* $\max(\|\ell\omega\|_{\mathbb{T}}, \|\tilde{\ell}\omega\|_{\mathbb{T}}) \leq \sigma$, *then* $|\ell - \tilde{\ell}| \geq 1/(6\sigma)$.

In order to control small divisors we will rely on the following two lemmas.

LEMMA 7.4. *For $\theta \in \mathbb{R}$, we have*

$$\left|\sin\left(\frac{\pi}{2}\theta\right)\right| \geq \|\theta\|_{\mathbb{T}}. \tag{7.8}$$

Proof. If $\theta \in \mathbb{Z} + \frac{1}{2}$, $|\sin((\pi/2)\theta)| = \sin(\pi/4) > \frac{1}{2} = \|\theta\|_{\mathbb{T}}$.

If $\theta \notin \mathbb{Z} + \frac{1}{2}$, then there exists a unique $k \in \mathbb{Z}$ such that $\theta = k + \|\theta\|_{\mathbb{T}}$ (if $\theta \in [k, k + \frac{1}{2})$), or $\theta = k - \|\theta\|_{\mathbb{T}}$ (if $\theta \in (k - \frac{1}{2}, k)$). If k is an even number, then $|\sin((\pi/2)\theta)| = \sin((\pi/2)\|\theta\|_{\mathbb{T}}) \geq \|\theta\|_{\mathbb{T}}$, in which we used $\sin x \geq (2/\pi)x$ for $0 \leq x \leq \pi/2$. If k is odd, then $|\sin((\pi/2)\theta)| = \cos((\pi/2)\|\theta\|_{\mathbb{T}}) \geq \cos(\pi/4) > \frac{1}{2} \geq \|\theta\|_{\mathbb{T}}$. □

We will also use the following two estimates.

LEMMA 7.5. *For any positive integer R,*

$$\left|\sum_{k=1}^{R} e(k\ell\omega)\right| \leq \min\left(R, \frac{2}{2\sin(\pi\|\ell\omega\|_{\mathbb{T}})}\right) \leq \min\left(R, \frac{1}{2\|\ell\omega\|_{\mathbb{T}}}\right) \tag{7.9}$$

and

$$\left|\sum_{k=1}^{R} e\left(\frac{1}{2}k\ell\omega\right)\right| \leq \min\left(R, \frac{1}{\|\ell\omega\|_{\mathbb{T}}}\right). \tag{7.10}$$

Proof. For $\theta \notin \mathbb{Z}$, we have

$$\left|\sum_{k=1}^{R} e(k\theta)\right| = \min\left(R, \left|\frac{e(\theta) - e((k+1)\theta)}{1 - e(\theta)}\right|\right) \leq \min\left(R, \frac{1}{|\sin(\pi\theta)|}\right).$$

Then (7.9) follows from taking $\theta = \ell\omega$, and using $\sin(\pi x) \geq 2x$ for $0 \leq x \leq \frac{1}{2}$. Inequality (7.10) follows from taking $\theta = \frac{1}{2}\ell\omega$ and employing (7.8). □

7.2. *Proof of Proposition 7.1.* Let $\hat{v}(\ell, y)$ and $\hat{v}(x, \ell)$ denote the Fourier coefficients relative to the first and second variables, respectively, and by $\hat{v}(\ell_1, \ell_2)$ we mean the Fourier transform in both variables. For simplicity, let us omit the dependence of $C(R, R_1, R_2)$, $B_3(R, R_1, R_2)$ on the radii.

First, we note the following estimate as a corollary of Lemma 3.13.

COROLLARY 7.6. *For any $\ell \neq 0$, we have*

$$\begin{aligned} \sup_{x\in\mathbb{T}} |\hat{v}(x, \ell)| &\leq \frac{C}{|\ell|}(B_5 - m_5), \\ \sup_{y\in\mathbb{T}} |\hat{v}(\ell, y)| &\leq \frac{C}{|\ell|}(B_4 - m_4), \end{aligned} \tag{7.11}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

and

$$\left(\sum_{\ell_1\in\mathbb{Z}} |\hat{v}(\ell_1,\ell_2)|^2\right)^{1/2} \le \frac{C}{|\ell_2|}(B_5-m_5) \quad \textit{for any } \ell_2\neq 0,$$
$$\left(\sum_{\ell_2\in\mathbb{Z}} |\hat{v}(\ell_1,\ell_2)|^2\right)^{1/2} \le \frac{C}{|\ell_1|}(B_4-m_4) \quad \textit{for any } \ell_1\neq 0. \tag{7.12}$$

Proof. Note that (7.11) follows directly from Lemma 3.13. On the other hand,

$$\left(\sum_{\ell_1\in\mathbb{Z}} |\hat{v}(\ell_1,\ell_2)|^2\right)^{1/2} = \|\hat{v}(\cdot,\ell_2)\|_{L^2(\mathbb{T})} \le \sup_{x\in\mathbb{T}} |\hat{v}(x,\ell_2)|.$$

Hence, (7.12) reduces to (7.11). □

With some positive integer p_1 to be determined, let

$$\begin{aligned} v(x,y) &= \sum_{|\ell_1|\le p_1} \hat{v}(\ell_1,y)e(\ell_1 x) + \sum_{|\ell_1|>p_1} \hat{v}(\ell_1,y)e(\ell_1 x) \\ &=: v_1(x,y)+\tilde{v}_1(x,y), \end{aligned} \tag{7.13}$$

where v_1 and $\tilde{v}_1$ are the low- and high-frequency parts, respectively.

By Corollary 7.6,

$$\begin{aligned} \sup_{y\in\mathbb{T}} \|\tilde{v}_1(\cdot,y)\|_{L^1(\mathbb{T})} &\le \left(\sum_{|\ell_1|>p_1} \sup_{y\in\mathbb{T}} |\hat{v}(\ell_1,y)|^2\right)^{1/2} \\ &\le C_5\left(\sum_{|\ell_1|>p_1} \frac{1}{\ell_1^2}\right)^{1/2}(B_4-m_4) \\ &\le \sqrt{2}C(B_4-m_4)p_1^{-1/2}. \end{aligned} \tag{7.14}$$

Next, we further decompose v_1 into low- and high-frequency parts in the y variable. With some positive integer p_2 to be determined, let

$$\begin{aligned} v_1(x,y) &= \sum_{\substack{|\ell_1|\le p_1\\ |\ell_2|>p_2}} \hat{v}(\ell_1,\ell_2)e(\ell_1 x+\ell_2 y) + \sum_{\substack{|\ell_1|\le p_1\\ |\ell_2|\le p_2}} \hat{v}(\ell_1,\ell_2)e(\ell_1 x+\ell_2 y) \\ &=: v_2(x,y)+v_3(x,y). \end{aligned} \tag{7.15}$$

By Corollary 7.6, we have

$$\begin{aligned} \|v_2(x,y)\|_{L^1(\mathbb{T}^2)} &\le \|v_2(x,y)\|_{L^2(\mathbb{T}^2)} = \left(\sum_{|\ell_1|\le p_1, |\ell_2|>p_2} |\hat{v}(\ell_1,\ell_2)|^2\right)^{1/2} \\ &\le \left(\sum_{\ell_1\in\mathbb{Z},\ |\ell_2|>p_2} |\hat{v}(\ell_1,\ell_2)|^2\right)^{1/2} \le \left(\sum_{|\ell_2|>p_2} \frac{C^2}{\ell_2^2}(B_5-m_5)^2\right)^{1/2} \\ &< \sqrt{2}C(B_5-m_5)p_2^{-1/2}. \end{aligned} \tag{7.16}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Hence, by Markov's inequality,

$$\left|\left\{y \in \mathbb{T} : \frac{1}{K}\int_{\mathbb{T}}\left|\sum_{k=1}^{K} v_2 \circ T_\omega^k(x, y)\right| dx > t\right\}\right| \leq \sqrt{2}C(B_5 - m_5)p_2^{-1/2}t^{-1}. \tag{7.17}$$

We denote the set on the left-hand side of (7.17) by $\mathcal{A}(t)$.

Now let us consider v_3, which will lead to small-divisor problems. By Corollary 7.6 and the fact that $\hat{v}(0, 0) = \langle v\rangle$, separating the cases $\ell_1 = 0$, $\ell_2 = 0$, and $\ell_1\ell_2 \neq 0$ yields

$$\begin{aligned}
&\sup_{(x,y)\in\mathbb{T}^2}\left|\frac{1}{K}\sum_{k=1}^{K} v_3 \circ T_\omega^k(x, y) - \langle v\rangle\right| \\
&\leq \frac{1}{K}\sum_{\substack{|\ell_1|\leq p_1\\|\ell_2|\leq p_2\\|\ell_1|+|\ell_2|\neq 0}} |\hat{v}(\ell_1, \ell_2)|\left|\sum_{k=1}^{K} e\left(\ell_1\left(ky + \frac{k(k-1)\omega}{2}\right) + \ell_2 k\omega\right)\right| \\
&\leq \frac{C(B_5 - m_5)}{K}\sum_{1\leq|\ell_2|\leq p_2}\frac{1}{|\ell_2|}\left|\sum_{k=1}^{K} e(\ell_2 k\omega)\right| \\
&\quad + \frac{C(B_4 - m_4)}{K}\sum_{1\leq|\ell_1|\leq p_1}\frac{1}{|\ell_1|}\left|\sum_{k=1}^{K} e\left(\ell_1\left(ky + \frac{k(k-1)\omega}{2}\right)\right)\right| \\
&\quad + \frac{1}{K}\sum_{1\leq|\ell_2|\leq p_2}\sum_{1\leq|\ell_1|\leq p_1}|\hat{v}(\ell_1, \ell_2)|\left|\sum_{k=1}^{K} e\left(\ell_1\left(ky + \frac{k(k-1)\omega}{2}\right) + \ell_2 k\omega\right)\right|.
\end{aligned} \tag{7.18}$$

We now separately consider the sums appearing on the previous three lines. First,

$$S_1 := \frac{1}{K}\sum_{1\leq|\ell_2|\leq p_2}\frac{1}{|\ell_2|}\left|\sum_{k=1}^{K} e(\ell_2 k\omega)\right|.$$

Second, by Cauchy–Schwarz,

$$\begin{aligned}
&K^{-1}\sum_{1\leq|\ell_1|\leq p_1}\frac{1}{|\ell_1|}\left|\sum_{k=1}^{K} e\left(\ell_1\left(ky + \frac{k(k-1)\omega}{2}\right)\right)\right| \\
&\leq K^{-1}\left(\sum_{1\leq|\ell_1|\leq p_1}\frac{1}{\ell_1^2}\right)^{1/2}\left(\sum_{1\leq|\ell_1|\leq p_1}\left|\sum_{k=1}^{K} e\left(\ell_1\left(ky + \frac{k(k-1)\omega}{2}\right)\right)\right|^2\right)^{1/2} \\
&\leq 2K^{-1}\left(\sum_{1\leq|\ell_1|\leq p_1}\left|\sum_{k=1}^{K} e\left(\ell_1\left(ky + \frac{k(k-1)\omega}{2}\right)\right)\right|^2\right)^{1/2} =: S_2.
\end{aligned} \tag{7.19}$$

And, finally,

$$\begin{aligned}
&K^{-1} \sum_{1 \le |\ell_2| \le p_2} \sum_{1 \le |\ell_1| \le p_1} |\hat{v}(\ell_1, \ell_2)| \bigg| \sum_{k=1}^{K} e\bigg(\ell_1\bigg(ky + \frac{k(k-1)\omega}{2}\bigg) + \ell_2 k\omega\bigg)\bigg| \\
&\le K^{-1} \sum_{1 \le |\ell_2| \le p_2} \bigg(\sum_{1 \le |\ell_1| \le p_1} |\hat{v}(\ell_1, \ell_2)|^2 \bigg)^{1/2} \\
&\quad \times \bigg(\sum_{1 \le |\ell_1| \le p_1} \bigg| \sum_{k=1}^{K} e\bigg(\ell_1\bigg(ky + \frac{k(k-1)\omega}{2}\bigg) + \ell_2 k\omega\bigg)\bigg|^2 \bigg)^{1/2} \\
&\le C(B_5 - m_5) K^{-1} \sum_{1 \le |\ell_2| \le p_2} \frac{1}{|\ell_2|} \\
&\quad \times \bigg(\sum_{1 \le |\ell_1| \le p_1} \bigg| \sum_{k=1}^{K} e\bigg(\ell_1\bigg(ky + \frac{k(k-1)\omega}{2}\bigg) + \ell_2 k\omega\bigg)\bigg|^2 \bigg)^{1/2} \\
&=: C(B_5 - m_5) S_3.
\end{aligned}$$

Returning to (7.18), we conclude that

$$\begin{aligned}
&\sup_{(x,y) \in \mathbb{T}^2} \bigg| \frac{1}{K} \sum_{k=1}^{K} v_3 \circ T_\omega^k(x, y) - \langle v \rangle \bigg| \\
&\quad \le C(B_5 - m_5) S_1 + C(B_4 - m_4) S_2 + C(B_5 - m_5) S_3.
\end{aligned} \tag{7.20}$$

7.2.1. *Estimate of S_1.* Applying (7.9) to S_1, we infer that

$$\begin{aligned}
S_1 &< 2 \sum_{1 \le \ell_2 \le p_2} \frac{1}{\ell_2} \min\bigg(1, \frac{1}{2K \|\ell_2 \omega\|_{\mathbb{T}}}\bigg) \\
&= 2 \sum_{1 \le \ell_2 \le p_2} \mathbb{1}_{\{\ell_2 : \|\ell_2 \omega\|_{\mathbb{T}} \le 1/2K\}} \frac{1}{\ell_2} \\
&\quad + 2 \sum_{j=1}^{2^j \le 2K} \sum_{1 \le \ell_2 \le p_2} \mathbb{1}_{\{\ell_2 : 2^{j-1}/2K < \|\ell_2 \omega\|_{\mathbb{T}} \le 2^j/2K\}} \frac{1}{\ell_2} \min\bigg(1, \frac{1}{2K \|\ell_2 \omega\|_{\mathbb{T}}}\bigg) \\
&=: S_{1,1} + S_{1,2}.
\end{aligned} \tag{7.21}$$

Estimate of $S_{1,1}$.

$$S_{1,1} = 2 \sum_{1 \le \ell_2 \le p_2} \mathbb{1}_{\{\ell_2 : \|\ell_2 \omega\|_{\mathbb{T}} \le 1/2K\}} \frac{1}{\ell_2}. \tag{7.22}$$

By Corollary 7.3, if, for some $\ell \ge 1$, $\|\ell\omega\|_{\mathbb{T}} \le 1/2K$, then

$$\ell \ge \tfrac{2}{3} K. \tag{7.23}$$

Moreover, if, for some distinct $\ell, \tilde{\ell} \ge 1$, $\max(\|\ell\omega\|_{\mathbb{T}}, \|\tilde{\ell}\omega\|_{\mathbb{T}}) \le 1/2K$, then

$$|\ell - \tilde{\ell}| \ge \tfrac{1}{3} K. \tag{7.24}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

By (7.22), (7.23) and (7.24), we have

$$S_{1,1} \leq 2 \sum_{\ell=1}^{\lfloor 3p_2/K \rfloor - 1} \frac{3}{\ell+1} \frac{1}{K} < \frac{6}{K} \log \frac{3p_2}{K}. \tag{7.25}$$

For this estimate, and from this point on, we assume that $p_2 \geq K$.

Estimate of $S_{1,2}$.

$$\begin{aligned} S_{1,2} &\leq \frac{1}{K} \sum_{j=1}^{2^j \leq 2K} \sum_{1 \leq \ell_2 \leq p_2} \mathbb{1}_{\{\ell_2 : 2^{j-1}/2K < \|\ell_2\omega\|_{\mathbb{T}} \leq 2^j/2K\}} \frac{1}{\ell_2 \|\ell_2\omega\|_{\mathbb{T}}} \\ &\leq 2 \sum_{j=1}^{2^j \leq 2K} \sum_{1 \leq \ell_2 \leq p_2} \mathbb{1}_{\{\ell_2 : 2^{j-1}/2K < \|\ell_2\omega\|_{\mathbb{T}} \leq 2^j/2K\}} \frac{1}{2^{j-1}\ell_2}. \end{aligned} \tag{7.26}$$

By Corollary 7.3, if, for some $\ell \geq 1$, $\|\ell\omega\|_{\mathbb{T}} \leq 2^j/2K$, then we have

$$\ell \geq \frac{2}{2^j 3} K. \tag{7.27}$$

Moreover, if, for some distinct $\ell, \tilde{\ell} \geq 1$, max $(\|\ell\omega\|_{\mathbb{T}}, \|\tilde{\ell}\omega\|_{\mathbb{T}}) \leq 2^j/2K$, then

$$|\ell - \tilde{\ell}| \geq \frac{1}{2^j 3} K. \tag{7.28}$$

By (7.26), (7.27) and (7.28), we have

$$\begin{aligned} S_{1,2} &\leq 2 \sum_{j=1}^{2^j \leq 2K} \sum_{\ell=1}^{\lfloor 2^j 3p_2/K \rfloor - 1} \frac{6}{\ell+1} \frac{1}{K} \\ &\leq \frac{12}{K} \sum_{j=1}^{2^j \leq 2K} \log \frac{2^j 3p_2}{K} \\ &\leq \frac{12}{K} \left(\frac{\log 2}{2} (\log_2 K + 1)^2 + (\log_2 K + 1) \log \frac{3p_2}{K} \right). \end{aligned} \tag{7.29}$$

Putting (7.21), (7.25) and (7.29) together, we have

$$S_1 \leq \frac{6}{K} \left(\log 2 (\log_2 K + 2)^2 + (2 \log_2 K + 3) \log \frac{3p_2}{K} \right). \tag{7.30}$$

Henceforth we assume $p_2 \geq K > 23$, which allows us to simplify (7.30) into

$$S_1 \leq 26 \frac{\log K}{K} \log p_2. \tag{7.31}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

7.2.2. *Estimate of S_2 and S_3.* In order to estimate S_2 and S_3, we use the well-known method of Weyl differencing; cf., for example, [**Mon**]. As a first step,

$$\begin{aligned}
&\left|\sum_{k=1}^{K} e\left(\ell_1\left(ky+\frac{k(k-1)\omega}{2}\right)+\ell_2 k\omega\right)\right|^2\\
&\quad=\left(\sum_{k=1}^{K} e\left(\ell_1\left(ky+\frac{k(k-1)\omega}{2}\right)+\ell_2 k\omega\right)\right)\\
&\qquad\times\left(\sum_{j=1}^{K} e\left(-\ell_1\left(jy+\frac{j(j-1)\omega}{2}\right)-\ell_2 j\omega\right)\right)\\
&\quad=\sum_{j,k=1}^{K} e\left(\ell_1\left(\left(y-\frac{\omega}{2}\right)(k-j)+\frac{\omega}{2}(k^2-j^2)\right)+\ell_2\omega(k-j)\right).
\end{aligned}$$

Let $\ell=k+j$ and $m=k-j$, hence $\ell\equiv m \pmod 2$. Let us denote by $\mathbb{1}_e$ the indicator function of even numbers, and by $\mathbb{1}_o$ the indicator function of odd numbers; that is,

$$\mathbb{1}_e(j)=\mathbb{1}_o(j+1)=\begin{cases}1 & \text{if } j \text{ is even},\\ 0 & \text{if } j \text{ is odd}.\end{cases}$$

Then

$$\begin{aligned}
&\left|\sum_{k=1}^{K} e\left(\ell_1\left(ky+\frac{k(k-1)\omega}{2}\right)+\ell_2 k\omega\right)\right|^2\\
&\quad=\sum_{m=1-K}^{K-1}\sum_{\ell=2+|m|}^{2K-|m|}\mathbb{1}_e(m)\mathbb{1}_e(\ell)e\left(\ell_1 m\left(y+\frac{(\ell-1)\omega}{2}\right)+\ell_2 m\omega\right)\\
&\qquad+\sum_{m=1-K}^{K-1}\sum_{\ell=2+|m|}^{2K-|m|}\mathbb{1}_o(m)\mathbb{1}_o(\ell)e\left(\ell_1 m\left(y+\frac{(\ell-1)\omega}{2}\right)+\ell_2 m\omega\right),
\end{aligned}\tag{7.32}$$

in which, by (7.9), with $m=2\tilde m$ and $\ell=2\tilde\ell$,

$$\begin{aligned}
&\left|\sum_{m=1-K}^{K-1}\sum_{\ell=2+|m|}^{2K-|m|}\mathbb{1}_e(m)\mathbb{1}_e(\ell)e\left(\ell_1 m\left(y+\frac{(\ell-1)\omega}{2}\right)+\ell_2 m\omega\right)\right|\\
&\quad=\left|\sum_{\tilde m=-\lfloor\frac{K-1}{2}\rfloor}^{\lfloor(K-1)/2\rfloor}\sum_{\tilde\ell=1+|\tilde m|}^{K-|\tilde m|} e\left(\ell_1\tilde m\left(2y+(2\tilde\ell-1)\omega\right)+2\ell_2\tilde m\omega\right)\right|\\
&\quad\le K+2\sum_{\tilde m=1}^{\lfloor(K-1)/2\rfloor}\min\left(K-2\tilde m,\frac{1}{2\|2\ell_1\tilde m\omega\|_{\mathbb{T}}}\right)\\
&\quad\le K+2K\sum_{\tilde m=1}^{\lfloor(K-1)/2\rfloor}\min\left(1,\frac{1}{2K\|2\ell_1\tilde m\omega\|_{\mathbb{T}}}\right),
\end{aligned}\tag{7.33}$$

and with $m = 2\tilde{m} - 1$, $\ell = 2\tilde{\ell} - 1$,

$$\bigg|\sum_{m=1-K}^{K-1} \sum_{\ell=2+|m|}^{2K-|m|} \mathbb{1}_o(m)\mathbb{1}_o(\ell) e\bigg(\ell_1 m\bigg(y + \frac{(\ell-1)\omega}{2}\bigg) + \ell_2 m\omega\bigg)\bigg|$$
$$\leq 2\sum_{m=1}^{K-1}\bigg|\sum_{\ell=2+|m|}^{2K-|m|} \mathbb{1}_o(m)\mathbb{1}_o(\ell) e\bigg(\ell_1 m\bigg(y + \frac{(\ell-1)\omega}{2}\bigg) + \ell_2 m\omega\bigg)\bigg|, \tag{7.34}$$

which is further equal to

$$2\sum_{\tilde{m}=1}^{\lfloor K/2\rfloor}\bigg|\sum_{\tilde{\ell}=\tilde{m}+1}^{K-\tilde{m}+1} e(\ell_1(2\tilde{m}-1)(y+\tilde{\ell}\omega) + \ell_2(2\tilde{m}-1)\omega)\bigg|$$
$$\leq 2\sum_{\tilde{m}=1}^{\lfloor K/2\rfloor} \min\bigg(K - 2\tilde{m} + 1, \frac{1}{2\|2\ell_1\tilde{m}\omega\|_{\mathbb{T}}}\bigg)$$
$$\leq 2K\sum_{\tilde{m}=1}^{\lfloor K/2\rfloor} \min\bigg(1, \frac{1}{2K\|2\ell_1\tilde{m}\omega\|_{\mathbb{T}}}\bigg).$$

Plugging the estimates of (7.33) and (7.34) into (7.32) yields

$$\bigg|\sum_{k=1}^{K} e\bigg(\ell_1\bigg(ky + \frac{k(k-1)\omega}{2}\bigg) + \ell_2 k\omega\bigg)\bigg|^2 \leq K + 4K\sum_{m=1}^{\lfloor K/2\rfloor} \min\bigg(1, \frac{1}{2K\|2\ell_1\tilde{m}\omega\|_{\mathbb{T}}}\bigg).$$

Hence we have

$$S_2 \leq \frac{2}{K}\bigg(2p_1K + 4K\sum_{1\leq|\ell_1|\leq p_1}\sum_{m=1}^{\lfloor K/2\rfloor} \min\bigg(1, \frac{1}{2K\|2\ell_1 m\omega\|_{\mathbb{T}}}\bigg)\bigg)^{1/2}$$
$$\leq \frac{2}{K}\bigg(2p_1K + 8K\sum_{\ell_1=1}^{p_1}\sum_{m=1}^{\lfloor K/2\rfloor} \min\bigg(1, \frac{1}{2K\|2\ell_1 m\omega\|_{\mathbb{T}}}\bigg)\bigg)^{1/2}$$
$$\leq \frac{2}{K}\bigg(2p_1K + 8KC^*_{p_1K}\sum_{j=1}^{p_1K} \min\bigg(1, \frac{1}{2K\|j\omega\|_{\mathbb{T}}}\bigg)\bigg)^{1/2}, \tag{7.35}$$

and similarly,

$$S_3 \leq \frac{2}{K}(\log(p_2)+1)\bigg(2p_1K + 8KC^*_{p_1K}\sum_{j=1}^{p_1K} \min\bigg(1, \frac{1}{2K\|j\omega\|_{\mathbb{T}}}\bigg)\bigg)^{1/2}. \tag{7.36}$$

The constant $C^*_{p_1K}$ comes from over-counting.

Next, we will need to bound $C^*_{p_1K}$ and $\sum_{j=1}^{p_1K} \min(1, 1/(2K\|j\omega\|_{\mathbb{T}}))$ separately.

Estimate of $C^*_{p_1K}$. We first note the following simple bound on $C^*_{p_1K}$:

$$C^*_{p_1K} \leq \min(p_1, \tau^*(p_1K)), \tag{7.37}$$

where $\tau^*(p_1K) := \max_{1\leq n\leq p_1K} \tau(n)$ with $\tau(n)$ be the divisor function of n. Standard divisor bound yields the following estimates.

LEMMA 7.7.
$$\tau^*(p_1K) \le \begin{cases} (p_1K)^{1.066\,02/(\log\log(p_1K))}, \\ C(\epsilon)(p_1K)^{\epsilon}. \end{cases} \tag{7.38}$$

The second inequality above holds for any integer $p_1K \ge 1$ with explicit constants $C(\frac{1}{2}) = 2$ and $C(\frac{1}{8}) = 42\,000$. It also holds for $p_1K \le 327\,680\,000$ with constant $C(\frac{1}{50}) = 702$.

Combining (7.37) with Lemma 7.7, we have that for any $0 \le \alpha \le 1$,
$$C^*_{p_1K} \le (C(\epsilon))^{\alpha}(p_1K)^{\alpha\epsilon}p_1^{1-\alpha}. \tag{7.39}$$

We will only use the case where $\alpha = 0$, but we keep this as a reference for the sake of completeness.

Estimate of $\sum_{\ell=1}^{p_1K} \min(1, 1/(2K\|\ell\omega\|_{\mathbb{T}}))$.

In analogy with (7.21), we will split the term $\sum_{\ell=1}^{p_1K} \min(1, 1/(2K\|\ell\omega\|_{\mathbb{T}}))$ appearing in (7.36) as follows:
$$\begin{aligned}
&\sum_{\ell=1}^{p_1K} \min\left(1, \frac{1}{2K\|\ell\omega\|_{\mathbb{T}}}\right) \\
&\quad = \sum_{\ell=1}^{p_1K} \mathbb{1}_{\{\ell:\|\ell\omega\|_{\mathbb{T}}<1/2K\}} + \sum_{j=1}^{2^j<2K}\sum_{\ell=1}^{p_1K} \mathbb{1}_{\{\ell:2^{j-1}/2K\le\|\ell\omega\|_{\mathbb{T}}<\frac{2^j}{2K}\}}\frac{1}{2K\|\ell\omega\|_{\mathbb{T}}} \\
&\quad \le \sum_{\ell=1}^{p_1K} \mathbb{1}_{\{\ell:\|\ell\omega\|_{\mathbb{T}}<1/2K\}} + \sum_{j=1}^{2^j<2K}\sum_{\ell=1}^{p_1K} \mathbb{1}_{\{\ell:2^{j-1}/2K\le\|\ell\omega\|_{\mathbb{T}}<2^j/2K\}}\frac{1}{2^{j-1}} \\
&\quad =: S_4 + S_5.
\end{aligned} \tag{7.40}$$

By Corollary 7.3, if, for some $\ell \ge 1$, $\|\ell\omega\|_{\mathbb{T}} < 2^j/2K$, then
$$\ell \ge \frac{2}{2^j 3}K. \tag{7.41}$$

By Corollary 7.3, if, for some distinct $\ell, \tilde{\ell} \ge 1$, $\max(\|\ell\omega\|_{\mathbb{T}}, \|\tilde{\ell}\omega\|_{\mathbb{T}}) < 2^j/K$, then
$$|\ell - \tilde{\ell}| \ge \frac{1}{2^j 3}K. \tag{7.42}$$

Combining (7.41) with (7.42), we have
$$S_4 \le 3p_1 \tag{7.43}$$

and
$$S_5 \le \sum_{j=1}^{2^j<2K} 6p_1 \le 6(\log_2 K + 1)p_1. \tag{7.44}$$

In view of (7.35), (7.36), (7.40), (7.43) and (7.44) together, one has
$$\begin{aligned}
S_2 &< \frac{2}{K}(2p_1K + 24p_1K\, C^*_{p_1K}(2\log_2 K + 3))^{1/2} \\
&< 8\sqrt{3}(C^*_{p_1K})^{1/2}p_1^{1/2}K^{-1/2}(\log_2 K + 2)^{1/2} \\
&< 20(C^*_{p_1K})^{1/2}p_1^{1/2}K^{-1/2}(\log K)^{1/2} \quad \text{for } K \ge 23,
\end{aligned} \tag{7.45}$$

and

$$\begin{aligned} S_3 &< 8\sqrt{3}(C^*_{p_1 K})^{1/2}(\log\,(p_2)+1)p_1^{1/2}K^{-1/2}(\log_2 K+2)^{1/2} \\ &< 25(\log p_2)(C^*_{p_1 K})^{1/2}p_1^{1/2}K^{-1/2}(\log K)^{1/2} \quad \text{for } p_2 \geq K \geq 38. \end{aligned} \tag{7.46}$$

7.2.3. *Combining S_1, S_2, S_3.* Taking $p_1 = \lfloor K^{\delta_1} \rfloor$ and $p_2 = \lfloor e^{4(\log K)^{\delta_2}} \rfloor$, the estimate of S_1, namely (7.31), becomes

$$S_1 < 105K^{-1}(\log K)^{\delta_2+1}. \tag{7.47}$$

Recall from the foregoing that we impose the conditions $K \geq 38$ and $\exp(4(\log K)^{\delta_2}) \geq K+1$; note that these are our assumptions (ii) and (iii). The estimate of S_2, (7.45), becomes

$$S_2 < 20(C^*_{K^{1+\delta_1}})^{1/2}K^{-(1-\delta_1)/2}(\log K)^{1/2}. \tag{7.48}$$

The estimate of S_3, (7.46), becomes

$$S_3 < 100(C^*_{K^{1+\delta_1}})^{1/2}K^{-(1-\delta_1)/2}(\log K)^{\delta_2+1/2}. \tag{7.49}$$

Combining (7.47), (7.48), (7.49), (7.20) with our assumption (i) that $C(B_5 - m_5) \leq K^{\delta}$ yields

$$\begin{aligned} &\sup_{(x,y)\in\mathbb{T}^2}\left|\frac{1}{K}\sum_{k=1}^{K} v_3 \circ T^k_\omega(x,y) - \langle v\rangle\right| \\ &\quad < 105K^{-1+\delta}(\log K)^{\delta_2+1} + 20C(B_4-m_4)(C^*_{K^{1+\delta_1}})^{1/2}K^{-(1-\delta_1)/2}(\log K)^{1/2} \\ &\qquad + 100(C^*_{K^{1+\delta_1}})^{1/2}K^{-(1-\delta_1)/2+\delta}(\log K)^{\delta_2+1/2}. \end{aligned} \tag{7.50}$$

By (7.39), with $\alpha = 0$, we have

$$\begin{aligned} &\sup_{(x,y)\in\mathbb{T}^2}\left|\frac{1}{K}\sum_{k=1}^{K} v_3 \circ T^k_\omega(x,y) - \langle v\rangle\right| \\ &\quad < 105K^{-1+\delta}(\log K)^{\delta_2+1} + 20C(B_4-m_4)K^{-(1-2\delta_1)/2}(\log K)^{1/2} \\ &\qquad + 100K^{-(1-2\delta_1)/2+\delta}(\log K)^{\delta_2+1/2}. \end{aligned}$$

By condition (iv) in our statement of the proposition, we have

$$21K^{-1/2+\delta-\delta_1}(\log K)^{\delta_2+1/2} + 4C(B_4-m_4) \leq K^{\delta}(\log K)^{\delta_2},$$

which implies

$$\sup_{(x,y)\in\mathbb{T}^2}\left|\frac{1}{K}\sum_{k=1}^{K} v_3 \circ T^k_\omega(x,y) - \langle v\rangle\right| < 105K^{-(1-2\delta_1)/2+\delta}(\log K)^{\delta_2+1/2} =: \varepsilon_0. \tag{7.51}$$

Let

$$t = C(B_4-m_4)K^{-(1/2)\delta_1} \tag{7.52}$$

in (7.17). Then for any $y \notin \mathcal{A}(t)$, with (7.14) we have

$$\begin{aligned} \left\|\frac{1}{K}\sum_{k=1}^{K}(\tilde{v}_1+v_2)\circ T^k_\omega(\cdot,y)\right\|_{L^1(\mathbb{T})} &\leq \sqrt{2}C(B_4-m_4)K^{-(1/2)\delta_1} + t \\ &= (\sqrt{2}+1)C(B_4-m_4)K^{-(1/2)\delta_1} =: \varepsilon_1. \end{aligned} \tag{7.53}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Recall that $v = \tilde{v}_1 + v_2 + v_3$. For any fixed $y \notin \mathcal{A}(t)$, consider the subharmonic function

$$v_y(z) := \frac{1}{K}\sum_{k=1}^{K} v \circ T_\omega^k(z, y) \quad \text{with } z \in D_R.$$

This subharmonic function will satisfy the bounds

$$v_y(z) \leq B_4 \ \text{ for all } z \in D_R \quad \text{and} \quad v_y(0) \geq m_4.$$

By (7.51) and (7.53), we know $v_y(x) - \langle v \rangle$ can be decomposed into two parts, one with small L^∞ norm ε_0, the other with small L^1 norm ε_1. We will choose δ_1 such that $\varepsilon_0 \sim \sqrt{\varepsilon_1}$, in the sense that

$$K^{-(1-2\delta_1)/2+\delta}(\log K)^{\delta_2+1/2} = K^{-(1/4)\delta_1}, \tag{7.54}$$

which yields

$$K^{\delta_1} = \frac{K^{2/5-4\delta/5}}{(\log K)^{2/5+4\delta_2/5}} \quad \text{with } 0 < \delta_1 < \frac{2}{5} - \frac{4\delta}{5}. \tag{7.55}$$

Then

$$\varepsilon_0 \leq 105K^{-1/10+\delta/5}(\log K)^{1/10+\delta_2/5} \tag{7.56}$$

and

$$\varepsilon_1 = (\sqrt{2}+1)C(B_4 - m_4)K^{-1/5+2\delta/5}(\log K)^{1/5+2\delta_2/5}. \tag{7.57}$$

Applying Corollary 3.2 to $v_y - \langle v \rangle$, we obtain that for

$$\varepsilon_4 = C_2 K^{-1/10+\delta/5}(\log K)^{1/10+\delta_2/5+\delta_3} \tag{7.58}$$

with some constant $C_2 > 0$,

$$\begin{aligned}\sup_{y\notin\mathcal{A}(t)}&\left|\left\{x \in \mathbb{T} : \left|\frac{1}{K}\sum_{k=1}^{K} v \circ T_\omega^k(x, y) - \langle v \rangle\right| > \varepsilon_4\right\}\right|\\ &\leq 2\sqrt{2}\exp\left(\frac{\pi}{4}\left[\frac{17}{36} + \frac{B_1}{4B_3^2} - \varepsilon_2\delta_0^{-1}\right]\right),\end{aligned} \tag{7.59}$$

where

$$\delta_0 = \left(472.5 + 2B_3(B_4 - m_4)\sqrt{(\sqrt{2}+1)C}\right)K^{-1/10+\delta/5}(\log K)^{1/10+\delta_2/5},$$

and hence

$$\begin{aligned}\varepsilon_4\delta_0^{-1} &= C_2\left(472.5 + 2B_3(B_4 - m_4)\sqrt{(\sqrt{2}+1)C}\right)^{-1}(\log K)^{\delta_3}\\ &\geq C_2\left(472.5 + 3.2B_3(B_4 - m_4)\sqrt{C}\right)^{-1}(\log K)^{\delta_3} =: \varepsilon_5^{-1}.\end{aligned}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Inequality (7.59), together with (7.52) and (7.17), implies

$$\begin{aligned}
&\left|\left\{(x, y) \in \mathbb{T}^2 : \left|\frac{1}{K}\sum_{k=1}^{K} v \circ T_\omega^k(x, y) - \langle v\rangle\right| > \varepsilon_4\right\}\right| \\
&\leq 2\sqrt{2}\exp\left(\frac{\pi}{4}\left[\frac{17}{36} + \frac{B_1}{4B_3^2} - \varepsilon_5^{-1}\right]\right) + |\mathcal{A}(t)| \\
&\leq 2\sqrt{2}\exp\left(\frac{\pi}{4}\left[\frac{17}{36} + \frac{B_1}{4B_3^2} - \varepsilon_5^{-1}\right]\right) \\
&\quad + \sqrt{2}(C(B_4 - m_4))^{-1} K^{1/5 - 2\delta/5}(\log K)^{-1/5 - 2\delta_2/5} \exp\left(-2(\log K)^{\delta_2}\right),
\end{aligned}$$

as claimed.

8. *Multi-scale estimates*

In this section we commence with the inductive arguments in our multi-scale Lyapunov exponent machinery. In analogy with [**GolSch**, **Bou1**], we proceed by combining the large-deviation estimates with the avalanche principle. We begin with the basic induction step, which provides a lower bound for the Lyapunov exponent at a large scale from information on the Lyapunov exponents at smaller scales, in combination with level-set estimates. In Proposition 8.4, which is the main result of this section, we will also invoke the quantitative control on the Birkhoff averages over the skew shift from the previous section in order to derive large-deviation estimates at the larger scale.

The following subsection will serve as an abstract multi-scale scheme to provide a lower bound on the (maximal) Lyapunov exponent, assuming large-deviation estimates. In our application to the skew shift, the large-deviation estimates will come from Proposition 7.1; see §8.2.

8.1. *Abstract multi-scale scheme.*

8.1.1. *Lyapunov exponent.* Let (X, μ, S) be an ergodic dynamical system. A linear cocycle over (X, μ, S) is a skew-product map

$$F_A : X \times \mathbb{R}^d \to X \times \mathbb{R}^d,$$

given by

$$X \times \mathbb{R}^d \ni (x, v) \to (Sx, A(x)v) \in X \times \mathbb{R}^d,$$

where

$$A : X \to \mathrm{SL}_d(\mathbb{R})$$

is a measurable function.

The forward iterates F_A^n of a linear cocycle F_A are given by $F_A^n(x, v) = (S^n x, M_n(x)v)$, where

$$M_n(x) := A(S^{n-1}x) \cdots A(Sx)A(x), \quad n \in \mathbb{N}.$$

A linear cocycle A is said to be μ-integrable if

$$\int_X \log \|A(x)\|\, d\mu < +\infty.$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Due to the fact that norms are submultiplicative with respect to matrix products, the sequence of functions $\log \|A^{(n)}(x)\|$ are subadditive. The Fürstenberg–Kesten theorem (or Kingman's ergodic theorem) implies that for a μ-integrable linear cocycle, the μ-almost everywhere limit

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

exists and is called the (maximal) Lyapunov exponent of A. Moreover,

$$L(A) := \lim_{n\to\infty} \int_X \frac{1}{n} \log \|M_n(x)\| \mu(dx) = \inf_{n\geq 1} \int_X \frac{1}{n} \log \|M_n(x)\| \mu(dx).$$

We point out the since $A \in \mathrm{SL}_d(\mathbb{R})$, we have $\|M_n(x)\| \geq 1$, hence $L(A) \geq 0$.

8.1.2. *Inductive scheme.* Let us denote

$$L_n(A) := \int_X \frac{1}{n} \log \|M_n(x)\| \, d\mu(x).$$

For simplicity, we may omit the dependence of $L(A)$, $L_n(A)$ on A, and simply write L and L_n.

Let us further assume that there exists a constant $C_3 > 0$, such that

$$\frac{1}{n} \log \|M_n(x)\| \leq C_3 < +\infty, \tag{8.1}$$

for μ-almost every x, uniformly in n. We point out that in our application to the skew-shift model, C_3 can be taken as $U(\lambda, 1)$; see (6.5) and (6.10).

Definition 8.1. In our multi-scale scheme, we quantify the failure of the Fürstenberg–Kesten theorem via the sets

$$\mathcal{B}_n := \left\{ x \in X : \left| \frac{1}{n} \log \|M_n(x)\| - L_n \right| > \frac{1}{10} L_n \right\}.$$

The lemma below shows how to inductively obtain estimates of L_N at a larger scale N, based on information at a smaller scale n. The key ingredient is the avalanche principle, Theorem 5.5.

LEMMA 8.2. *Let $n, N/n \in \mathbb{N}$ be positive integers, and $\delta \in (0, 1/2)$. Let C_3 be as in (8.1). Assume the following three conditions:*

(a) $nL_n \geq 7$;

(b) $L_n - L_{2n} \leq \frac{1}{8} L_n$;

(c) $\max(\mu(\mathcal{B}_n), \mu(\mathcal{B}_{2n})) \leq N^{-12/5+4\delta/5}$.

Then we have

$$L_N \geq L_n - \left(2 - \frac{2n}{N}\right)(L_n - L_{2n}) - \frac{11}{n} e^{-(1/2)nL_n} - 8C_3 N^{-7/5+4\delta/5} \tag{8.2}$$

and

$$L_N - L_{2N} \leq \frac{n}{N}(L_n - L_{2n}) + \frac{22}{n} e^{-(1/2)nL_n} + 24C_3 N^{-7/5+4\delta/5}. \tag{8.3}$$

8.1.3. *Multi-scale scheme.* The following lemma shows how information on a sequence of larger and larger scales determines the limit L.

LEMMA 8.3. *Let $\delta \in (0, 1/2)$ be a constant, and C_3 be as in (8.1). Let $\{N_m\}_{m=0}^{\infty} \in \mathbb{N}$ be a sequence of positive integers, such that $10 \le N_m/N_{m-1} \in \mathbb{N}$ for $1 \le m$. Assume that the following hold for an integer $j \ge 0$ (note that (2)–(4) below are empty conditions for $j = 0$):*
(1) $N_0 L_{N_0} \ge 7$ *and* $L_{N_0} - L_{2N_0} \le \frac{1}{8} L_{N_0}$*;*
(2) $\sum_{m=0}^{j-1} (1/N_m) e^{-(1/2)N_m L_{N_m}} < \frac{1}{512} L_{N_0}$*;*
(3) $\sum_{m=1}^{j} N_m^{-7/5+4\delta/5} < (1/1280 C_3) L_{N_0}$*;*
(4) $\max\,(\mu(\mathcal{B}_{N_m}), \mu(\mathcal{B}_{2N_m})) \le N_{m+1}^{-12/5+4\delta/5}$*, for $0 \le m \le j-1$.*
Then we have the following four estimates for $j \ge 0$.

First,

$$\begin{aligned} L_{N_j} \ge L_{N_0} &- \left(2 - \frac{2N_0}{N_j}\right)(L_{N_0} - L_{2N_0}) \\ &- \sum_{m=1}^{j} \left(\frac{11}{N_{m-1}} e^{-(1/2)N_{m-1}L_{N_{m-1}}} + 8C_3 N_m^{-7/5+4\delta/5}\right) \\ &- \sum_{m=1}^{j-1} \left(2 - \frac{2N_m}{N_j}\right)\left(\frac{22}{N_{m-1}} e^{-(1/2)N_{m-1}L_{N_{m-1}}} + 24C_3 N_m^{-7/5+4\delta/5}\right), \end{aligned} \tag{8.4}$$

in which $\sum_{m=1}^{0} = \sum_{m=1}^{-1} :\equiv 0$.

Second,

$$\begin{aligned} L_{N_j} - L_{2N_j} \le &\frac{N_0}{N_j}(L_{N_0} - L_{2N_0}) \\ &+ \sum_{m=1}^{j} \frac{N_m}{N_j}\left(\frac{22}{N_{m-1}} e^{-(1/2)N_{m-1}L_{N_{m-1}}} + 24C_3 N_m^{-7/5+4\delta/5}\right), \end{aligned} \tag{8.5}$$

in which $\sum_{m=1}^{0} :\equiv 0$.

Third,

$$L_{N_j} - L_{2N_j} \le \tfrac{1}{8} L_{N_j}. \tag{8.6}$$

Fourth,

$$L_{2N_j} \ge \tfrac{1}{2} L_{N_0} \quad \textit{and} \quad N_j L_{N_j} \ge 7. \tag{8.7}$$

8.2. *Application to the skew-shift model.* The two cornerstones of the abstract multi-scale scheme are:

- initial scale N_0 estimates, including (1) $N_0 L_{N_0} \ge 7$, (2) $L_{N_0} - L_{2N_0} \le \frac{1}{8} L_{N_0}$ as well as (3) large-deviation estimates of $\mu(\mathcal{B}_{N_0})$ and $\mu(\mathcal{B}_{2N_0})$;
- large-deviation estimates of $\mu(\mathcal{B}_{N_j})$ and $\mu(\mathcal{B}_{2N_j})$ for $j \ge 1$.

In this subsection, we will present machinery that inductively provides large-deviation estimates for scales N_j, $j \ge 1$, thus reducing the problem to the initial scale only. The key ingredients are the avalanche principle and the quantitative control of the ergodic averages of plurisubharmonic functions over a skew-shift orbit, Proposition 7.1.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Let us recall some notation from §6. We have $u_n(\lambda, E; x, y) = (1/n)\log\|M_n(\lambda, E; x, y)\|$, and $v_n(\lambda, E; z, w)$ is the complexification of u_n from $\mathbb{T}^2$ to $\mathbb{C}^2$, as in (6.5). The constant $U(\lambda, 1)$, as in (6.10), is a uniform (in n and E) L^∞ upper bound on $u_n(\lambda, E; x, y)$. For simplicity, we will omit the dependence of $u_n(x, y)$, $v_n(z, w)$, $\log\|M_n(x, y)\|$ and L_n on λ, E, since λ will be fixed, and our estimates are uniform in $E \in [-2-2\lambda, 2+2\lambda]$. Recall also from Lemma 6.1 with $R_3 = R$ that the bounds with respect to v_n satisfy

$$\begin{aligned} B_4 - m_4 &= 2\log R + U(\lambda, R) - \log\lambda, \\ B_5^{(n)} - m_5^{(n)} &= (n+1)\log R + U(\lambda, R) - \log\lambda. \end{aligned} \tag{8.8}$$

Let us finally also recall the constants $B_3(R, R_1, R_2)$ as in (2.2), $C(R, R_1, R_2)$ as in (3.47), and $C_0(R, R_1, R_2)$ is as in (3.2). In the following we will write B_3, C, C_0 for simplicity.

PROPOSITION 8.4. *Let $\omega = (\sqrt{5}-1)/2$ be the golden ratio. Let $\delta \in (0, 1/2)$ and*

$$\delta_2, \delta_3, \delta_4, C_2, C_4, C_5 > 0$$

be constants. Let $n, N \in \mathbb{N}$ be positive integers and assume that n divides N. In addition to conditions (a)–(c) in Lemma 8.2 and Definition 3.5, assume further that the following properties hold for both $\tilde{N} = N$ and $2N$:

(I) $C((2n+1)\log R + U(\lambda, R) - \log\lambda) \le \tilde{N}^{\delta}$;
(II) $\tilde{N} \ge 38$;
(III) $\exp(4(\log\tilde{N})^{\delta_2}) \ge \tilde{N} + 1$;
(IV) $21\tilde{N}^{-9/10+9/5\delta}(\log\tilde{N})^{9/10+9/5\delta_2} + 4C(B_4 - m_4) \le \tilde{N}^{\delta}(\log\tilde{N})^{\delta_2}$;
(V) $2n\tilde{N}^{-1}(L_n - L_{2n}) + 8U(\lambda, 1)\tilde{N}^{-7/5+4\delta/5} + 5U(\lambda, 1)n\tilde{N}^{-1} < C_2\tilde{N}^{-1/10+\delta/5}(\log\tilde{N})^{1/10+\delta_2/5+\delta_3}$;
(VI) $22n^{-1}\exp(-nL_n/2) < C_2\tilde{N}^{-1/10+\delta/5}(\log\tilde{N})^{1/10+\delta_2/5+\delta_3}$;
(VII) $4\sqrt{2}\exp(\pi/4[\frac{17}{36} + B_1/(4B_3^2) - C_2(472.5 + 3.2B_3(B_4 - m_4)\sqrt{C})^{-1}(\log\tilde{N})^{\delta_3}]) \le \tilde{N}^{-7/5+4\delta/5}$;
(VIII) $2\sqrt{2}(C(B_4 - m_4))^{-1}\tilde{N}^{1/5-2\delta/5}(\log\tilde{N})^{-1/5-2\delta_2/5}\exp(-2(\log\tilde{N})^{\delta_2}) \le \tilde{N}^{-7/5+4\delta/5}$;
(IX) $\tilde{N} > (\log R + U(\lambda, R) - \log\lambda)(\log R)^{-1}$;
(X) $C_4(\log\tilde{N})^{\delta_4} > 4$;
(XI) $C_5(\log\tilde{N})^{\delta_4} > C_4$.

Then the following holds for both $\tilde{N} = N$ and $2N$:

$$\begin{aligned} &|\{(x, y) \in \mathbb{T}^2 : |v_{\tilde{N}}(x, y) - L_{\tilde{N}}| > C_2C_5\tilde{N}^{-1/10+\delta/5}(\log\tilde{N})^{1/10+\delta_2/5+\delta_3+2\delta_4}\}| \\ &\quad\le 2(2C_0)^{1/2}\exp\left(\frac{-\pi C_2C_4(\log\tilde{N})^{\delta_4}}{144C_2 + 48B_3\sqrt{2U(\lambda, 1)(B_4 - m_4)}(\log\tilde{N})^{-1/10-\delta_2/5-\delta_3}}\right) \\ &\qquad + C_0\exp\left(\frac{-\pi C_2C_5(\log\tilde{N})^{\delta_4}}{18C_2C_4 + 96B_3U(\lambda, 1)\sqrt{\log R}(\log\tilde{N})^{-1/10-\delta_2/5-\delta_3-\delta_4}}\right). \end{aligned}$$

Remark 8.5. Note that our conditions (I)–(IV) correspond to (i)–(iv) of Proposition 7.1. In particular, (I) is (i) of Proposition 7.1 with $B_5^{(2n)} - m_5^{(2n)}$ given in (8.8).

8.3. *Proofs.* Before proving Lemma 8.2 and Proposition 8.4, we will first give a quick proof of Lemma 8.3 based on Lemma 8.2.

Proof of Lemma 8.3. For $m \geq 1$, let us denote

$$\alpha_m := 11 N_{m-1}^{-1} e^{-(1/2) N_{m-1} L_{N_{m-1}}} + 8 C_3 N_m^{-7/5+4\delta/5},$$
$$\beta_m := 22 N_{m-1}^{-1} e^{-(1/2) N_{m-1} L_{N_{m-1}}} + 24 C_3 N_m^{-7/5+4\delta/5}.$$

Note that in terms of α and β, our conditions (2) and (3) in the statement of the lemma become

$$\sum_{m=1}^{j} \alpha_m \leq \frac{11}{512} L_{N_0} + \frac{8}{1280} L_{N_0} = \frac{71}{2560} L_{N_0} \tag{8.9}$$

and

$$\sum_{m=1}^{j} \beta_m \leq \frac{22}{512} L_{N_0} + \frac{24}{1280} L_{N_0} = \frac{79}{1280} L_{N_0}. \tag{8.10}$$

Our proof is based on induction on j. Note that, for the induction base case $j = 0$, inequalities (8.5), (8.6) and (8.7) follow directly from condition (1). Inequality (8.4) follows from the fact that $L_{N_0} - L_{2N_0} \geq 0$.

Now let us suppose Lemma 8.3 holds for $j = J$ for some $J \geq 0$. Note that conditions (2)–(4) with $j = J + 1$ already imply those with $j = J$. Hence, by our inductive assumption, (8.4)–(8.7) hold for $j = J$, whence

$$L_{N_J} \geq L_{N_0} - \left(2 - \frac{2N_0}{N_J}\right)(L_{N_0} - L_{2N_0}) - \sum_{m=1}^{J} \alpha_m - \sum_{m=1}^{J-1} \left(2 - \frac{2N_m}{N_J}\right) \beta_m, \tag{8.11}$$

$$L_{N_J} - L_{2N_J} \leq \frac{N_0}{N_J}(L_{N_0} - L_{2N_0}) + \sum_{m=1}^{J} \frac{N_m}{N_J} \beta_m, \tag{8.12}$$

$$L_{N_J} - L_{2N_J} \leq \tfrac{1}{8} L_{N_J} \tag{8.13}$$

and

$$N_J L_{N_J} \geq 7. \tag{8.14}$$

Note that (8.13), (8.14) and our condition (4) in the statement of the lemma with $m = J$ verify the conditions of Lemma 8.2 for $n = N_J$ and $N = N_{J+1}$. Therefore Lemma 8.2 implies

$$L_{N_{J+1}} \geq L_{N_J} - \left(2 - \frac{2N_J}{N_{J+1}}\right)(L_{N_J} - L_{2N_J}) - \alpha_{J+1} \tag{8.15}$$

and

$$L_{N_{J+1}} - L_{2N_{J+1}} \leq \frac{N_J}{N_{J+1}}(L_{N_J} - L_{2N_J}) + \beta_{J+1}. \tag{8.16}$$

Plugging (8.12) into (8.16), we obtain

$$\begin{aligned} L_{N_{J+1}} - L_{2N_{J+1}} &\leq \frac{N_0}{N_{J+1}}(L_{N_0} - L_{2N_0}) + \frac{N_J}{N_{J+1}}\left(\sum_{m=1}^{J} \frac{N_m}{N_J} \beta_m\right) + \beta_{J+1} \\ &= \frac{N_0}{N_{J+1}}(L_{N_0} - L_{2N_0}) + \sum_{m=1}^{J+1} \frac{N_m}{N_{J+1}} \beta_m. \end{aligned} \tag{8.17}$$

This proves (8.5) for $j = J + 1$.

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Plugging (8.11) and (8.12) with $j = J+1$ into (8.15), we have

$$\begin{aligned} L_{N_{J+1}} &\geq L_{N_0} - \left(2 - \frac{2N_0}{N_J}\right)(L_{N_0} - L_{2N_0}) - \sum_{m=1}^{J} \alpha_m - \sum_{m=1}^{J-1}\left(2 - \frac{2N_m}{N_J}\right)\beta_m \\ &\quad - \left(2 - \frac{2N_J}{N_{J+1}}\right)\left(\frac{N_0}{N_J}(L_{N_0} - 2L_{2N_0}) + \sum_{m=1}^{J} \frac{N_m}{N_J}\beta_m\right) - \alpha_{J+1} \\ &= L_{N_0} - \left(2 - \frac{2N_0}{N_{J+1}}\right)(L_{N_0} - L_{2N_0}) - \sum_{m=1}^{J+1} \alpha_m - \sum_{m=1}^{J}\left(2 - \frac{2N_m}{N_{J+1}}\right)\beta_m. \end{aligned} \tag{8.18}$$

This proves (8.4) for $j = J+1$.

Combining (8.17), (8.18) with the fact that $10 \leq N_{j+1}/N_j$ for any $j \geq 0$ yields

$$\begin{aligned} &8(L_{N_{J+1}} - L_{2N_{J+1}}) - L_{N_{J+1}} \\ &\quad \leq -L_{N_0} + \left(2 + \frac{6N_0}{N_{J+1}}\right)(L_{N_0} - L_{2N_0}) + \sum_{m=1}^{J+1} \alpha_m + \sum_{m=1}^{J+1}\left(2 + \frac{6N_m}{N_{J+1}}\right)\beta_m \\ &\quad \leq -L_{N_0} + \left(2 + \frac{6}{10}\right)(L_{N_0} - L_{2N_0}) + \sum_{m=1}^{J+1} \alpha_m + 8\sum_{m=1}^{J+1} \beta_m. \end{aligned}$$

Using (8.9), (8.10) and the fact that $L_{N_0} - L_{2N_0} \leq \frac{1}{8}L_{N_0}$, we conclude that

$$8(L_{N_{J+1}} - L_{2N_{J+1}}) - L_{N_{J+1}} \leq -\tfrac{393}{2560}L_{N_0} < 0. \tag{8.19}$$

This proves (8.6) for $j = J+1$.

By (8.18) and the fact that $N_{m+1} \geq 10N_m$ for any $m \geq 0$, we have

$$L_{N_{J+1}} \geq L_{N_0} - \left(2 - \frac{1}{5}\right)(L_{N_0} - L_{2N_0}) - \sum_{m=1}^{J+1} \alpha_m - \left(2 - \frac{1}{5}\right)\sum_{m=1}^{J} \beta_m. \tag{8.20}$$

Plugging (8.9) and (8.10) with $j = J+1$ into (8.20), and using that $L_{N_0} - L_{2N_0} \leq \frac{1}{8}L_{N_0}$, yields

$$L_{N_{J+1}} \geq \tfrac{8143}{12\,800}L_{N_0}, \tag{8.21}$$

which also implies $N_J L_{N_J} \geq 7$. Inequality (8.7) with $j = J+1$ then follows from (8.19) and (8.21), indeed,

$$L_{2N_{J+1}} \geq \tfrac{7}{8}L_{N_{J+1}} \geq \tfrac{7}{8} \times \tfrac{8143}{12\,800}L_{N_0} \geq \tfrac{1}{2}L_{N_0},$$

as desired. □

Proof of Lemma 8.2. Let $\tilde{N} = N$ or $2N$. Let us define

$$\mathcal{B}^{(\tilde{N})} := \left(\bigcup_{j=0}^{\tilde{N}-1} S^{-j}\mathcal{B}_n\right) \cup \left(\bigcup_{j=0}^{\tilde{N}-n-1} S^{-j}\mathcal{B}_{2n}\right). \tag{8.22}$$

We have the following measure estimates for $\mathcal{B}^{(\tilde{N})}$ by condition (c) of the statement of the lemma:

$$\mu(\mathcal{B}^{(N)}) \leq \frac{2N-n}{N^{12/5-4\delta/5}} \leq 2N^{-7/5+4\delta/5} \quad \text{and} \quad \mu(\mathcal{B}^{(2N)}) \leq 4N^{-7/5+4\delta/5}. \tag{8.23}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Taking any $x \notin \mathcal{B}^{(\tilde{N})}$, by our definitions of $\mathcal{B}_n$ and $\mathcal{B}_{2n}$, we have

$$e^{(11/10)nL_n} \geq \|M_n(S^j x)\| \geq e^{(9/10)nL_n} =: \kappa^{-1/2} \quad \text{for any } 0 \leq j \leq \tilde{N} - 1, \tag{8.24}$$

and

$$\|M_{2n}(S^j x)\| \geq e^{(9/5)nL_{2n}} \quad \text{for any } 0 \leq j \leq \tilde{N} - n - 1. \tag{8.25}$$

Hence, for any $0 \leq j \leq \tilde{N} - n - 1$, (8.24) and (8.25) imply that

$$\frac{\|M_{2n}(S^j x)\|}{\|M_n(S^{j+n}x)\| \, \|M_n(S^j x)\|} \geq \exp\left(2n(L_{2n} - L_n) - \frac{1}{5}n(L_n + L_{2n})\right) =: \epsilon.$$

We now need to verify the assumptions of the avalanche principle, Theorem 5.5. First, by subadditivity of $\log \|M_n(x)\|$, we have $L_{2n} \leq L_n$. This, together with our assumptions (a) and (b), yields

$$\epsilon = \exp\left(\tfrac{9}{5}nL_{2n} - \tfrac{11}{5}nL_n\right) \leq e^{-(2/5)nL_n} \leq e^{-14/5} < \tfrac{1}{10}, \tag{8.26}$$

and

$$\kappa\epsilon^{-2} = \exp\left(-2nL_n + \tfrac{3}{5}nL_n + \tfrac{2}{5}nL_{2n} + 4n(L_n - L_{2n})\right) \leq e^{-(1/2)nL_n} \leq e^{-7/2} < \tfrac{1}{10}.$$

Applying Theorem 5.5 to $x \notin \mathcal{B}^{(\tilde{N})}$, we conclude that for each $0 \leq k \leq n-1$,

$$\frac{1}{\tilde{N}}\left|\log \|M_{\tilde{N}}(S^k x)\| + \sum_{j=1}^{(\tilde{N}-2n)/n} \log \|M_n(S^{jn+k}x)\| \right.$$
$$\left. - \sum_{j=0}^{(\tilde{N}-2n)/n} \log \|M_{2n}(S^{jn+k}x)\|\right| \leq \frac{11}{n}\kappa\epsilon^{-2} \leq \frac{11}{n}e^{-(1/2)nL_n}.$$

Summing over $k \in [0, n-1]$ and dividing by n, and finally applying the triangle inequality yields

$$\left|\frac{1}{n}\sum_{k=0}^{n-1} \frac{1}{\tilde{N}} \log \|M_{\tilde{N}}(S^k x)\| + \frac{1}{\tilde{N}} \sum_{j=n}^{\tilde{N}-n-1} \frac{1}{n} \log \|M_n(S^j x)\| \right.$$
$$\left. - \frac{2}{\tilde{N}} \sum_{j=0}^{\tilde{N}-n-1} \frac{1}{2n} \log \|M_{2n}(S^j x)\|\right| \leq \frac{11}{n}e^{-(1/2)nL_n}. \tag{8.27}$$

Integrating over $x \in X$, and using our definition of C_3 (8.1), we infer due to (8.23) that

$$\left|L_N + \frac{N-2n}{N}L_n - 2\frac{N-n}{N}L_{2n}\right| \leq \frac{11}{n}e^{-(1/2)nL_n} + 4C_3\mu(\mathcal{B}^{(N)})$$
$$< \frac{11}{n}e^{-(1/2)nL_n} + 8C_3N^{-7/5+4\delta/5} \tag{8.28}$$

and

$$\left|L_{2N} + \frac{N-n}{N}L_n - \frac{2N-n}{N}L_{2n}\right| \leq \frac{11}{n}e^{-(1/2)nL_n} + 4C_3\mu(\mathcal{B}^{(2N)})$$
$$< \frac{11}{n}e^{-(1/2)nL_n} + 16C_3N^{-7/5+4\delta/5}. \tag{8.29}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

From (8.28), we conclude that

$$L_N \geq L_n - \left(2 - \frac{2n}{N}\right)(L_n - L_{2n}) - \frac{11}{n}e^{-(1/2)nL_n} - 8C_3 N^{-7/5+4\delta/5}.$$

This proves (8.2).

Taking the difference between (8.28) and (8.29), we obtain

$$L_N - L_{2N} \leq \frac{22}{n}e^{-(1/2)nL_n} + 24C_3 N^{-7/5+4\delta/5} + \frac{n}{N}(L_n - L_{2n}).$$

This proves (8.3). □

Proof of Proposition 8.4. This will be a continuation of the proof of Lemma 8.2. Note that all the constants C_3 will be replaced by $U(\lambda, 1)$. Let $\tilde{N}$ be either N or $2N$. Let us consider the first term in (8.27):

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{\tilde{N}}\log\|M_{\tilde{N}}(T_\omega^k(x,y))\| - \frac{1}{\tilde{N}}\log\|M_{\tilde{N}}(x,y)\|\right| \\
&\quad\leq \frac{1}{n\tilde{N}}\sum_{k=0}^{n-1}|\log\|M_{\tilde{N}}(T_\omega^k(x,y))\| - \log\|M_{\tilde{N}}(x,y)\|| \\
&\quad\leq \frac{1}{n\tilde{N}}\sum_{k=0}^{n-1}(\log\|M_k(x,y) + \log\|M_k(T_\omega^{\tilde{N}}(x,y))\|) \\
&\quad\leq \frac{1}{n\tilde{N}}\sum_{k=0}^{n-1}2kU(\lambda,1) \\
&\quad\leq \frac{U(\lambda,1)n}{\tilde{N}}.
\end{aligned} \tag{8.30}$$

Hence (8.27) leads to

$$\begin{aligned}
&\left|\frac{1}{\tilde{N}}\log\|M_{\tilde{N}}(x,y)\| + \frac{1}{\tilde{N}}\sum_{j=n}^{\tilde{N}-n-1}\frac{1}{n}\log\|M_n(T_\omega^j(x,y))\|\right. \\
&\quad\left.-\frac{2}{\tilde{N}}\sum_{j=0}^{\tilde{N}-n-1}\frac{1}{2n}\log\|M_{2n}(T_\omega^j(x,y))\|\right| \leq \frac{11}{n}e^{-(1/2)nL_n} + \frac{U(\lambda,1)n}{\tilde{N}},
\end{aligned} \tag{8.31}$$

which holds for $(x, y) \notin \mathcal{B}^{(\tilde{N})}$. This implies

$$\begin{aligned}
&\left|\frac{1}{\tilde{N}}\log\|M_{\tilde{N}}(x,y)\| + \frac{1}{\tilde{N}}\sum_{j=0}^{\tilde{N}-1}\frac{1}{n}\log\|M_n(T_\omega^j(x,y))\|\right. \\
&\quad\left.-\frac{2}{\tilde{N}}\sum_{j=0}^{\tilde{N}-1}\frac{1}{2n}\log\|M_{2n}(T_\omega^j(x,y))\|\right| \leq \frac{11}{n}e^{-(1/2)nL_n} + \frac{5U(\lambda,1)n}{\tilde{N}}.
\end{aligned} \tag{8.32}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

Now we apply Proposition 7.1 to v_n and v_{2n} with $K = \tilde{N}$. Note that conditions (I)–(IV) ensure the applicability of that proposition. Therefore, following (7.1), we define

$$\begin{aligned}\varepsilon_4 &= C_2 \tilde{N}^{-1/10+\delta/5} (\log \tilde{N})^{1/10+\delta_2/5+\delta_3},\\ \varepsilon_5 &= C_2^{-1} \Big(472.5 + 3.2 B_3 (B_4 - m_4)\sqrt{C}\Big) (\log \tilde{N})^{-\delta_3}.\end{aligned} \tag{8.33}$$

For $\tilde{n} = n$ or $2n$, denote

$$\mathcal{C}_{\tilde{n}} := \bigg\{ (x, y) \in \mathbb{T}^2 : \bigg| \frac{1}{\tilde{N}} \sum_{j=0}^{\tilde{N}-1} v_{\tilde{n}} \circ T_\omega^j (x, y) - L_{\tilde{n}} \bigg| > \varepsilon_4 \bigg\}.$$

Then Proposition 7.1 implies that

$$\begin{aligned}\max\,(|\mathcal{C}_n|, |\mathcal{C}_{2n}|) &\le 2\sqrt{2} \exp\bigg(\frac{\pi}{4} \bigg[\frac{17}{36} + \frac{B_1}{4B_3^2} - \varepsilon_5^{-1} \bigg] \bigg)\\ &\quad + \sqrt{2} (C(B_4 - m_4))^{-1} \tilde{N}^{1/5-2\delta/5} (\log \tilde{N})^{-1/5-2\delta_2/5} \exp\,(-2(\log \tilde{N})^{\delta_2}).\end{aligned} \tag{8.34}$$

Let

$$\mathcal{E} := \mathcal{C}_n \cup \mathcal{C}_{2n} \cup \mathcal{B}^{(\tilde{N})}.$$

For $(x, y) \notin \mathcal{E}$, by (8.32) we have that

$$\bigg| \frac{1}{\tilde{N}} \log \|M_{\tilde{N}}(x, y)\| + L_n - 2L_{2n} \bigg| \le \frac{11}{n} e^{-(1/2)nL_n} + \frac{5U(\lambda, 1)n}{\tilde{N}} + 2\varepsilon_4. \tag{8.35}$$

Together with (8.28), this implies that for any $(x, y) \notin \mathcal{E}$,

$$\begin{aligned}&\bigg\| \mathbb{1}_{\mathcal{E}^c}(x, y) \bigg(\frac{1}{\tilde{N}} \log \|M_{\tilde{N}}(x, y)\| - L_{\tilde{N}} \bigg) \bigg\|_{L^\infty(\mathbb{T}^2)}\\ &\quad \le |L_{\tilde{N}} + L_n - 2L_{2n}| + \frac{11}{n} e^{-(1/2)nL_n} + \frac{5U(\lambda, 1)n}{\tilde{N}} + 2\varepsilon_4\\ &\quad \le \frac{2n}{\tilde{N}} (L_n - L_{2n}) + \frac{22}{n} e^{-(1/2)nL_n} + 8U(\lambda, 1)\tilde{N}^{-7/5+4\delta/5} + \frac{5U(\lambda, 1)n}{\tilde{N}} + 2\varepsilon_4\\ &\quad =: \varepsilon_0.\end{aligned} \tag{8.36}$$

Recall (8.23) states that

$$|\mathcal{B}^{(N)}| < 2N^{-7/5+4\delta/5} \quad \text{and} \quad |\mathcal{B}^{(2N)}| \le 4N^{-7/5+4\delta/5}.$$

Since $\delta > 0$, this clearly leads to

$$|\mathcal{B}^{(\tilde{N})}| \le 2^{17/5-4\delta/5} \tilde{N}^{-7/5+4\delta/5} < 16\tilde{N}^{-7/5+4\delta/5}. \tag{8.37}$$

Combining (8.37) with (8.34), we obtain

$$\begin{aligned}&\bigg\| \mathbb{1}_{\mathcal{E}}(x, y) \bigg(\frac{1}{\tilde{N}} \log \|M_{\tilde{N}}(x, y)\| - L_{\tilde{N}} \bigg) \bigg\|_{L^1(\mathbb{T}^2)}\\ &\quad \le U(\lambda, 1) |\mathcal{C}_n \cup \mathcal{C}_{2n} \cup \mathcal{B}^{(\tilde{N})}|\\ &\quad \le U(\lambda, 1) \bigg\{ 4\sqrt{2} \exp\bigg(\frac{\pi}{4} \bigg[\frac{17}{36} + \frac{B_1}{4B_3^2} - \varepsilon_5^{-1} \bigg] \bigg) + 16\tilde{N}^{-7/5+4\delta/5}\\ &\qquad + 2\sqrt{2} (C(B_4 - m_4))^{-1} \tilde{N}^{1/5-2\delta/5} (\log \tilde{N})^{-1/5-2\delta_2/5} \exp\,(-2(\log \tilde{N})^{\delta_2}) \bigg\} =: \varepsilon_1.\end{aligned} \tag{8.38}$$

By our conditions (V)–(VIII), we have

$$\begin{aligned} \varepsilon_0 &\le 4\varepsilon_4, \\ \varepsilon_1 &\le 18U(\lambda, 1)\tilde{N}^{-7/5+4\delta/5} =: 18U(\lambda, 1)\tilde{N}^{\eta}. \end{aligned} \tag{8.39}$$

Indeed, note that the right-hand sides of (V) and (VI) are precisely ε_4, which allows us to bound ε_0 by $4\varepsilon_4$. On the other hand, (VII) and (VIII) simply state that the sum of the first two terms in the braces defining ε_1 are bounded by the third term, $2\tilde{N}^{-7/5+4\delta/5}$.

Let

$$\begin{aligned} \varepsilon_3 &:= C_4\varepsilon_4(\log \tilde{N})^{\delta_4}, \\ \varepsilon_2 &:= C_5\varepsilon_4(\log \tilde{N})^{2\delta_4}, \\ r &= \frac{1-2\delta}{7-4\delta} \in (0, 1). \end{aligned} \tag{8.40}$$

Our conditions (X) and (XI) ensure that $\varepsilon_0 \le 4\varepsilon_4 < \varepsilon_3 < \varepsilon_2$. Recall that $B_4 - m_4$ and $B_5^{(\tilde{N})} - m_5^{(\tilde{N})}$ are as in (8.8). Therefore, by our condition (IX) we have

$$(B_5^{(\tilde{N})} - m_5^{(\tilde{N})})\tilde{N}^{-1} < 2\log R. \tag{8.41}$$

Applying Lemma 4.2 to $v_{\tilde{N}}$, and taking (8.41) into account, we obtain

$$\begin{aligned} &|\{(x, y) \in \mathbb{T}^2 : |v_{\tilde{N}}(x, y) - L_{\tilde{N}}| > \varepsilon_2\}| \\ &\le 2(2C_0)^{1/2} \exp\left(-\pi\left(36\varepsilon_0 + 16B_3\sqrt{\varepsilon_1^r(B_4 - m_4)}\right)^{-1}\varepsilon_3\right) \\ &\quad + C_0 \exp\left(-\pi\left(18\varepsilon_3 + 16B_3\sqrt{U(\lambda, 1)}\sqrt{\varepsilon_1^{1-r}(B_5^{(\tilde{N})} - m_5^{(\tilde{N})})}\right)^{-1}\varepsilon_2\right) \\ &\le 2(2C_0)^{1/2} \exp\left(-\pi\left(36\varepsilon_0 + 16B_3\sqrt{\varepsilon_1^r(B_4 - m_4)}\right)^{-1}\varepsilon_3\right) \\ &\quad + C_0 \exp\left(-\pi\left(18\varepsilon_3 + 16B_3\sqrt{U(\lambda, 1)}\sqrt{2\varepsilon_1^{1-r}\tilde{N}\log R}\right)^{-1}\varepsilon_2\right). \end{aligned} \tag{8.42}$$

Note that we changed B_6 into $U(\lambda, 1)$ in this expression. Inserting our estimates of ε_0, ε_1 (see (8.39)) and choices of ε_2, ε_3, r (see (8.40)) into (8.42), we arrive at

$$\begin{aligned} &|\{(x, y) \in \mathbb{T}^2 : |v_{\tilde{N}}(x, y) - L_{\tilde{N}}| > C_5\varepsilon_4(\log \tilde{N})^{2\delta_4}\}| \\ &\le 2(2C_0)^{1/2} \exp\left(\frac{-\pi C_4\varepsilon_4(\log \tilde{N})^{\delta_4}}{144\varepsilon_4 + 16B_3\sqrt{(18U(\lambda, 1))^r\tilde{N}^{\eta r}(B_4 - m_4)}}\right) \\ &\quad + C_0 \exp\left(\frac{-\pi C_5\varepsilon_4(\log \tilde{N})^{2\delta_4}}{18C_4\varepsilon_4(\log \tilde{N})^{\delta_4} + 16B_3\sqrt{U(\lambda, 1)}\sqrt{2(18U(\lambda, 1))^{1-r}\tilde{N}^{\eta(1-r)+1}\log R}}\right). \end{aligned}$$

Using (6.11), and $0 \le r \le 1$, we estimate

$$(18U(\lambda, 1))^r \le 18U(\lambda, 1) \quad \text{and} \quad (18U(\lambda, 1))^{1-r} \le 18U(\lambda, 1),$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

respectively. Hence we have

$$
\begin{aligned}
&|\{(x, y) \in \mathbb{T}^2 : |v_{\tilde{N}}(x, y) - L_{\tilde{N}}| > C_5\varepsilon_4(\log \tilde{N})^{2\delta_4}\}| \\
&\quad\le 2(2C_0)^{1/2} \exp\left(\frac{-\pi C_4\varepsilon_4(\log \tilde{N})^{\delta_4}}{144\varepsilon_4 + 48B_3\sqrt{2U(\lambda, 1)\tilde{N}^{\eta r}(B_4 - m_4)}}\right) \\
&\qquad + C_0 \exp\left(\frac{-\pi C_5\varepsilon_4(\log \tilde{N})^{2\delta_4}}{18C_4\varepsilon_4(\log \tilde{N})^{\delta_4} + 96B_3U(\lambda, 1)\sqrt{\tilde{N}^{\eta(1-r)+1}\log R}}\right).
\end{aligned}
$$

Plugging in our choice of ε_4 (see (8.33)), and noting that the powers of $\tilde{N}$ in numerators and denominators cancel out due to our choice of r, we infer that

$$
\begin{aligned}
&|\{(x, y) \in \mathbb{T}^2 : |v_{\tilde{N}}(x, y) - L_{\tilde{N}}| > C_5\varepsilon_4(\log \tilde{N})^{2\delta_4}\}| \\
&\le 2(2C_0)^{1/2} \exp\left(\frac{-\pi C_2C_4(\log \tilde{N})^{1/10+\delta_2/5+\delta_3+\delta_4}}{144C_2(\log \tilde{N})^{1/10+\delta_2/5+\delta_3} + 48B_3\sqrt{2U(\lambda, 1)(B_4 - m_4)}}\right) \\
&\quad + C_0 \exp\left(\frac{-\pi C_2C_5(\log \tilde{N})^{1/10+\delta_2/5+\delta_3+2\delta_4}}{18C_2C_4(\log \tilde{N})^{1/10+\delta_2/5+\delta_3+\delta_4} + 96B_3U(\lambda, 1)\sqrt{\log R}}\right) \\
&= 2(2C_0)^{1/2} \exp\left(\frac{-\pi C_2C_4(\log \tilde{N})^{\delta_4}}{144C_2 + 48B_3\sqrt{2U(\lambda, 1)(B_4 - m_4)}(\log \tilde{N})^{-1/10-\delta_2/5-\delta_3}}\right) \\
&\quad + C_0 \exp\left(\frac{-\pi C_2C_5(\log \tilde{N})^{\delta_4}}{18C_2C_4 + 96B_3U(\lambda, 1)\sqrt{\log R}(\log \tilde{N})^{-1/10-\delta_2/5-\delta_3-\delta_4}}\right),
\end{aligned}
$$

as desired. □

Readers will note that the constants were chosen in such a way that in the final steps of the proof only powers of $\log N$ remained inside of the exponential. We have found this to be more efficient over intermediate scales. The following, final, section of this paper will show how our work up to this point allows for such concrete estimates with specific numbers.

9. *Explicit numbers and proof of Theorem 1.1*

Our goal here is to make concrete choices of our parameters so as to arrive at an actual multi-scale scheme for the skew-shift operator from §6. Let $\mathcal{B}_n$ be as in Definition 8.1. The values below were found to be convenient ones, but clearly many other choices could have been made.

Definition 9.1. Set

$$R := 4, \quad R_1 := 3, \quad R_2 := 2$$

in Definition 2.1. The coupling constant in (6.1) is required to obey $\lambda \in [\frac{1}{2}, 1]$. Further, in Proposition 8.4 set

$$\delta := \tfrac{1}{8}, \quad \delta_2 := 1, \quad \delta_3 := 2, \quad \delta_4 := \tfrac{3}{2},$$

as well as

$$C_2 := 203, \quad C_4 := \frac{145}{\pi}, \quad C_5 := \frac{850}{\pi}.$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

By an explicit computation, the condition in Definition 3.5 is satisfied. In fact, one has

$$B_3^2 - 289\bigg(B_0 + \frac{13}{20\log(R/R_1)}\bigg) > 61 > 0.$$

PROPOSITION 9.2. *Let $\omega = (\sqrt{5}-1)/2$ be the golden ratio, and consider model (6.1) with $\lambda \in [\frac{1}{2}, 1]$ arbitrary but fixed. Let $a \geq 7$ and let n, N be positive integers such that $N \geq 10^{12}$, n divides N, and*

$$10^{13}(n+1)^8 \leq N, \qquad \frac{N}{(\log N)^{92/3}} < \frac{1}{2}\bigg(\frac{203}{22}e^{a/2}n\bigg)^{40/3}. \tag{9.1}$$

Impose the following conditions:
(a) $nL_n \geq a$;
(b) $L_n - L_{2n} \leq \frac{1}{8}L_n$;
(c) $\max(|\mathcal{B}_n|, |\mathcal{B}_{2n}|) \leq N^{-23/10}$.
Then we have

$$\begin{aligned}&|\{(x, y) \in \mathbb{T}^2 : |v_{\tilde{N}}(x, y) - L_{\tilde{N}}| > 5.5 \times 10^4 \tilde{N}^{-3/40}(\log \tilde{N})^{53/10}\}| \\ &\quad\leq\ 10\exp(-(\log \tilde{N})^{3/2})\end{aligned} \tag{9.2}$$

for $\tilde{N} = N$ and $2N$.

Remark 9.3. We will choose the constant $a = 7$ along the inductive multi-scale procedure. The only exception is the first step of the induction, which goes from the scale N_0 to N_1, where for some of our main results we use a larger value of a. This is made possible by assumption (i) on the Lyapunov exponent at the initial scale and it is the reason behind the relatively small values of N_0 in Theorems 1.3 and 1.4.

Proof. We need to check the hypotheses of Proposition 8.4. We already verified (3.14), and the conditions of Lemma 8.2 hold by assumption. Let $\tilde{N} = N$ or $2N$. The function

$$[0.5, 1] \to \mathbb{R} : \lambda \mapsto U(\lambda, 4) - \log\lambda$$

is decreasing and positive. Hence

$$0.5 < U(1, 4) \leq U(\lambda, 4) - \log\lambda \leq U(\tfrac{1}{2}, 4) - \log\tfrac{1}{2} < 1. \tag{9.3}$$

Further, the constant C in (I) satisfies $C < 11.97$. So that condition is implied by the stronger one,

$$12^8(4\log(2)n + 2\log(2) + 1)^8 \leq \tilde{N},$$

which we may further strengthen to

$$36^8(n+1)^8 < 10^{13}(n+1)^8 \leq N,$$

which is the left-hand side of (9.1). Condition (II) holds, as does (III) since $\exp(4(\log\tilde{N})^{\delta_2}) = \tilde{N}^4 \geq \tilde{N} + 1$. Condition (IX) is implied by the stronger one,

$$\tilde{N} > \frac{1 + \log R}{\log R} = \frac{2\log(2) + 1}{2\log(2)} \sim 1.721,$$

which clearly holds. In view of (8.8) and (9.3), we have

$$4\log(2) + 0.5 \leq B_4 - m_4 \leq 4\log(2) + 1. \tag{9.4}$$

Condition (IV) will therefore hold provided

$$\tilde{N}^{1/8}\log(\tilde{N}) - 21\tilde{N}^{-27/40}(\log(\tilde{N}))^{27/10} - 181 > 0.$$

The left-hand side is increasing in $\tilde{N}$, and one checks by explicit computation that it is positive if $N \geq 10^8$. So this condition holds as well. Condition (VIII) is implied by the condition

$$\tilde{N}^{-13/10} > \frac{2\sqrt{2}}{C(4\log(2)+0.5)}\tilde{N}^{-37/20}(\log(\tilde{N}))^{-3/5}.$$

Simplifying this, one obtains the stronger condition,

$$\tilde{N}^{1/5}(\log(\tilde{N}))^{3/5} > 0.08,$$

which holds provided $N \geq 2$. So condition (VIII) holds.

Next, we look at condition (VI). Using the assumed lower bound $nL_n \geq a$, we find the condition

$$22e^{-a/2} < 203n\tilde{N}^{-3/40}(\log(\tilde{N}))^{23/10.}$$

We recall that $\tilde{N} \in \{N, 2N\}$ and estimate $\log(\tilde{N}) \geq \log(N)$. This inequality follows from the upper bound in (9.1). For condition (VII), one checks that it follows from the slightly stronger

$$\tilde{N}^{-13/10} - 5.66\exp(0.374 - 0.05(\log(\tilde{N}))^2) > 0$$

which holds for $\tilde{N} \geq 10^{12}$ (but fails for 10^{11}). Hence we impose the second lower bound in (9.1). For condition (V), we use

$$L_n - L_{2n} \leq \frac{1}{8}L_n \leq \frac{U(\lambda, 1)}{8} \leq \frac{1}{4},$$

and so it suffices to check that

$$\frac{21n}{2\tilde{N}} + 16\tilde{N}^{-13/10} < 203\tilde{N}^{-3/40}(\log(\tilde{N}))^{23/10}.$$

Bounding n in terms of N via (9.1) and discarding the $\log\tilde{N}$ on the right-hand side reduces us to

$$\frac{21}{2\cdot 10^{13/8}}\cdot N^{-7/8} + 16\tilde{N}^{-13/10} < 203\tilde{N}^{-3/40}.$$

This holds for all $N \geq 1$, so we are done with (V). Finally, we turn to (X) and (XI). Using $N \geq 10^{12}$, they hold provided

$$C_4 \geq 0.03, \quad C_5 \geq 0.007\cdot C_4.$$

Our actual values assigned to these constants satisfy

$$46 < C_4 < 47, \quad C_5 > 270,$$

and so all conditions of Proposition 8.4 hold.

As for the conclusion of that proposition, we first compute $C_2C_5 < 5.5 \times 10^4$. Thus, the sizes of the deviations satisfy

$$C_2C_5\tilde{N}^{-1/10+\delta/5}(\log \tilde{N})^{1/10+\delta_2/5+\delta_3+2\delta_4} < 5.5 \times 10^4\tilde{N}^{-3/40}(\log \tilde{N})^{53/10},$$

as stated in (9.2). As for the measure bound, we calculate that

$$2(2C_0)^{1/2} + C_0 < 10,$$
$$U(\lambda, 1) = \tfrac{1}{2}\log((4\lambda+2)^2+2) \le \tfrac{1}{2}\log 38.$$

Thus, in view of (9.4), one has $48B_3\sqrt{2U(\lambda, 1)(B_4 - m_4)} < 11\,518$, and

$$144 + 11\,518C_2^{-1}(\log \tilde{N})^{-23/10} \le 144 + \tfrac{11\,518}{203}(12\log 10)^{-23/10} < 145.$$

Hence the first exponential in the measure bound of Proposition 8.4 contributes less than

$$\exp(-\pi C_4(\log \tilde{N})^{3/2}/145) < \exp(-(\log \tilde{N})^{3/2}).$$

For the second exponential, we have $18C_4 < 831$, $96B_3U(\lambda, 1)\sqrt{2\log 2} < 13\,317$, and

$$831 + 13\,317C_2^{-1}(\log \tilde{N})^{-38/10} \le 831 + \tfrac{13\,317}{203}(12\log 10)^{-38/10} < 850.$$

Hence, the second exponential contributes less than

$$\exp(-\pi C_5(\log \tilde{N})^{3/2}/850) < \exp(-(\log \tilde{N})^{3/2}),$$

and we are done. □

9.1. *Proof of Theorem 1.1.* Let $N_0 := 2 \times 10^{37}$. We define a sequence of scales $N_j := N_{j-1}^9$ for $j \ge 1$. In particular, $N_1 > 5 \times 10^{335}$. The proof is based on an induction on scales, where at every step we first apply Lemma 8.3 to control the Lyapunov exponent at the next scale. Afterwards, we apply Proposition 9.2 to obtain the large-deviation estimate at the next scale and then we continue the induction.

For later purposes, we note some properties of this choice of scales. The last inequality is the main reason why we need to choose the scale so that N_1 is large.

LEMMA 9.4. *Recall that we defined $N_{j+1} := N_j^9$ with $N_0 = 2 \times 10^{37}$. For all $j \ge 1$, we have the following bounds:*

$$10^{13}(N_{j-1}+1)^8 \le N_j, \quad \frac{N_j}{(\log N_j)^{92/3}} < \frac{1}{2}\left(\frac{203}{22}e^{7/2}N_{j-1}\right)^{40/3} \tag{9.5}$$

as well as

$$10\exp(-(\log N_j)^{3/2}) \le (N_j^9)^{-2.3} = N_{j+1}^{-2.3} \tag{9.6}$$

and

$$5.5 \times 10^4 N_j^{-3/40}(\log N_j)^{53/10} \le \tfrac{1}{20}L_{N_0}. \tag{9.7}$$

Proof. From the definition of the N_j, we have

$$10^{13}(N_{j-1}+1)^8 \le N_j < N_{j-1}^{13}, \tag{9.8}$$

and this implies (9.5). Notice that we have $N_j \geq N_1 > 5 \times 10^{335}$ for all $j \geq 1$. Then (9.6) follows from the inequality

$$10 \exp(-(\log x)^{3/2}) \leq (x^9)^{-2.3}, \tag{9.9}$$

which holds for all $x \geq 2.06 \times 10^{186}$. For (9.7), we note that

$$5.5 \times 10^4 x^{-3/40} (\log x)^{53/10} \leq 10^{-5} = \tfrac{1}{20} \times 2 \times 10^{-4} \leq \tfrac{1}{20} L_{N_0},$$

where the first inequality holds for all $x \geq 10^{334}$. This proves the lemma. □

We will inductively apply Lemma 8.3 to $j = 1, 2, 3, \ldots$. We begin with $j = 1$. Condition (1) of Lemma 8.3 follows from our assumptions (i) and (ii) and the fact that $N_0 = 2 \times 10^{37}$. Condition (2) with $j = 1$ is fulfilled since

$$\frac{1}{N_0} \exp\left(-\frac{1}{2} N_0 L_{N_0}\right) \leq \frac{1}{2 \times 10^{37}} < \frac{2}{512} 10^{-4} \leq \frac{1}{512} L_{N_0}.$$

For condition (3), recall that $N_1 > 5 \times 10^{335}$, $\delta = \frac{1}{8}$ and $C_3 = U(\lambda, 1) \leq \frac{1}{2} \log 38$. Then we have

$$N_1^{-13/10} < (5 \times 10^{335})^{-1.3} < \frac{1}{320 \log 38} 10^{-4} \leq \frac{1}{1280 U(\lambda, 1)} L_{N_0}.$$

For condition (4),

$$\max\, (|\mathcal{B}_{N_0}|, |\mathcal{B}_{2N_0}|) \leq N_0^{-21} = (N_1)^{-7/3} \leq N_1^{-2.3}. \tag{9.10}$$

Hence Lemma 8.3 applies to $j = 1$ and yields

$$N_1 L_{N_1} \geq 7, \quad L_{N_1} - L_{2N_1} \leq \tfrac{1}{8} L_{N_1} \tag{9.11}$$

and

$$L_{2N_1} \geq \tfrac{1}{2} L_{N_0}. \tag{9.12}$$

We would like to apply Lemma 8.3 for $j = 2$. This requires measure estimates for $\mathcal{B}_{N_1}$ and $\mathcal{B}_{2N_1}$. To this end, we invoke Proposition 9.2 with $n = N_0$, $N = N_1$ and $a = 7$. Let us check that its conditions are satisfied. First, we have (9.1) by applying (9.5) with $j = 0$. Moreover, conditions (a)–(c) hold by assumptions (i)–(iii) and (9.10). Hence, we can apply Proposition 9.2 and obtain that, for $\tilde{N} = N_1$ and $2N_1$,

$$\begin{aligned}
&|\{(x, y) \in \mathbb{T}^2 : |v_{\tilde{N}}(x, y) - L_{\tilde{N}}| > 5.5 \times 10^4 \tilde{N}^{-3/40} (\log \tilde{N})^{53/10}\}| \\
&\quad \leq 10 \exp\, (-(\log N_1)^{2.3}) \leq N_2^{-2.3}.
\end{aligned}$$

In the second step, we used (9.6) with $j = 1$. To turn this into measure estimates for $\mathcal{B}_{\tilde{N}}$, notice that (9.7) and (9.12) imply

$$5.5 \times 10^4 \tilde{N}^{-3/40} (\log \tilde{N})^{53/10} \leq \tfrac{1}{20} L_{N_0} \leq \tfrac{1}{10} L_{\tilde{N}}.$$

Therefore,

$$\max\, (|\mathcal{B}_{N_1}|, |\mathcal{B}_{2N_1}|) \leq N_2^{-2.3}. \tag{9.13}$$

We have shown how to pass from scale N_0 to N_1 via Lemma 8.3 and Proposition 9.2, by using the properties (9.5)–(9.7).

We can now iterate this procedure. We apply Lemma 8.3 with $j = 2$. The main input is the measure estimate (9.13), which verifies condition (4). The remaining conditions hold by our choice of scales, (9.11), (9.12) and assumptions (i) and (ii). (Notice that the sums in conditions (2) and (3) are rapidly convergent.) From Lemma 8.3, we obtain estimates of L_{N_2} and L_{2N_2}, in particular $L_{2N_2} \geq \frac{1}{2} L_{N_0}$. Then Proposition 9.2 yields the measure estimates for $\mathcal{B}_{N_2}$ and $\mathcal{B}_{2N_2}$, which is the key input for Lemma 8.3 with $j = 3$, etc. We conclude that, after k steps of this procedure, we have

$$L_{2N_k} \geq \tfrac{1}{2} L_{N_0}.$$

This yields

$$L \geq \tfrac{1}{2} L_{N_0},$$

by taking $k \to \infty$, and we have proved Theorem 1.1.

9.2. *Proof of Theorem 1.3.* We follow the general line of argumentation of Theorem 1.1. The only difference is that in the sequence of scales N_j, we take the first step to be very large. Namely, while $N_0 = 3 \times 10^5$, we define

$$N_1 := 3 \times 10^{334}, \quad N_{j+1} := N_j^9 \quad \text{for all } j \geq 1. \tag{9.14}$$

Notice that for $j \geq 1$, the scales N_j are essentially the ones used in the proof of Theorem 1.1 above. Therefore we have the following analog of Lemma 9.4, in which (9.6) for $j = 1$ is replaced.

LEMMA 9.5. *We have*

$$10^{13}(N_0 + 1)^8 \leq N_1, \quad \frac{N_1}{(\log N_1)^{92/3}} < \frac{1}{2}\left(\frac{203}{22} e^{30} N_0\right)^{40/3}. \tag{9.15}$$

Moreover, for all $j \geq 1$, we have the bounds

$$10^{13}(N_j + 1)^8 \leq N_{j+1}, \quad \frac{N_{j+1}}{(\log N_{j+1})^{92/3}} < \frac{1}{2}\left(\frac{203}{22} e^{7/2} N_j\right)^{40/3}, \tag{9.16}$$

as well as (9.6) and (9.7).

Except for (9.15), the bounds are only concerned with N_j, $j \geq 1$, and therefore follow in the same way as for Lemma 9.5. The new bound (9.15) follows from

$$\left(2 \frac{N_1}{(\log N_1)^{92/3}}\right)^{3/40} \frac{1}{(203/22) e^{30}} < 29\,974 < N_0.$$

This establishes Lemma 9.5. As before, we will successively apply Lemma 8.3 and Proposition 9.2 and iterate. We begin by applying Lemma 8.3 with $j = 1$. Condition (1) is immediate from assumption (i) and $N_0 = 3 \times 10^5$; indeed,

$$N_0 L_{N_0} \geq 2N_0 10^{-4} = 60. \tag{9.17}$$

We can use this inequality to verify condition (2) as well:

$$\frac{1}{N_0} e^{-N_0 L_{N_0}/2} \leq \frac{1}{3} 10^{-5} e^{-30} < 10^{-18} < \frac{2}{512} 10^{-4} < \frac{L_{N_0}}{512}. \tag{9.18}$$

Condition (3) holds by our choice of N_1. Finally, condition (4) holds by assumption (iii):

$$\max\,(|\mathcal{B}_{N_0}|,\,|\mathcal{B}_{2N_0}|) \le N_0^{-141} < (3\times 10^{334})^{-2.3} = N_1^{-2.3}. \tag{9.19}$$

Hence, Lemma 8.3 applies and yields (9.11) and (9.12) as before. Next, we verify the assumption of Proposition 9.2 with $n = N_0$ and $N = N_1$. The key difference is that we now take $a = 60$. This is made possible by (9.17), since it verifies condition (a) of Proposition 9.2. Condition (b) is immediate from assumption (ii), and condition (c) was checked in (9.19). The bounds (9.1) hold by (9.15). Therefore, we can apply Proposition 9.2. Combining the resulting estimate with (9.7) for $j = 1$, (9.12) and (9.6) for $j = 1$, we obtain the measure estimate

$$\max\,(|\mathcal{B}_{N_1}|,\,|\mathcal{B}_{2N_1}|) \le N_2^{-2.3}.$$

At this point, we have moved completely from scale N_0 to scale N_1 and can follow the argument from Theorem 1.1 verbatim. In particular, we take $a = 7$ in every subsequent application of Proposition 9.2. The only difference is the $m = 0$ term in condition (2) of Lemma 8.3, which now involves $N_0 = 3\times 10^{-5}$. By (9.18), we can replace condition (2) by the stronger bound

$$\sum_{m=1}^{j-1} N_m^{-1} e^{-(1/2)N_m L_{N_m}} < \frac{10^{-4}}{256} - 10^{-18},$$

and this holds by our choice of scales and the estimates (9.12) along the induction (notice again the rapid convergence of the series). We conclude that

$$L \ge \tfrac{1}{2} L_{N_0}$$

and this proves Theorem 1.3.

9.3. *Proof of Theorem 1.4.* Again, we follow the same steps for a different sequence of scales. We have $N_0 = 3\times 10^4$. We define the sequence of scales N_j, $j \ge 1$, by

$$N_1 := 3\times 10^{320}, \qquad N_{j+1} := N_j^9 \quad \text{for all } j \ge 1.$$

We still have Lemma 9.5 for this choice of scales. Indeed, (9.16) and (9.6) still follow from the inequalities (9.8) and (9.9) given in the proof of Lemma 9.4. For (9.7), we now use assumption (i) to find

$$5.5\times 10^4 x^{-3/40}(\log x)^{53/10} \le 10^{-4} = \tfrac{1}{20}\times 2\times 10^{-3} \le \tfrac{1}{20} L_{N_0},$$

where the first inequality holds for all $x \ge 10^{320}$, so in particular for all N_j with $j \ge 1$. Finally, (9.15) follows from

$$\left(2\frac{N_1}{(\log N_1)^{92/3}}\right)^{3/40} \frac{1}{(203/22)e^{30}} < 2938 < N_0.$$

This establishes Lemma 9.5 for the new choice of scales.

Next we check the hypotheses for Lemma 8.3 with $j = 1$. Condition (1) is immediate from assumption (i) and $N_0 = 3\times 10^4$:

$$N_0 L_{N_0} \ge 2N_0 10^{-3} = 60 \tag{9.20}$$

https://doi.org/10.1017/etds.2019.19 Published online by Cambridge University Press

(compare this to (9.17)). Condition (2) holds by

$$\frac{1}{N_0}e^{-N_0 L_{N_0}/2} \leq \frac{1}{3}10^{-4}e^{-30} < 10^{-17} < \frac{2}{512}10^{-3} < \frac{L_{N_0}}{512}, \tag{9.21}$$

where we used (9.20). Condition (3) holds by our choice of scales N_j, $j \geq 1$, and condition (4) holds by assumption (iii):

$$\max\,(|\mathcal{B}_{N_0}|, |\mathcal{B}_{2N_0}|) \leq N_0^{-165} < (3 \times 10^{320})^{-2.3} = N_1^{-2.3}. \tag{9.22}$$

Therefore we can apply Lemma 8.3 and obtain (9.11) and (9.12). As in the proof of Theorem 1.3, the first application of Proposition 9.2 utilizes $a = 60$. This is made possible by (9.17), since it verifies condition (a) of Proposition 9.2. Now we iterate the argument in the same way as was done for Theorems 1.1 and 1.3. (Notice that the series in conditions (2) and (3) of Lemma 8.3 are still rapidly convergent.) The end result is the lower bound

$$L \geq \tfrac{1}{2}L_{N_0},$$

and Theorem 1.4 is proved.

Acknowledgements. The authors thank the Institute for Advanced Study, Princeton, for its hospitality during the 2017–18 academic year. This material is based upon work supported by the National Science Foundation under Grant No. DMS-1638352. The third author was partially supported by NSF grant DMS-1500696. The authors thank James Maynard and Silvius Klein for helpful conversations, and Jean Bourgain and Thomas Spencer for their interest in this work.

REFERENCES

[BelSim] J. Béllissard and B. Simon. Cantor spectrum for the almost Mathieu equation. *J. Funct. Anal.* **48**(3) (1982), 408–419.

[Bou1] J. Bourgain. *Green's Function Estimates for Lattice Schrödinger Operators and Applications (Annals of Mathematics Studies, 158)*. Princeton University Press, Princeton, NJ, 2005.

[Bou2] J. Bourgain. On the spectrum of lattice Schrödinger operators with deterministic potential. *J. Anal. Math.* **87** (2002), 37–75.

[Bou3] J. Bourgain. Positive Lyapounov exponents for most energies. *Geometric Aspects of Functional Analysis (Lecture Notes in Mathematics, 1745)*. Springer, Berlin, 2000, pp. 37–66.

[BouGol] J. Bourgain and M. Goldstein. On nonperturbative localization with quasi-periodic potential. *Ann. of Math. (2)* **152**(3) (2000), 835–879.

[BouGolSch] J. Bourgain, M. Goldstein and W. Schlag. Anderson localization for Schrödinger operators on $\mathbb{Z}$ with potentials given by the skew-shift. *Comm. Math. Phys.* **220**(3) (2001), 583–621.

[Dam] D. Damanik. Lyapunov exponents and spectral analysis of ergodic Schrödinger operators: a survey of Kotani theory and its applications. *Spectral Theory and Mathematical Physics: A Festschrift in Honor of Barry Simon's 60th Birthday (Proceedings of Symposia in Pure Mathematics, 76, Part 2)*. American Mathematical Society, Providence, RI, 2007, pp. 539–563.

[DuaKle] P. Duarte and S. Klein. Lyapunov exponents of linear cocycles. *Continuity via Large Deviations (Atlantis Studies in Dynamical Systems, 3)*. Atlantis Press, Paris, 2016.

[DuaKle2] P. Duarte and S. Klein. *Continuity of the Lyapunov Exponents of Linear Cocycles*. Associação Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 2017.

[Fur] H. Fürstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.* **108** (1963), 377–428.

[GolSch] M. Goldstein and W. Schlag. Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions. *Ann. of Math. (2)* **154**(1) (2001), 155–203.

[**Hea**] D. R. Heath-Brown. Pair correlation for fractional parts of αn^2. *Math. Proc. Cambridge Philos. Soc.* **148**(3) (2010), 385–407.

[**Her**] M.-R. Herman. Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnol'd et de Moser sur le tore de dimension 2. *Comment. Math. Helv.* **58**(3) (1983), 453–502.

[**Kat**] Y. Katznelson. *An Introduction to Harmonic Analysis (Cambridge Mathematical Library)*, 3rd edn. Cambridge University Press, Cambridge, 2004.

[**Kru1**] H. Krüger. Multiscale analysis for ergodic Schrödinger operators and positivity of Lyapunov exponents. *J. Anal. Math.* **115** (2011), 343–387.

[**Kru2**] H. Krüger. On positive Lyapunov exponent for the skew-shift potential. *Preprint*.

[**MarStr**] J. Marklof and A. Strömbergsson. Equidistribution of Kronecker sequences along closed horocycles. *Geom. Funct. Anal.* **13**(6) (2003), 1239–1280.

[**Mon**] H. L. Montgomery. *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis (CBMS Regional Conference Series in Mathematics, 84)*. American Mathematical Society, Providence, RI, 1994.

[**RudSarZah**] Z. Rudnick, P. Sarnak and A. Zaharescu. The distribution of spacings between the fractional parts of $n^2\alpha$. *Invent. Math.* **145**(1) (2001), 37–57.

[**Sch**] W. Schlag. Regularity and convergence rates for the Lyapunov exponents of linear cocycles. *J. Mod. Dyn.* **7**(4) (2013), 619–637.

[**SorSpe**] E. Sorets and T. Spencer. Positive Lyapunov exponents for Schrödinger operators with quasi-periodic potentials. *Comm. Math. Phys.* **142**(3) (1991), 543–566.

[**Via**] M. Viana. *Lectures on Lyapunov Exponents (Cambridge Studies in Advanced Mathematics, 145)*. Cambridge University Press, Cambridge, 2014.