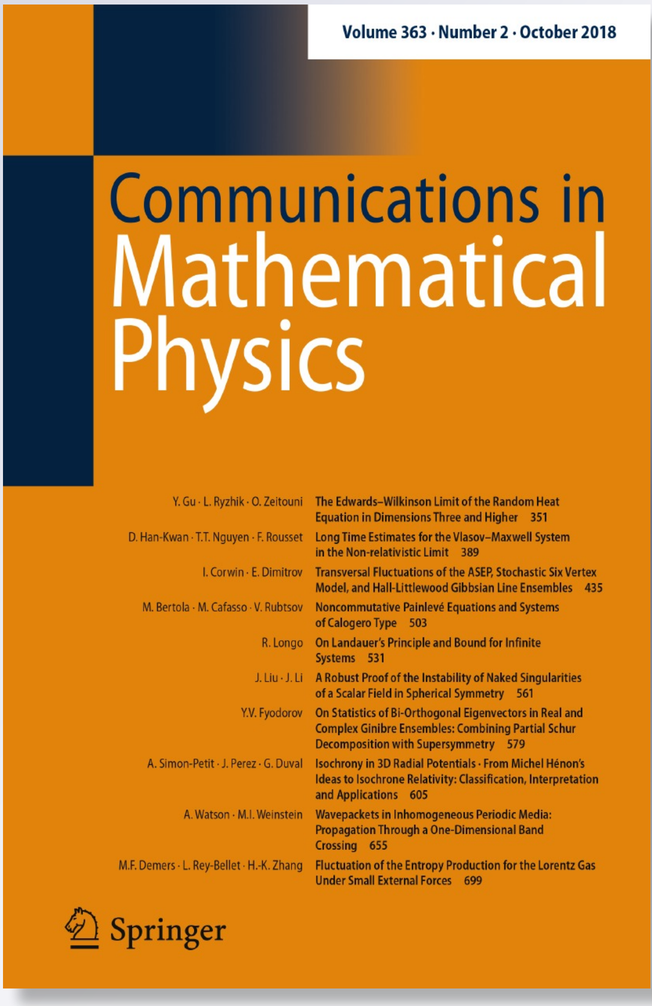
Fluctuation of the Entropy Production for the Lorentz Gas Under Small External Forces

Mark F. Demers, Luc Rey-Bellet & Hong-Kun Zhang

Communications in Mathematical Physics

ISSN 0010-3616
Volume 363
Number 2

Commun. Math. Phys. (2018)
363:699-740
DOI 10.1007/s00220-018-3228-3



Springer

Commun. Math. Phys. 363, 699–740 (2018)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-018-3228-3

Communications in Mathematical Physics

CrossMark

Fluctuation of the Entropy Production for the Lorentz Gas Under Small External Forces

Mark F. Demers[1], **Luc Rey-Bellet**[2], **Hong-Kun Zhang**[2]

[1] Department of Mathematics, Fairfield University, Fairfield, CT 06824, USA. E-mail: mdemers@fairfield.edu
[2] Department of Mathematics and Statistics, University of Massachusetts Amherst, Amherst, MA 01003, USA. E-mail: luc@math.umass.edu; hongkun@math.umass.edu

Received: 13 October 2017 / Accepted: 18 June 2018
Published online: 24 August 2018 –

Abstract: In this paper we study the physical and statistical properties of the periodic Lorentz gas with finite horizon driven to a non-equilibrium steady state by the combination of non-conservative external forces and deterministic thermostats. A version of this model was introduced by Chernov, Eyink, Lebowitz, and Sinai and subsequently generalized by Chernov and the third author. Non-equilibrium steady states for these models are SRB measures and they are characterized by the positivity of the steady state entropy production rate. Our main result is to establish that the entropy production, in this context equal to the phase space contraction, satisfies the Gallavotti–Cohen fluctuation relation. The main tool needed in the proof is the family of anisotropic Banach spaces introduced by the first and third authors to study the ergodic and statistical properties of billiards using transfer operator techniques.

1. Introduction

The periodic Lorentz gas (or Sinai billiard) is obtained by placing finitely many disjoint scatterers with smooth boundaries of strictly positive curvature on the 2-torus. The dynamics is the motion of a point particle traveling at unit speed and undergoing elastic reflections at the boundaries and is purely Hamiltonian. The associated two-dimensional collision map (the billiard map) preserves a smooth invariant measure μ_0 with very strong ergodic properties: see the works by Sinai, Bunimovich and Chernov [S,BS,BSC,Ch1] on ergodicity, mixing and the central limit theorem, the proof by Young [Y] of exponential decay of correlations, and many other statistical properties [RY,MN1,MN2] as well as the recent proof by Baladi, Liverani and one of the authors [BDL] for the exponential decay of correlations for the billiard flow. Of particular importance for this paper are the recent papers by two of the authors [DZ1,DZ2,DZ3] who introduced Banach spaces

M. F. Demers is partially supported by NSF Grant DMS-1362420; L. Rey-Bellet is partially supported by NSF Grant DMS-1515712; H.-K. Zhang is partially supported by NSF Grant DMS-1151762. The authors wish to thank the anonymous referee for helpful comments and suggestions.

suitable for a direct analysis of the dynamics by transfer operators, which bypasses the construction of symbolic dynamics (Markov partitions and Young towers). These functional analytic tools will turn out to be crucial to prove the large deviation theorems needed in this paper.

Suitable perturbations of this model where the particle is submitted to external non-conservative forces in between or during collisions and to a suitable thermostatting mechanism have been put forward as simple, yet realistic, models in non-equilibrium statistical mechanics. With a constant external electric field and an iso-energetic thermostat, this kind of model was first studied by Chernov, Eyink, Lebowitz and Sinai [CELS1,CELS2] who proved the existence of a unique SRB measure μ_+ for the system: for μ_0-almost-every initial condition the system converges to an invariant measure μ_+ which is ergodic and mixing, and singular with respect to μ_0. In addition, they established linear response formulas for this system. In subsequent papers, Chernov and one of the authors [Ch2,Ch4,CZZ,Z] generalized and strengthened these results to cover a large class of perturbations and our work will rely on these results extensively. In a more general context the use of thermostats and SRB measures as good models of non-equilibrium steady states has been advocated, see e.g. the book by Evans and Morriss [EM] and the papers by Gallavotti and Cohen [GC1,GC2] and Ruelle [R3] (more on this in Sect. 1.1.)

One of the main results in this paper is to establish a version of the so-called Gallavotti–Cohen fluctuation theorem [GC1,GC2] for the entropy production for the Lorentz gas driven out of equilibrium by external forces. The concept of entropy production in non-equilibrium statistical mechanics, in this context, was best formalized by Ruelle [R1,R2,R3] (see also the earlier work by Andrey [A]) and we will discuss it in Sect. 1.1. The fluctuation theorem asserts that for time-reversible systems the time fluctuations (of large deviation type) of the entropy production have a universal symmetry: the ratio of the probabilities of observing an average entropy production rate over a time interval of length T equal to a and equal to $-a$ is equal to e^{aT}. The study of the fluctuations of the entropy production for systems driven out of equilibrium originated in the numerical observation by Evans, Cohen, and Morris [ECM] for a thermostatted system driven by external shear. The symmetry of the transient fluctuations of entropy production, that is when the system starts in the equilibrium (but not stationary) state (μ_0 in our notation) was first noted by Evans and Searles [ES1] (see Proposition 1.4 in Sect. 1.1). On the other hand, using Markov partitions, in [GC1,GC2] Cohen and Gallavotti established the fluctuation symmetry for time-reversible smooth uniformly hyperbolic systems starting in a stationary non-equilibrium state. The relation between the transient and stationary fluctuation theorem is discussed further in [CG,ES2,JPR]. From a slightly different point of view, Kurchan [Ku], Lebowitz and Spohn [LS] proved the fluctuation theorem for general stochastic (Markovian and/or Gibbsian) dynamics and Maes [M1] recast the fluctuation theorem as following from the Gibbs property of an equilibrium state by considering the distribution of the time series of the process. Also in a related work, Jarzynski [Ja] established a very influential transient relation for the fluctuations of work of a system driven by time-dependent forces. These (and other) seminal works have given rise to a substantial amount of research in the past 20 years, and the fluctuation theorems and relations now stand as one of the pillars in the modern theory of non-equilibrium statistical mechanics. There have been a number of recent reviews, among them [M2,MN,ChGa,JPR], to which we direct the reader for some of the recent developments in this subject. Among these reviews, Jaksic, Pillet, and one of the authors [JPR] present a general formalism to understand the transient and stationary fluctuation theorems, and the relation between them, in the general framework of dynamical systems; to some extent, we will follow the approach taken in that paper.

In this paper, we prove the steady state fluctuation relations for the periodic Lorentz gas with an external electric field and an iso-energetic thermostat [CELS1,CELS2] as well as several classes of related models with different forcing mechanisms [Ch2,Ch4, CZZ,Z]. While the models at hand are uniformly hyperbolic, the singularities of the billiard dynamics (due to grazing collisions) preclude the use of Markov partitions to study the fluctuation properties of ergodic averages. Instead, we follow a direct approach using suitable transfer operators to express the cumulant generating function of ergodic averages. This approach to large deviations was used for hyperbolic dynamical systems in [RY] using Young towers [Y], especially for the Lorentz gas with finite horizon. Our approach consists of proving that the fluctuation properties of ergodic averages are the same for a large class of initial distributions, which contains both the stationary distribution μ_0 of the Lorentz gas without external forces used to verify the transient fluctuation theorem, and the invariant SRB measure for the perturbed Lorentz gas. Since the symmetry of fluctuations when starting from μ_0 is easy to establish (see Proposition 1.4) a proof of the fluctuation theorem follows then immediately. The key new tool needed is the family of Banach spaces introduced by two of the authors [DZ1,DZ2,DZ3] to study the ergodic properties of billiards without using the symbolic dynamics tools used in earlier approaches (Markov partitions [BS], Markov sieves [BSC], Young towers [Y,Ch1]). These Banach spaces are devised for the exact purpose to be large enough to contain the SRB invariant measure, singular with respect to μ_0 but smooth along unstable directions, yet small enough for the transfer operator to have a spectral gap. They also have the advantage of being stable under perturbations: since all the relevant transfer operators act on a single Banach space, we are able to show that important spectral quantities vary smoothly as functions of certain system parameters, and from these properties we derive the necessary control to prove the desired limit theorems.

This paper is organized as follows. In Sect. 1.1 we give a brief overview of the ideas and concepts of non-equilibrium statistical mechanics needed for the paper. In Sect. 2 we introduce our model and state our main results. In Sect. 2.2, following [DZ2], we discuss a general family of maps with singularities, to which our dynamical results apply. In Sect. 3 we introduce the Banach spaces and transfer operators needed in our analysis. In Sect. 4 we prove the key analytical estimates needed to establish a spectral gap for the family of transfer operators associated with the entropy production. Finally, in Sect. 5 we establish the analyticity and (strict) convexity of the logarithmic moment generating function, allowing us to conclude the proof of the fluctuation theorem. In Appendix A we provide the Lasota–Yorke estimates needed to establish a spectral gap for the relevant operators.

1.1. Entropy production and fluctuation theorems. In this section, for the convenience of the reader, we provide a general (and somewhat informal) discussion, following [JPR], of the concepts of non-equilibrium steady states, entropy production, and the fluctuation relations.

The starting point is an invertible dynamical system (M, T), i.e. a measurable space M and an invertible measurable map $T : M \to M$. We also postulate the existence of a reference measure μ_0 which, in general, is not an invariant measure for T.

In a physical context one may write $T = T_{\mathbf{E}}$ depending on some external non-equilibrium forces $\mathbf{E}$ with T_0 (for $\mathbf{E} = 0$) being the equilibrium dynamics without external forces. One may think of μ_0 as the invariant measure for the dynamics T_0

without external forces; in this context, μ_0 is the equilibrium steady state. If we think of μ_0 as describing the initial state of the system, we then define μ_n as the state of the system at time $n \in \mathbb{Z}$, i.e. we have

$$\mu_n(f) = \mu_0(f \circ T^n), \tag{1.1}$$

for any bounded measurable f.

We introduce next the concept of a non-equilibrium steady state following Ruelle [R3].

Definition 1.1. A probability measure μ_+ is called a non-equilibrium steady state for the dynamical system (M, T) with reference measure μ_0 if:

(1) the measure μ_+ is an ergodic invariant measure for T;
(2) for μ_0-almost every initial condition $x \in M$ the empirical measure $\frac{1}{n}\sum_{k=0}^{n-1} \delta_{T^k(x)}$ converges weakly to μ_+ as $n \to \infty$;
(3) the measure μ_+ is singular with respect to μ_0.

Item (2) in the definition selects one invariant measure μ_+ among the usually many invariant measures of the dynamical system (M, T) and it is essentially equivalent to the SRB property in the theory of hyperbolic dynamical systems if M is a smooth manifold and μ_0 is Lebesgue measure. Measures satisfying (2) are also often called "physical measures" as they describe the statistics of "most" initial conditions. Item (3) in the definition ensures that the invariant measure is truly a "non-equilibrium" steady state in the sense of statistical mechanics, while if μ_+ were equivalent to μ_0 it should rather be called an equilibrium steady state. Finally in a physical context where $T = T_{\mathbf{E}}$ depends on external forces, the non-equilibrium steady state μ_+ depends on $\mathbf{E}$ and we will use the notation $\mu_{\mathbf{E}}$ in that case.

Next we turn to the concept of the entropy production observable $s : M \to \mathbb{R}$ which plays a central role in non-equilibrium statistical mechanics. We make the (rather weak) regularity assumption that μ_n and μ_0 are mutually absolutely continuous and denote by l_n the logarithm of the Radon-Nykodym derivative,

$$l_n = \log \frac{d\mu_n}{d\mu_0}.$$

Since $\mu_{n+m}(f) = \mu_m(f \circ T^n) = \mu_0(e^{l_m} f \circ T^n) = \mu_n(e^{l_m \circ T^{-n}} f) = \mu_0(e^{l_n} e^{l_m \circ T^{-n}} f)$, we have the chain rule, $l_{n+m} = l_n + l_m \circ T^{-n}$, and in particular, $l_{-1} = -l_1 \circ T$. Therefore, we have

$$l_n = \sum_{k=0}^{n-1} l_1 \circ T^{-k}.$$

For two probability measures μ and ν on M, let us denote by $R(\mu|\nu)$ the relative entropy of μ with respect to ν (also known as the Kullback–Leibler divergence) which is defined by

$$R(\mu|\nu) = \begin{cases} \int \log \frac{d\mu}{d\nu} d\mu & \text{if } \mu \ll \nu \\ +\infty & \text{otherwise} \end{cases}.$$

We have then

$$R(\mu_n|\mu_0) = \mu_n(l_n) = \mu_n\left(\sum_{k=0}^{n-1} l_1 \circ T^{-k}\right) = \mu_0\left(\sum_{k=1}^{n} l_1 \circ T^k\right),$$

using (1.1). This leads to the following definition.

Definition 1.2. The entropy production observable for the dynamical (M, T) with reference measure μ_0 is is given by

$$s = l_1 \circ T.$$

If we assume the existence of a non-equilibrium steady state and if the entropy production observable s is regular enough we have

$$\lim_{n\to\infty} \frac{1}{n} R(\mu_n|\mu_0) = \lim_{n\to\infty} \mu_0 \left(\frac{1}{n} \sum_{k=0}^{n-1} s \circ T^k \right) = \mu_+(s) \geq 0\,,$$

since the relative entropy is non-negative. This general fact is known as the non-negativity of the entropy production rate in non-equilibrium steady states. It is shown in [JPR, Section 5] that, under quite general conditions, we have

$$\mu_+(s) > 0 \ \text{ if and only if } \mu_+ \text{ is singular with respect to } \mu_0.$$

We expect in any case that, for a bona fide non-equilibrium steady state, we have positivity of entropy production, i.e., $\mu_+(s) > 0$, a fact which usually requires some non-trivial analysis. We prove this result in the context of the Lorentz gas under external forces as part of Theorem 2.4.

An important example in the context of this paper is when the state space M is a smooth manifold, μ_0 is a measure with a smooth density with respect to Lebesgue measure on M, and T is a (piecewise) smooth transformation. In this case the change of variable formula gives

$$e^{l_n} = \frac{1}{J_{\mu_0} T^n \circ T^{-n}}\,,$$

where $J_{\mu_0} T$ is the Jacobian of the map T with respect to μ_0 and therefore

$$s = -\log J_{\mu_0} T\,,$$

which can be interpreted as describing a phase space contraction rate. We refer to [JPR] for various other examples.

The fluctuation theorem asserts that the fluctuations of the ergodic averages of the entropy production have a universal symmetry under the condition that the system is invariant under-time reversal.

Definition 1.3. The dynamical system (M, T) with reference measure μ_0 is time-reversal invariant if there exists an involution $i : M \to M$ (that is, $i \circ i$ is the identity) such that,

(1) μ_0 is invariant under i, i.e., $\mu_0(f \circ i) = \mu_0(f)$;
(2) $i \circ T \circ i = T^{-1}$.

Using the time reversal property, we have for any bounded measurable f,

$$\begin{aligned}\mu_0(e^{l_{-n}} f) = \mu_0(f \circ T^{-n}) = \mu_0(f \circ T^{-n} \circ i) = \mu_0(f \circ i \circ T^n) = \mu_0(e^{l_n} f \circ i)\\ = \mu_0(e^{l_n \circ i} f),\end{aligned}$$

and hence

$$l_{-n} = l_n \circ i. \tag{1.2}$$

Using this it is straightforward to derive the so-called transient fluctuation theorem [ES1,JPR] (also called the Evans-Searles fluctuation theorem). We give a proof here for the convenience of the reader.

Proposition 1.4 (Transient fluctuation theorem). *Suppose the dynamical system* (M, T) *with reference measure* μ_0 *is time-reversal invariant and* s *is the entropy production observable. Then we have the symmetry*

$$\mu_0\left(e^{-a\sum_{k=0}^{n-1} s\circ T^k}\right) = \mu_0\left(e^{-(1-a)\sum_{k=0}^{n-1} s\circ T^k}\right),$$

for any $a \in \mathbb{R}$ *for which both integrals are finite.*

Proof. First we use that by the chain rule,

$$l_{-n} = -l_n \circ T^n = -\sum_{k=1}^{n} l_1 \circ T^k = -\sum_{k=0}^{n-1} s \circ T^k.$$

Thus without the assumption of time reversal, we have by (1.1)

$$\mu_0(e^{-a\sum_{k=0}^{n-1} s\circ T^k}) = \mu_0(e^{-al_n\circ T^n}) = \mu_n(e^{-al_n}) = \mu_0(e^{(1-a)l_n}). \tag{1.3}$$

On the other hand time reversal implies by (1.2) that,

$$\mu_0(e^{(1-a)l_n}) = \mu_0(e^{(1-a)l_{-n}\circ i}) = \mu_0(e^{(1-a)l_{-n}}) = \mu_0(e^{-(1-a)\sum_{k=0}^{n-1} s\circ T^k}). \tag{1.4}$$

Combining (1.3) and (1.4) gives the desired symmetry. □

The transient fluctuation theorem has the following interpretation (Proposition 3.3 of [JPR]): if $P_n(z)$ denotes the probability distribution of $\sum_{k=0}^{n-1} s \circ T^k$ with initial distribution μ_0 and $\tau(z) = -z$ then we have

$$\frac{dP_n}{dP_n \circ \tau} = e^{nz},$$

which gives a universal ratio for the probabilities to observe an average entropy production rate equal to $+z$ or $-z$.

By contrast the Gallavotti–Cohen (steady state) fluctuation relation deals with the fluctuation starting in the non-equilibrium steady state μ_+. To state it we define, for any probability measure ν, the logarithmic moment generating function

$$e_\nu(a) = \lim_{n\to\infty} \frac{1}{n} \log \nu\left(e^{-a\sum_{k=0}^{n-1} s\circ T^k}\right),$$

provided the limit exists.

Steady state fluctuation relation. *The dynamical system* (M, T) *with reference measure* μ_0 *and non-equilibrium steady state* μ_+ *satisfies the steady state fluctuation relation if for some* $a_0 > 0$ *and all* $a \in [-a_0, 1 + a_0]$*:*

(1) *the limit defining the logarithmic moment generating function exists,*

$$e_{\mu_+}(a) = \lim_{n\to\infty} \frac{1}{n} \log \mu_+\left(e^{-a\sum_{k=0}^{n-1} s\circ T^k}\right);$$

(2) *the moment generating function has the following symmetry,*

$$e_{\mu_+}(a) = e_{\mu_+}(1-a).$$

The transient and steady state fluctuation relations look similar, yet are distinct statements. In particular, the transient fluctuation theorem is a finite time statement, valid even in the absence of a steady state. Even if we assume that the limit $e_{\mu_0}(a)$ exists (a nontrivial statement), one cannot expect, in general, that $e_{\mu_0}(a) = e_{\mu_+}(a)$ even if μ_+ is a steady state (with reference measure μ_0) (see e.g. [CG] for a counterexample). There certainly are examples where these two functions coincide, e.g. for Anosov diffeormorphisms (see e.g. [JPR]) and indeed one of the main contributions of this paper is to prove that for billiards under small external forces the limits $e_{\mu_0}(a)$ and $e_{\mu_+}(a)$ exist and coincide for a non-perturbative range of values of the parameter a.

To conclude we briefly discuss the large deviation interpretation of the symmetries. From the theory of large deviations, it is well known that if $e_\nu(a)$ is $\mathcal{C}^1$ on an interval $a \in [-a_0, 1+a_0]$, then by the Gartner-Ellis theorem (see [DZe]) we have a large deviation principle for the ergodic averages $\frac{1}{n}\sum_{k=0}^{n-1} s \circ T^k$, with initial condition distributed according to ν, i.e.,

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}\log\nu\left(x : \frac{1}{n}\sum_{k=0}^{n-1} s\circ T^k \in [z-\delta, z+\delta]\right) = -I(z),$$

for any $z \in [e'_\nu(-a_0), e'_\nu(1+a_0)]$, where $I : \mathbb{R} \to [0,\infty]$ is the rate function given by the Legendre transform

$$I(z) = \sup_{-a_0\le a\le 1+a_0}\{az - e_\nu(a)\}.$$

The symmetry $e_\nu(a) = e_\nu(1-a)$ implies that rate function $I(z)$ has the symmetry

$$\begin{aligned} I(z) &= \sup_{-a_0\le a\le 1+a_0}\{az - e_\nu(a)\} = \sup_{-a_0\le a\le 1+a_0}\{az - e_\nu(1-a)\} \\ &= \sup_{-a_0\le b\le 1+a_0}\{(1-b)z - e_\nu(b)\} = I(-z) - z. \end{aligned} \tag{1.5}$$

The symmetry of the rate function $I(z) - I(-z) = -z$ implies that the ratio of probabilities to observe an entropy production rate equal to z and equal to $-z$ over a time interval of length n is asymptotically equal to e^{nz}.

One can also show that the fluctuation relation does imply the Kubo formula for the linear response of currents, but we shall not discuss this further here (see e.g. [LS,M1, M2,JPR]).

2. Description of Model and Main Results

Letting $d \geq 1$, we define a periodic Lorentz gas by placing finitely many closed, convex regions (scatterers) $\Gamma_i, i = 1,\dots d$, on a Torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, which are pairwise disjoint and have $\mathcal{C}^3$ boundaries with strictly positive curvature. The classical billiard flow on the table $\mathbb{T}^2\setminus\cup_i\{\text{interior } \Gamma_i\}$ is defined by the motion of a particle traveling at unit speed and undergoing elastic collisions at the boundaries. In this paper we will also consider the motion of particles subject to external forces, as well as certain types of collisions which do not obey the usual law of reflection.

The discrete-time billiard map T associated with the flow is the Poincaré map corresponding to collisions with the scatterers. At each collision, we record the position according to an arclength parameter r (oriented clockwise on the boundary of each scatterer) and the angle φ made by the outgoing (post-collision) velocity with the unit normal to the boundary at the point of collision. The phase space of the map is thus $M = \cup_{i=1}^{d} I_i \times [-\pi/2, \pi/2]$, where each I_i is an interval with endpoints identified and with length equal to the arclength of $\partial\Gamma_i$.

For any $x = (r, \varphi) \in M$, define $\tau(x)$ to be the free path of the first collision of the trajectory starting at x under the billiard flow. The billiard map is defined wherever $\tau(x) < \infty$. We say that the billiard has finite horizon if there is an upper bound on the function τ. Otherwise, we say the billiard has infinite horizon. Notice that the function τ depends on the (possibly curved) trajectories of particles in $\mathbb{T}^2$, while M is independent of the trajectories; thus we may study many classes of perturbations of a billiard flow while fixing M.

We will denote by $d\mu_0 = c_0 \cos\varphi dr d\varphi$ the smooth invariant probability measure which is preserved by the unperturbed billiard map, where c_0 is the normalizing constant.

2.1. Assumptions. In this subsection we first state the assumptions on the model, following [CZZ] (which in turn combines the assumptions in [CZ,Z,DZ2]).

Let $\mathbf{q} = (x, y)$ be the position of a particle in the billiard table $Q := \mathbb{T}^2 \backslash (\cup_i \Gamma_i)$ and $\mathbf{p}$ be the velocity vector. We may define a perturbed billiard flow on Q as follows. Between collisions, the position and velocity obey the following differential equation,

$$\frac{d\mathbf{q}}{dt} = \mathbf{p}(t), \qquad \frac{d\mathbf{p}}{dt} = \mathbf{F}(\mathbf{q}, \mathbf{p}), \tag{2.1}$$

where $\mathbf{F} : \mathbb{T}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ is a C^2 stationary external force. At collisions, the trajectory experiences possibly nonelastic reflections with slipping along the boundary,

$$(\mathbf{q}^+(t_i), \mathbf{p}^+(t_i)) = (\mathbf{q}^-(t_i), \mathcal{R}\mathbf{p}^-(t_i)) + \mathbf{G}(\mathbf{q}^-(t_i), \mathbf{p}^-(t_i)), \tag{2.2}$$

where $\mathcal{R}\mathbf{p}^-(t_i) = \mathbf{p}^-(t_i) + 2(n(\mathbf{q}^-) \cdot \mathbf{p}^-)n(\mathbf{q}^-)$ is the usual reflection operator, $n(\mathbf{q})$ is the unit normal vector to the billiard wall ∂Q at $\mathbf{q}$ pointing inside the table Q, and $\mathbf{q}^-(t_i)$, $\mathbf{p}^-(t_i)$, $\mathbf{q}^+(t_i)$ and $\mathbf{p}^+(t_i)$ refer to the incoming and outgoing position and velocity vectors, respectively.[1] $\mathbf{G}$ is an external force acting on the incoming trajectories. We allow $\mathbf{G}$ to change both the position and the velocity of the particle at the moment of collision. The change in velocity can be thought of as a kick or twist while a change in position can model a slip along the boundary at collision, or even reflection by a soft billiard potential [BT].

In [Ch2,Ch4], Chernov considered billiards under small external forces $\mathbf{F}$ with $\mathbf{G} = 0$, and $\mathbf{F}$ to be stationary. In [Z] a twist force was considered assuming $\mathbf{F} = 0$ and $\mathbf{G}$ depending on and affecting only the velocity, not the position. Here we follow [DZ2, CZZ] and consider a combination of these two cases for systems under more general forces $\mathbf{F}$ and $\mathbf{G}$.

Let $\mathbf{E} = (\mathbf{F}, \mathbf{G})$, where $\mathbf{F}$ and $\mathbf{G}$ are the two external forces during the flight and at collisions, respectively. Let $\Phi^t_{\mathbf{E}}$ be the induced billiard flow on $Q \times \mathbb{R}^2$ and denote by $T_{\mathbf{E}} = T_{\mathbf{F},\mathbf{G}}$ the corresponding billiard map.

(A1) (Invariant space) *The perturbed flow $\Phi^t_{\mathbf{E}}$ preserves a smooth function $\mathcal{E}(\mathbf{q}, \mathbf{p})$, such that the level surface $\mathcal{M} := \{\mathcal{E}(\mathbf{q}, \mathbf{p}) = c\}$ is a compact 3-D manifold, for some $c > 0$.*

[1] Since we identify $\mathbb{T}^2$ with $\mathbb{R}^2/\mathbb{Z}^2$, we define addition of vectors $(\mathbf{q}, \mathbf{p}) \in \mathbb{T}^2 \times \mathbb{R}^2$ as addition mod 1 in each coordinate of $\mathbf{q}$ and standard vector addition for $\mathbf{p}$.

Moreover, $\|\mathbf{p}\| > 0$ on $\mathcal{M}$, and for each $\mathbf{q} \in Q$ and $\mathbf{p} \in S^1$, the ray $\{(\mathbf{q}, t\mathbf{p}), t > 0\}$ intersects the manifold $\mathcal{M}$ in exactly one point.

Under assumption (**A1**), the system has an additional integral of motion and we will consider the restricted system on a compact phase space, $\mathcal{M} \subset Q \times \mathbb{R}^2$. For example, if we add a Gaussian thermostat (a heat bath) to the system such that the billiard moves at constant speed (constant temperature if there are a large number of particles), then $\mathcal{M} := \{\|\mathbf{p}\| = c\}$ is an invariant compact level set. More generally, the speed $p = \|\mathbf{p}\|$ of the billiard along all trajectories on $\mathcal{M}$ at time t satisfies

$$0 < p_{\min} \le p(t) \le p_{\max} < \infty,$$

for some constants $p_{\min} \le p_{\max}$. In addition, $\mathcal{M}$ admits a global coordinate system $\{(x, y, \theta) : (x, y) \in Q, 0 \le \theta < 2\pi\}$, where θ is the angle between $\mathbf{p}$ and the positive x-axis. Thus the speed $p = \|\mathbf{p}\|$ on $\mathcal{M}$ can be represented as a function $p = p(x, y, \theta)$ and the velocity $\mathbf{p}$ at $\mathbf{q}$ can be expressed as $\mathbf{p} = p\mathbf{v}$, where $\mathbf{v} = (\cos\theta, \sin\theta)$ is the unit vector in the direction of $\mathbf{p}$. We can then rewrite Eq. (2.1) for the dynamics between collisions as

$$\dot{\mathbf{q}} = \mathbf{p}, \quad \dot{p}\mathbf{v} + p\dot{\mathbf{v}} = \mathbf{F}. \tag{2.3}$$

Multiplying both sides of the second equation in (2.3) by $\mathbf{v}$ using the dot product and cross product respectively, we obtain

$$\dot{p} = \mathbf{v} \cdot \mathbf{F}, \quad \text{and} \quad p\mathbf{v} \times \dot{\mathbf{v}} = \mathbf{v} \times \mathbf{F}. \tag{2.4}$$

Therefore, using the notation $\mathbf{F} = (F_1, F_2)$, the equations in (2.1) have the following coordinate representations at any $(x, y, \theta) \in \mathcal{M}$,

$$\begin{cases} \dot{x} = p\cos\theta, \\ \dot{y} = p\sin\theta, \\ \dot{\theta} = (-F_1 \sin\theta + F_2 \cos\theta)/p. \end{cases} \tag{2.5}$$

Next, consider a trajectory $\tilde{\gamma} \subset \mathcal{M}$ of the flow passing through the point $(x, y, \theta) \in \mathcal{M}$, which projects down to a smooth curve $\gamma \subset Q$. We denote by $\kappa = \kappa(x, y, \theta)$ the (signed) geometric curvature of γ at $(x, y) \in Q$. It follows that

$$\kappa(x, y, \theta) = \pm\frac{\|\dot{\mathbf{q}} \times \ddot{\mathbf{q}}\|}{\|\dot{\mathbf{q}}\|^3} = \pm\frac{\|\mathbf{v} \times \mathbf{F}\|}{p^2} = \frac{-F_1 \sin\theta + F_2 \cos\theta}{p^2}, \tag{2.6}$$

where the sign should be chosen accordingly. Combining this with (2.5), we have

$$\dot{\theta} = p\kappa. \tag{2.7}$$

Note that the angle $\theta = \theta(t)$ is discontinuous at reflection times: it jumps from θ^- to θ^+. In the case of elastic collisions, the quantities x, y and p remain unchanged. By contrast, under the twisting force $\mathbf{G}$, all quantities may change at collisions.

For any point $(x, y, \theta) \in \mathcal{M}$, let $\tau(x, y, \theta)$ be the time for the trajectory starting from (x, y, θ) to make its next non-tangential collision at ∂Q.

(**A2**) (**Finite horizon**) *There exist $\tau_{\max} > \tau_{\min} > 0$ such that free paths between successive non-tangential reflections are uniformly bounded: $\tau_{\min} \le \tau(x, y, \theta) \le \tau_{\max}$, for all $(x, y, \theta) \in \mathcal{M}$ with $(x, y) \in \partial Q$.*

(**A3**) (**Smallness of the external forces**). *There exists* $\varepsilon > 0$ *small enough such that the forces* $\mathbf{E} = (\mathbf{F}, \mathbf{G})$ *satisfy*

$$\|\mathbf{F}\|_{C^1} < \varepsilon, \quad \|\mathbf{G}\|_{C^1} < \varepsilon.$$

Moreover, there exist constants $\alpha_0 > 1/3$ *and* $C_{\mathbf{E}} > 0$ *such that* $\|\mathbf{F}\|_{C^{1+\alpha_0}}, \|\mathbf{G}\|_{C^{1+\alpha_0}} \le C_{\mathbf{E}}$.

Remark 2.1. Note that (**A2**) also puts some implicit constraints on the smallness of forces. In fact, the existence of $\tau_{\min}$ not only prevents touching scatterers, but also implies the trajectory cannot be bent too much such that the particle falls back to the same scatterer immediately.

Let $\mathcal{I} : \mathcal{M} \to \mathcal{M}$ be the involution defined by $\mathcal{I}(x, y, \theta) = (x, y, \pi + \theta)$. For a general flow $\Phi^t : \mathcal{M} \to \mathcal{M}$, the reversed flow of Φ^t is defined by $\Phi^t_- = \mathcal{I} \circ \Phi^{-t} \circ \mathcal{I}$. The flow Φ^t is said to be *time-reversible*, if $\Phi^t_- = \Phi^t$. It is well known that the unforced billiard flow is time-reversible.

(**A4**) (**Time-reversibility**) *Both forces* **F** *and* **G** *are stationary, and the forced billiard flow* $\Phi^t_{\mathbf{E}}$ *is time-reversible. Moreover, we assume that the addition of* **G** *preserves tangential collisions:* $\mathbf{G}(r, \pm\frac{\pi}{2}) = (0, 0)$.

Note that due to (**A4**), the singularity set of $T^{-1}_{\mathbf{F},\mathbf{G}}$ is the same as that of the untwisted map $T^{-1}_{\mathbf{F},\mathbf{0}}$. It also implies that the billiard map $T_{\mathbf{E}}$ is time-reversible.

Fix $\varepsilon_0 > 0$, $\tau_* \in (0, 1)$, and $C_0 > 0$. For the fixed billiard table Q, let $\mathcal{F}(\varepsilon_0, \tau_*, C_0)$ denote the collection of all forced billiard maps defined by the dynamics (2.1) and (2.2) under the external forces $\mathbf{E} = (\mathbf{F}, \mathbf{G})$ and satisfying assumptions (**A1**)–(**A4**), such that $\tau_* \le \tau_{\min} \le \tau_{\max} \le \tau_*^{-1}$, $C_{\mathbf{E}} \le C_0$, and $\varepsilon \le \varepsilon_0$ in (**A3**).

In Sect. 2.2.1 we define a class of maps satisfying uniform properties regarding hyperbolicity and singularities, **(H1)–(H5)**. The following lemma from [DZ2] is crucial in that respect.

Lemma 2.2 ([DZ2, Theorem 2.10]). *Fix* $\tau_* \in (0, 1)$. *There exist* $\varepsilon_0, C_0 > 0$ *such that the family of maps* $\mathcal{F}(\varepsilon_0, \tau_*, C_0)$ *satisfy* ***(H1)–(H5)*** *with uniform constants.*

2.2. Abstract framework. In this section, we identify a set of uniform properties **(H1)–(H5)** enjoyed by the class of perturbed billiard maps defined in Sect. 2.1; these properties guarantee the Lasota–Yorke inequalities (2.17) with uniform constants. These conditions are a simplified version of the abstract framework appearing in [DZ2] since here we consider only finite horizon billiards, so the technical difficulties associated with the infinite horizon case are excluded.

We also introduce general conditions **(C1)–(C4)** to verify that a perturbation is small in the sense required for Theorem 2.3. These conditions are sufficient to establish the framework of [KL]. As mentioned above, the fact that the specific classes of perturbations we consider in Sect. 2.1 satisfy **(H1)–(H5)** follows from Lemma 2.2.

2.2.1. A class of maps with uniform properties. We fix the phase space $M = \cup_{i=1}^d I_i \times [-\frac{\pi}{2}, \frac{\pi}{2}]$ of a billiard map associated with a periodic Lorentz gas as in Sect. 2.4. We will denote (normalized) Lebesgue measure on M by m, i.e., $dm = \frac{1}{\pi L} dr d\varphi$, where $L = \sum_{i=1}^d |I_i|$.

We define the set $\mathcal{S}_0 = \{\varphi = \pm\frac{\pi}{2}\}$ and for a fixed $k_0 \in \mathbb{N}$, we define for $k \ge k_0$, the homogeneity strips,

$$\mathbb{H}_k = \{(r,\varphi) : \pi/2 - k^{-2} < \varphi < \pi/2 - (k+1)^{-2}\}. \tag{2.8}$$

The strips $\mathbb{H}_{-k}$ are defined similarly near $\varphi = -\pi/2$. We also define $\mathbb{H}_0 = \{(r,\varphi) : -\pi/2 + k_0^{-2} < \varphi < \pi/2 - k_0^{-2}\}$. The set $\mathcal{S}_{0,H} = \mathcal{S}_0 \cup (\cup_{|k|\geq k_0} \partial\mathbb{H}_{\pm k})$ is therefore fixed and will give rise to the singularity sets for the maps that we define below, i.e. for any map T that we consider, we define $\mathcal{S}^T_{\pm n} = \cup_{i=0}^n T^{\mp i}\mathcal{S}_0$ to be the singularity sets for $T^{\pm n}$, $n \geq 0$. We assume that $\mathcal{S}^T_{\pm n}$ comprises finitely many smooth curves for each $n \in \mathbb{N}$. We also define the extended singularity sets $\mathcal{S}^{T,\mathbb{H}}_{\pm n} = \cup_{i=0}^n T^{\mp i}\mathcal{S}_{0,H}$ to include the boundaries of the homogeneity strips. When the map T is fixed, we sometimes write $\mathcal{S}^{\mathbb{H}}_{\pm n}$ to simplify notation.

Suppose there exists a class of invertible maps $\mathcal{F}$ such that for each $T \in \mathcal{F}$, $T : M\setminus\mathcal{S}^T_1 \to M\setminus\mathcal{S}^T_{-1}$ is a C^2 diffeomorphism on each connected component of $M\setminus\mathcal{S}^T_1$. We assume that elements of $\mathcal{F}$ enjoy the following uniform properties.

(H1) Hyperbolicity and singularities. There exist continuous families of stable and unstable cones $C^s(x)$ and $C^u(x)$, defined on all of M, which are strictly invariant for the class $\mathcal{F}$, i.e., $DT(x)C^u(x) \subset C^u(Tx)$ and $DT^{-1}(x)C^s(x) \subset C^s(T^{-1}x)$ for all $T \in \mathcal{F}$ wherever DT and DT^{-1} are defined.

The cones $C^s(x)$ and $C^u(x)$ are uniformly transverse on M and $\mathcal{S}^T_{-n}$ is uniformly transverse to $C^s(x)$ for each $n \in \mathbb{N}$ and all $T \in \mathcal{F}$. We assume in addition that $C^s(x)$ is uniformly transverse to the horizontal and vertical directions on all of M.[2]

Moreover, there exist constants $C_e > 0$ and $\Lambda > 1$ such that for all $T \in \mathcal{F}$,

$$\|DT^n(x)v\| \geq C_e^{-1}\Lambda^n\|v\|, \forall v \in C^u(x), \quad \text{and} \quad \|DT^{-n}(x)v\| \geq C_e^{-1}\Lambda^n\|v\|, \forall v \in C^s(x), \tag{2.9}$$

for all $n \geq 0$, where $\|\cdot\|$ is the Euclidean norm on the tangent space $\mathcal{T}_x M$.

We also assume a similar unbounded expansion in a neighborhood of $\mathcal{S}_0$. We assume there exists $C_c > 0$ such that

$$C_c[\cos\varphi(T^{-1}x)]^{-1}\|v\| \leq \|DT^{-1}(x)v\| \leq C_c^{-1}[\cos\varphi(T^{-1}x)]^{-1}\|v\|, \\ \forall x \in M\setminus\mathcal{S}^T_{-1}, \forall v \in C^s(x), \tag{2.10}$$

where $\varphi(y)$ denotes the angle at the point $y = (r,\varphi) \in M$. Let $\exp_x$ denote the exponential map from $\mathcal{T}_x M$ to M. We require the following bound on the second derivative,

$$C_c[\cos\varphi(T^{-1}x)]^{-3} \leq \|D^2T^{-1}(x)v\| \leq C_c^{-1}[\cos\varphi(T^{-1}x)]^{-3}, \quad \forall x \in M\setminus\mathcal{S}^T_{-1}, \tag{2.11}$$

for all $v \in \mathcal{T}_x M$ such that $T^{-1}(\exp_x(v))$ and $T^{-1}x$ lie in the same homogeneity strip.

(H2) Families of stable and unstable curves. We call W a *stable curve* for a map $T \in \mathcal{F}$ if the tangent line to W, $\mathcal{T}_x W$ lies in $C^s(x)$ for all $x \in W$. We call W *homogeneous* if W is contained in one homogeneity strip $\mathbb{H}_k$. Unstable curves are defined similarly. Let $\widehat{\mathcal{W}}^s$ denote the set of $\mathcal{C}^2$ homogeneous stable curves in M whose curvature is bounded above by a uniform constant $B > 0$. We assume there exists a choice of B such that $\widehat{\mathcal{W}}^s$ is invariant under $\mathcal{F}$ in the following sense: For any $W \in \widehat{\mathcal{W}}^s$ and $T \in \mathcal{F}$, the connected components of $T^{-1}W$ are again elements of $\widehat{\mathcal{W}}^s$. A family of unstable curves $\widehat{\mathcal{W}}^u$ is

[2] This is not a restrictive assumption for perturbations of the Lorentz gas since the standard cones $\hat{C}^s$ and $\hat{C}^u$ for the billiard map satisfy this property (see for example [CM, Section 4.5]); the common cones $C^s(x)$ and $C^u(x)$ shared by all maps in the class $\mathcal{F}$ must therefore lie inside $\hat{C}^s(x)$ and $\hat{C}^u(x)$ and therefore satisfy this property.

defined analogously, with obvious modifications: For example, we require the connected components of TW to be elements of $\widehat{\mathcal{W}}^u$ for all $W \in \widehat{\mathcal{W}}^u$ and $T \in \mathcal{F}$.

(H3) Complexity bounds (One-step expansion).[3] We assume that there exists an adapted norm $\|\cdot\|_*$, uniformly equivalent to $\|\cdot\|$, in which the constant C_e in (2.9) can be taken to be 1, i.e. we have expansion and contraction in one step in the adapted norm for all maps in the class $\mathcal{F}$ (for example, the norm from [CM, Sect. 5.10]).

Let $W \in \widehat{W}^s$. For any $T \in \mathcal{F}$, we partition the connected components of $T^{-1}W$ into maximal pieces $V_i = V_i(T)$ such that each V_i is a homogeneous stable curve in some $\mathbb{H}_k$, $k \geq k_0$, or $\mathbb{H}_0$. Let $|J_{V_i}T|_*$ denote the minimum contraction on V_i under T in the metric induced by the adapted norm $\|\cdot\|_*$. We assume that for some choice of k_0,

$$\limsup_{\delta\to 0} \sup_{T\in\mathcal{F}} \sup_{|W|<\delta} \sum_i |J_{V_i}T|_* < 1, \tag{2.12}$$

where $|W|$ denotes the arclength of W.

(H4) Bounded distortion. There exists a constant $C_d > 0$ with the following properties. Let $W' \in \widehat{\mathcal{W}}^s$ and for any $T \in \mathcal{F}$, $n \in \mathbb{N}$, let $x, y \in W$ for some connected component $W \subset T^{-n}W'$ such that T^iW is a homogeneous stable curve for each $0 \leq i \leq n$. Then,

$$\left|\frac{J_{\mu_0}T^n(x)}{J_{\mu_0}T^n(y)} - 1\right| \;\leq\; C_d d_W(x,y)^{1/3} \quad\text{and}\quad \left|\frac{J_W T^n(x)}{J_W T^n(y)} - 1\right| \;\leq\; C_d d_W(x,y)^{1/3}, \tag{2.13}$$

where as before $J_{\mu_0}T^n$ is the Jacobian of T^n with respect to the smooth measure $d\mu_0 = c\cos\varphi dr d\varphi$.

We assume the analogous bound along unstable leaves: If $W \in \widehat{\mathcal{W}}^u$ is an unstable curve such that T^iW is a homogeneous unstable curve for $0 \leq i \leq n$, then for any $x, y \in W$,

$$\left|\frac{J_{\mu_0}T^n(x)}{J_{\mu_0}T^n(y)} - 1\right| \;\leq\; C_d d(T^n x, T^n y)^{1/3}. \tag{2.14}$$

(H5) Control of Jacobian. Let $\beta, \gamma, p < 1$ be from the definition of the norms in Sect. 3 and let $\theta_* < 1$ be from (2.15). Assume there exists a constant $1 \leq \eta < \min\{\Lambda^\beta, \Lambda^\gamma, \theta_*^{p-1}\}$ such that for any $T \in \mathcal{F}$,

$$(J_{\mu_0}T(x))^{-1} \leq \eta \qquad \text{wherever } J_{\mu_0}T \text{ is defined.}$$

Recall the family of stable curves $\widehat{\mathcal{W}}^s$ defined by **(H2)**. We define a subset $\mathcal{W}^s \subset \widehat{\mathcal{W}}^s$ as follows. By **(H3)** we may choose $\delta_0 > 0$ for which there exists $\theta_* < 1$ such that

$$\sup_{T\in\mathcal{F}} \sup_{|W|\leq\delta_0} \sum_i |J_{V_i}T|_* \leq \theta_*. \tag{2.15}$$

We shrink δ_0 further if necessary so that the graph transform argument needed in the proof of Lemma A.2(a) holds. The set $\mathcal{W}^s$ comprises all those stable curves $W \in \widehat{\mathcal{W}}^s$ such that $|W| \leq \delta_0$.

---

[3] In [DZ2], a 'weakened one-step expansion' was also assumed: $\limsup_{\delta\to 0} \sup_{T\in\mathcal{F}} \sup_{|W|<\delta} \sum_i |J_{V_i}T|^\varsigma < \infty$ for some $\varsigma < 1$, where the norm of the Jacobian is measured in the Euclidean norm. Since here we restrict to finite horizon, however, this property follows from **(H1)**.

2.2.2. Distance in $\mathcal{F}$. We define a distance in $\mathcal{F}$ as follows. For $T_1, T_2 \in \mathcal{F}$ and $\varepsilon > 0$, let $N_\varepsilon(\mathcal{S}^i_{-1})$ denote the ε-neighborhood in M of the singularity set $\mathcal{S}^i_{-1}$ of T_i^{-1}, $i = 1, 2$. We say $d_\mathcal{F}(T_1, T_2) = \varepsilon'$ if the maps are close away from their singularity sets in the following sense: ε' is the infimum over $\varepsilon > 0$ such that for all $x \notin N_\varepsilon(\mathcal{S}^1_{-1} \cup \mathcal{S}^2_{-1})$,

(C1) $d(T_1^{-1}(x), T_2^{-1}(x)) \le \varepsilon$;

(C2) $\left|\dfrac{J_{\mu_0}T_i(x)}{J_{\mu_0}T_j(x)} - 1\right| \le \varepsilon, i, j = 1, 2;$

(C3) $\left|\dfrac{J_W T_i(x)}{J_W T_j(x)} - 1\right| \le \varepsilon,$ for any $W \in \mathcal{W}^s$, $i, j = 1, 2$, and $x \in W$;

(C4) $\|DT_1^{-1}(x)v - DT_2^{-1}(x)v\| \le \sqrt{\varepsilon}$, for any unit vector $v \in \mathcal{T}_x W$, $W \in \mathcal{W}^s$.

We remark that while this notion of distance requires T_1 and T_2 to be $\mathcal{C}^1$-close outside an ε-neighborhood of $\mathcal{S}^1_{-1} \cup \mathcal{S}^2_{-1}$, it does not require $\mathcal{S}^1_{-1}$ and $\mathcal{S}^2_{-1}$ to be close as subsets of M.

Due to the exclusion of $x \in N_\varepsilon(\mathcal{S}^1_{-1} \cup \mathcal{S}^2_{-1})$, our notion of distance between maps does not satisfy the triangle inequality: to compare the relevant quantities for 3 maps T_1, T_2 and T_3, we would have to exclude the ε-neighborhoods of all three singularity sets. Nevertheless, as in [DZ2], this does not create problems in our use of this distance: we fix a map T_0 for which the associated transfer operator has a spectral gap, then compare transfer operators for maps T near to T_0 with respect to the quantity $d_\mathcal{F}(T, T_0)$. Lemma 4.6 summarizes the key use of this distance.

2.3. Transfer operators. In this section, we fix a class of maps $\mathcal{F}$ with uniform properties **(H1)**–**(H5)** as defined Sect. 2.2.1. Later, we will specialize to a particular family $\mathcal{F} = \mathcal{F}(\varepsilon, \tau_*, C_0)$ satisfying **(A1)**–**(A4)** above.

Let $\widehat{\mathcal{W}^s}$ be the set of stable curves invariant under maps in $\mathcal{F}$ according to **(H2)**, and let $\mathcal{W}^s \subset \widehat{\mathcal{W}^s}$ denote those stable curves having length less than δ_0, where δ_0 is from (2.15). For any $T \in \mathcal{F}$, we define scales of spaces using the set of stable curves $\mathcal{W}^s$ on which the *transfer operator* $\mathcal{L}_T$ associated with T will act. Define $T^{-n}\mathcal{W}^s$ to be the set of homogeneous stable curves W such that T^n is smooth on W and $T^i W \in \mathcal{W}^s$ for $0 \le i \le n$. It follows from **(H2)** that $T^{-n}\mathcal{W}^s \subset \mathcal{W}^s$.

For $W \in T^{-n}\mathcal{W}^s$, a complex-valued test function $\psi : M \to \mathbb{C}$, and $0 < \alpha \le 1$ define $H^\alpha_W(\psi)$ to be the Hölder constant of ψ on W with exponent α measured in the Euclidean metric. Define $H^\alpha_n(\psi) = \sup_{W \in T^{-n}\mathcal{W}^s} H^\alpha_W(\psi)$ and let $\tilde{\mathcal{C}}^\alpha(T^{-n}\mathcal{W}^s) = \{\psi : M \to \mathbb{C} \mid H^\alpha_n(\psi) < \infty\}$, denote the set of complex-valued functions which are Hölder continuous on elements of $T^{-n}\mathcal{W}^s$. The set $\tilde{\mathcal{C}}^\alpha(T^{-n}\mathcal{W}^s)$ equipped with the norm $|\psi|_{\mathcal{C}^\alpha(T^{-n}\mathcal{W}^s)} = |\psi|_\infty + H^\alpha_n(\psi)$ is a Banach space. Similarly, we define $\tilde{\mathcal{C}}^\alpha(\widehat{\mathcal{W}}^u)$ to be the set of functions which are Hölder continuous with exponent α on unstable curves $\widehat{\mathcal{W}}^u$.

It follows from the uniform hyperbolicity of T (see **(H1)**) that if $\psi \in \tilde{\mathcal{C}}^\alpha(T^{-(n-1)}\mathcal{W}^s)$, then $\psi \circ T \in \tilde{\mathcal{C}}^\alpha(T^{-n}\mathcal{W}^s)$. Thus if $h \in (\tilde{\mathcal{C}}^\alpha(T^{-n}\mathcal{W}^s))'$, is an element of the dual of $\tilde{\mathcal{C}}^\alpha(T^{-n}\mathcal{W}^s)$, then $\mathcal{L}_T : (\tilde{\mathcal{C}}^\alpha(T^{-n}\mathcal{W}^s))' \to (\tilde{\mathcal{C}}^\alpha(T^{-(n-1)}\mathcal{W}^s))'$ acts on h by

$$\mathcal{L}_T h(\psi) := h(\psi \circ T) \quad \forall \psi \in \tilde{\mathcal{C}}^\alpha(T^{-(n-1)}\mathcal{W}^s).$$

Recall that $d\mu_0 = c \cos\varphi dr d\varphi$ denotes the smooth invariant measure for the billiard map corresponding to the unperturbed periodic Lorentz gas. If $h \in L^1(M, \mu_0)$, then h is canonically identified with a signed measure absolutely continuous with respect to μ_0, which we shall also call h, i.e., $h(\psi) = \int_M \psi h \, d\mu_0$. With the above identification, we write $L^1(M, \mu_0) \subset (\tilde{\mathcal{C}}^\alpha(T^{-n}\mathcal{W}^s))'$ for each $n \in \mathbb{N}$. Then restricted to $L^1(M, \mu_0)$, $\mathcal{L}_T$ acts according to the familiar expression

$$\mathcal{L}_T^n h = \frac{h \circ T^{-n}}{J_{\mu_0} T^n \circ T^{-n}} \quad \text{for any } n \geq 0 \text{ and } h \in L^1(M, \mu_0),$$

where $J_{\mu_0} T$ is the Jacobian of T with respect to μ_0.

In Sect. 3, we define Banach spaces of distributions $(\mathcal{B}, \| \cdot \|_{\mathcal{B}})$ and $(\mathcal{B}_w, | \cdot |_w)$, preserved under the action of $\mathcal{L}_T$, such that the unit ball of $\mathcal{B}$ is compactly embedded in $\mathcal{B}_w$. It follows from [DZ2, Corollary 2.4] that for ε sufficiently small, $\mathcal{L}_T$ has a spectral gap on $\mathcal{B}$.

To study large deviations we will need a suitable weighted transfer operator. In order to have a well defined operator on $\mathcal{B}$ we will assume that $g : M \to \mathbb{R}$ is (piecewise) Hölder continous on the connected components of $M \backslash \mathcal{S}_1^T$ where $\mathcal{S}_1^T$ is the set of discontinuities of T (see Sects. 2.2 and 3.1 for details). Under these assumptions it is shown in Lemma 3.3 that we can define the weighted transfer operator $\mathcal{L}_{T,g}$ associated with T and g on $\mathcal{B}$ and $\mathcal{B}_w$ by

$$\mathcal{L}_{T,g} h(\psi) := \mathcal{L}_T(he^g)(\psi) = h(e^g \cdot \psi \circ T), \quad \text{for } h \in \mathcal{B}_w \text{ and suitable test functions } \psi. \tag{2.16}$$

The family of transfer operators $\mathcal{L}_{T,ag}$ parametrized by $a \in \mathbb{R}$ occurs naturally in studying the large deviations of Birkhoff sums $S_n g = g + \cdots + g \circ T^{n-1}$: since we have

$$\mathcal{L}_{T,ag}^n h(\psi) = h(e^{aS_n g} \psi \circ T^n) \,,$$

the logarithmic moment generating function of $S_n g$ with initial distribution $\nu \in \mathcal{B}$ is then given by

$$\log \nu(e^{aS_n g}) = \log \mathcal{L}_{T,ag}^n \nu(1).$$

Suitable spectral gap conditions on $\mathcal{L}_{T,ag}$ imply that the limit

$$e_\nu(a) = \lim_{n \to \infty} \frac{1}{n} \log \nu(e^{aS_n g}) = \lim_{n \to \infty} \frac{1}{n} \log \mathcal{L}_{T,ag}^n \nu(1)$$

exists and is smooth and then large deviation estimates follow from the Gärtner-Ellis theorem [DZe]. In this paper we shall be interested in particular in the choices $\nu = \mu_0$, the SRB measure for the unperturbed Lorentz gas, and $\nu = \mu_{\mathbf{E}}$, the SRB measure for the perturbed Lorentz gas $T_{\mathbf{E}}$, both measures belonging to $\mathcal{B}$.

2.4. Statement of results. In [DZ2], local large deviation estimates for (piecewise) smooth observables g were obtained for small ε (small forces) and small a (deviations very close to the mean of g); these were essentially perturbative results in a and ε. By contrast here we concentrate on the observable $s = -\log J_{\mu_0} T_{\mathbf{E}}$, which is the entropy production observable defined in Sect. 1.1. For the fluctuation symmetry to make sense we will need the moment generating function to be well-defined for a in a neighborhood of $[0, 1]$. To this end, we will fix $a_0 > 0$ and consider the interval $a \in [-a_0, 1 + a_0]$. We study the dependence of the spectral gap of $\mathcal{L}_{T,-as}$ as a function of the two parameters, ε and a. Since s is fixed, in what follows we will use the more concise notation,

$\mathcal{L}_{T,a} = \mathcal{L}_{T,-as}$. Note also that in the absence of external forces, μ_0 is an invariant measure and $J_{\mu_0}T_0 = 1$. More generally, for $T_{\mathbf{E}} = T_{(\mathbf{F},\mathbf{G})}$, we show in Lemma 4.2 that

$$J_{\mu_0}T_{\mathbf{E}} = 1 + \varepsilon H,$$

where H is bounded uniformly in ε, a key fact in our analysis.

The following spectral result is key to proving the existence and smoothness of the limiting logarithmic moment generating function.

Theorem 2.3 (Spectral gap). *Choose $a_0 > 0$ and fix the parameters C_0, τ_* from Sect. 2.1. There exists $\varepsilon_0 > 0$ such that for any $T \in \mathcal{F} := \mathcal{F}(\varepsilon_0, \tau_*, C_0)$, the operator $\mathcal{L}_{T,a}$ is well defined as a bounded linear operator on $\mathcal{B}$ for all $a \in [-a_0, 1 + a_0]$. In addition, there exists $C > 0$, such that for any $T \in \mathcal{F}$ and $n \geq 0$,*

$$\begin{aligned} |\mathcal{L}^n_{T,a}h|_w &\leq C(1 + sign(a-1)C_H\varepsilon)^{n(a-1)}|h|_w \qquad \text{for all } h \in \mathcal{B}_w, \\ \|\mathcal{L}^n_{T,a}h\|_{\mathcal{B}} &\leq C\sigma^n(1 + sign(a-1)C_H\varepsilon)^{n(a-1)}\|h\|_{\mathcal{B}} + C\eta^n|h|_w \qquad \text{for all } h \in \mathcal{B}, \end{aligned} \tag{2.17}$$

where $C_H > 0$ is from Lemma 4.2 and $\sigma \in (0, 1)$ is from (4.4). Moreover, for each $T \in \mathcal{F}$,

(i) *$\mathcal{L}_{T,a}$ is quasi-compact as an operator on $\mathcal{B}$: The spectral radius $\rho(\mathcal{L}_{T,a})$ lies in $[(1-sign(a-1)C_H\varepsilon_0)^{a-1}, (1+sign(a-1)C_H\varepsilon_0)^{a-1}]$, while the essential spectral radius $\rho_{ess}(\mathcal{L}_{T,a})$ is at most $\sigma(1+sign(a-1)C_H\varepsilon_0)^{a-1} < (1-sign(a-1)C_H\varepsilon_0)^{a-1}$.*
(ii) *There exists $\varepsilon_1 \leq \varepsilon_0$ such that for all $T \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$ and all $a \in [-a_0, 1 + a_0]$, $\mathcal{L}_{T,a}$ has a spectral gap: there exists exactly one simple real eigenvalue $\lambda_a = \rho(\mathcal{L}_{T,a})$; the corresponding eigenfunction h_a is a positive Borel measure.*

For $T_{\mathbf{E}} \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$, we discuss next the existence and properties of the logarithmic moment generating function for the entropy production observable $s = -\log J_{\mu_0}T_{\mathbf{E}}$ with respect to the non-equilibrium steady state $\mu_{\mathbf{E}}$,

$$e_{\mathbf{E}}(a) = \lim_{n\to\infty} \frac{1}{n} \log \mu_{\mathbf{E}}\left((J_{\mu_0}T^n_{\mathbf{E}})^a\right). \tag{2.18}$$

We denote by $\sigma^2_{\mathbf{E}}$ the diffusion constant for the sequence $\{\log J_{\mu_0}T_{\mathbf{E}} \circ T^n_{\mathbf{E}}\}_{n\geq 0}$ distributed according to $\mu_{\mathbf{E}}$, and by σ^2_H the diffusion constant for the sequence $\{H \circ T^n_0\}_{n\geq 0}$ distributed according to μ_0. Our main results are summarized in the following theorem.

Theorem 2.4 (Logarithmic moment generating function and fluctuation relation). *Under the assumptions of Theorem 2.3, we have the following.*

(1) *The map T_E has a unique SRB measure (non-equilibrium steady state) $\mu_{\mathbf{E}}$.*
(2) *The logarithmic moment generating function $e_{\mathbf{E}}(a)$ for the entropy production exists and is analytic in the disk $|a| \leq 1 + a_0$. Moreover we have*

$$e_{\mathbf{E}}(a) = \lim_{n\to\infty} \frac{1}{n} \log \mu_{\mathbf{E}}\left((J_{\mu_0}T^n_{\mathbf{E}})^a\right) = \lim_{n\to\infty} \frac{1}{n} \log \mu_0\left((J_{\mu_0}T^n_{\mathbf{E}})^a\right)$$

and, as a consequence, for $a \in [-a_0, 1 + a_0]$ we have the non-equilibrium steady state fluctuation relation

$$e_{\mathbf{E}}(a) = e_{\mathbf{E}}(1 - a). \tag{2.19}$$

(3) *The logarithmic moment generating function* $e_{\mathbf{E}}(a)$ *is strictly convex if and only if* $\log J_{\mu_0} T_{\mathbf{E}}$ *is not a coboundary for some* $\psi \in L^2(\mu_{\mathbf{E}})$*, in which case we have*

$$0 > e'_{\mathbf{E}}(0) = \mu_{\mathbf{E}}(\log J_{\mu_0} T_{\mathbf{E}}) = \varepsilon \mu_0(H) + o(\varepsilon) \text{ (Positivity of entropy production)},$$

and

$$0 < e''_{\mathbf{E}}(0) = \sigma^2_{\mathbf{E}} = \sigma^2_H \varepsilon^2 + o(\varepsilon^2) \text{ (Positivity of diffusion coefficients)}.$$

Remark 2.5. The expansion of $\mu_{\mathbf{E}}(\log J_{\mu_0} T_{\mathbf{E}})$ in item (3) of Theorem 2.4 is related to the linear response of the periodic Lorentz gas to the external forces $\mathbf{E} = (\mathbf{F}, \mathbf{G})$. For more explicit relations valid for this class of perturbations, see [CELS2,CZZ].

We prove Theorem 2.4 in Sect. 5. The main technical elements in the proof are first to establish the spectral gap, and then to derive the existence of the relevant limit(s) and the analyticity of the moment generating function. The proof of strict convexity also requires substantial work related to the Central Limit Theorem. Once these two properties are established, the fluctuation relation (2.19) follows immediately from the transient fluctuation relation, Proposition 1.4.

By using standard large deviation techniques [DZe] we obtain immediately a version of the Gallavotti–Cohen fluctuation theorem.

Theorem 2.6. *Under the assumptions of Theorem 2.3, for all* $z \in [e'_{\mathbf{E}}(-a_0), e'_{\mathbf{E}}(1+a_0)]$*, we have*

$$\lim_{\delta \to 0} \lim_{n \to \infty} \frac{1}{n} \log \frac{\mu_{\mathbf{E}}\Big(x : \frac{1}{n} S_n s(x) \in [z - \delta, z + \delta]\Big)}{\mu_{\mathbf{E}}\Big(x : \frac{1}{n} S_n s(x) \in [-z - \delta, -z + \delta]\Big)} = z.$$

The proof is immediate as soon as we recall that the symmetry of the logarithmic moment generating function implies the symmetry $I(z) = I(-z) - z$ from (1.5) for the rate function.

3. Definition of the Norms

The norms we will use are defined via integration on the set of stable curves $\mathcal{W}^s$. Before defining the norms, we define the notion of a distance $d_{\mathcal{W}^s}(\cdot, \cdot)$ between such curves as well as a distance $d_\beta(\cdot, \cdot)$ defined among functions supported on these curves.

Due to the transversality condition on the stable cones $C^s(x)$ given by **(H1)**, each stable curve W can be viewed as the graph of a function $\varphi_W(r)$ of the arc length parameter r. For each $W \in \mathcal{W}^s$, let I_W denote the interval on which φ_W is defined and set $G_W(r) = (r, \varphi_W(r))$ to be its graph so that $W = \{G_W(r) : r \in I_W\}$. We let m_W denote the unnormalized arclength measure on W, defined using the Euclidean metric.

Let $W_1, W_2 \in \mathcal{W}^s$ and identify them with the graphs G_{W_i} of their functions φ_{W_i} defined on I_{W_i}, $i = 1, 2$. Suppose W_1, W_2 lie in the same component of M and denote by $\ell(I_{W_1} \triangle I_{W_2})$ the length of the symmetric difference between I_{W_1} and I_{W_2}. Let $\mathbb{H}_{k_i}$ be the homogeneity strip containing W_i. We define the distance[4] between W_1 and W_2 to be,

[4] Notice that $d_{\mathcal{W}^s}$ is not a metric since it does not satisfy the triangle inequality. However, we have

$$|\varphi_{W_1} - \varphi_{W_2}|_{\mathcal{C}^1(I_{W_1} \cap I_{W_2})} \le |\varphi_{W_1} - \varphi_{W_3}|_{\mathcal{C}^1(I_{W_1} \cap I_{W_3})} + |\varphi_{W_3} - \varphi_{W_2}|_{\mathcal{C}^1(I_{W_3} \cap I_{W_2})} + K\ell(I_{W_1} \cap I_{W_2} \setminus I_{W_3}),$$

$$d_{\mathcal{W}^s}(W_1, W_2) = \eta(k_1, k_2) + \ell(I_{W_1} \triangle I_{W_2}) + |\varphi_{W_1} - \varphi_{W_2}|_{\mathcal{C}^1(I_{W_1} \cap I_{W_2})}$$

where $\eta(k_1, k_2) = 0$ if $k_1 = k_2$ and $\eta(k_1, k_2) = \infty$ otherwise, i.e., we only compare curves which lie in the same homogeneity strip.

For $0 \le \alpha \le 1$, denote by $\tilde{\mathcal{C}}^\alpha(W)$ the set of continuous complex-valued functions on W with Hölder exponent α, measured in the Euclidean metric, which we denote by $d_W(\cdot, \cdot)$. We then denote by $\mathcal{C}^\alpha(W)$ the closure of $\mathcal{C}^\infty(W)$ in the $\tilde{\mathcal{C}}^\alpha$-norm[5]: $|\psi|_{\mathcal{C}^\alpha(W)} = |\psi|_{\mathcal{C}^0(W)} + H^\alpha_W(\psi)$, where $H^\alpha_W(\psi)$ is the Hölder constant of ψ along W. Notice that with this definition, $|\psi_1\psi_2|_{\mathcal{C}^\alpha(W)} \le |\psi_1|_{\mathcal{C}^\alpha(W)}|\psi_2|_{\mathcal{C}^\alpha(W)}$. We define $\tilde{\mathcal{C}}^\alpha(M)$ and $\mathcal{C}^\alpha(M)$ similarly.

Given two functions $\psi_i \in \mathcal{C}^\beta(W_i, \mathbb{C})$, $\beta > 0$, we define the distance between ψ_1, ψ_2 as

$$d_\beta(\psi_1, \psi_2) = |\psi_1 \circ G_{W_1} - \psi_2 \circ G_{W_2}|_{\mathcal{C}^\beta(I_{W_1} \cap I_{W_2})}.$$

We will define the required Banach spaces by closing $\mathcal{C}^1(M)$ with respect to the following set of norms.

Fix $0 < \alpha \le \min\{\frac{1}{3}, \frac{\alpha_1}{2}\}$, where α_1 is from Lemma 4.2. Given a function $h \in \mathcal{C}^1(M)$, define the *weak norm* of h by

$$|h|_w := \sup_{W \in \mathcal{W}^s} \sup_{\substack{\psi \in \mathcal{C}^\alpha(W) \\ |\psi|_{\mathcal{C}^\alpha(W)} \le 1}} \int_W h\psi \, dm_W. \tag{3.1}$$

Choose $\beta, \gamma, p > 0$ such that $\beta < \alpha$, $p \le 1/3$ and $\gamma < \min\{p, \alpha - \beta, 1/7\}$. We define the *strong stable norm* of h as

$$\|h\|_s := \sup_{W \in \mathcal{W}^s} \sup_{\substack{\psi \in \mathcal{C}^\beta(W) \\ |W|^p |\psi|_{\mathcal{C}^\beta(W)} \le 1}} \int_W h\psi \, dm_W \tag{3.2}$$

and the *strong unstable norm* as

$$\|h\|_u := \sup_{\varepsilon \le \varepsilon_0} \sup_{\substack{W_1, W_2 \in \mathcal{W}^s \\ d_{\mathcal{W}^s}(W_1, W_2) \le \varepsilon}} \sup_{\substack{\psi_i \in \mathcal{C}^\alpha(W_i) \\ |\psi_i|_{\mathcal{C}^\alpha(W_i)} \le 1 \\ d_\beta(\psi_1, \psi_2) \le \varepsilon}} \frac{1}{\varepsilon^\gamma} \left| \int_{W_1} h\psi_1 \, dm_W - \int_{W_2} h\psi_2 \, dm_W \right| \tag{3.3}$$

where $\varepsilon_0 > 0$ is chosen less than δ_0, the maximum length of $W \in \mathcal{W}^s$ which is determined by (2.15). We then define the *strong norm* of h by

$$\|h\|_{\mathcal{B}} = \|h\|_s + b\|h\|_u$$

where b is a small constant chosen in (4.4).

Footnote 4 continued
where $K = 2\sup_{W \in \mathcal{W}^s} |\varphi_W|_{\mathcal{C}^1}$. If we define $\tilde{d}_{\mathcal{W}^s}(W_1, W_2) = \eta(k_1, k_2) + \ell(I_{W_1} \triangle I_{W_2}) + \frac{1}{K}|\varphi_{W_1} - \varphi_{W_2}|_{\mathcal{C}^1(I_{W_1} \cap I_{W_2})}$, then $\tilde{d}_{\mathcal{W}^s}$ does satisfy the triangle inequality. We do not introduce such a modification since we do not need this property: the unstable norm defined in (3.3) satisfies the triangle inequality with the current definition.

[5] While $\mathcal{C}^\alpha(W)$ may not contain all of $\tilde{\mathcal{C}}^\alpha(W)$, it does contain $\mathcal{C}^{\alpha'}(W)$ for all $\alpha' > \alpha$. Defining $\mathcal{C}^\alpha(W)$ in this manner ensures the injectivity of the inclusion $\mathcal{B} \hookrightarrow \mathcal{B}_w$.

We define $\mathcal{B}$ to be the completion of $\mathcal{C}^1(M)$ in the strong norm and $\mathcal{B}_w$ to be the completion of $\mathcal{C}^1(M)$ in the weak norm. We remark that as a measure, $h \in \mathcal{C}^1(M)$ is identified with $hd\mu_0$ according to our earlier convention. As a consequence, Lebesgue measure $dm = (\cos\varphi)^{-1}d\mu_0$ is not automatically included in $\mathcal{B}$ since $(\cos\varphi)^{-1} \notin \mathcal{C}^1(M)$. It follows from [DZ2, Lemma 5.5] that in fact, $m \in \mathcal{B}$ (and $\mathcal{B}_w$).

3.1. Properties of the Banach spaces. We recall some properties of our Banach spaces which demonstrate that although they are spaces of distributions defined as closures of $\mathcal{C}^1$ functions in the stated norms, they enjoy some natural relations with more familiar spaces of functions and distributions. Recall $H_n^\alpha(\psi) := \sup_{W \in T^{-n}\mathcal{W}^s} H_W^\alpha(\psi)$ from Sect. 2.3.

Lemma 3.1. *The following properties hold.*

(i) ([DZ2, Lemma 5.4]) *There exists $C > 0$ such that for any $h \in \mathcal{B}_w$, $T \in \mathcal{F}$, $n \geq 0$ and $\psi \in \mathcal{C}^\alpha(T^{-n}\mathcal{W}^s)$,*

$$|h(\psi)| \leq C|h|_w(|\psi|_\infty + H_n^\alpha(\psi)).$$

(ii) ([DZ3, Lemma 2.1]) *There is a sequence of continuous inclusions $\mathcal{C}^q(M) \hookrightarrow \mathcal{B} \hookrightarrow \mathcal{B}_w \hookrightarrow (\mathcal{C}^\alpha(M))'$, for all $q > \gamma/(1-\gamma)$. The inclusions are injective, except possibly the last.*[6]

(iii) ([DZ1, Lemma 3.10]) *The unit ball of $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ is compactly embedded in $(\mathcal{B}_w, |\cdot|_w)$.*

We shall need the following result, which is [DZ3, Lemma 3.5]. Let $N_\varepsilon(\cdot)$ denote the ε-neighborhood of a set in M.

Lemma 3.2. *Let $\mathcal{P}$ be a (mod 0) countable partition of M into open, simply connected sets such that:*
(1) There are constants $K, C_1 > 0$ such that for each $P \in \mathcal{P}$ and $W \in \mathcal{W}^s$, $P \cap W$ consists of at most K connected components and for any $\varepsilon > 0$, $m_W(N_\varepsilon(\partial P) \cap W) \leq C_1\varepsilon$; (2) Each homogeneity strip $\mathbb{H}_k$ intersects at most finitely many $P \in \mathcal{P}$.

Let $q > \gamma/(1-\gamma)$. Suppose f is a function on M such that $\sup_{P\in\mathcal{P}} |f|_{C^q(P)} < \infty$ and let $h \in \mathcal{B}$. Then $hf \in \mathcal{B}$ and

$$\|hf\|_{\mathcal{B}} \leq C\|h\|_{\mathcal{B}} \sup_{P\in\mathcal{P}} |f|_{C^q(P)}$$

for some uniform constant C.

We call a potential g *admissible* for a map $T \in \mathcal{F}$ if g is at least 1/3 Hölder continuous[7] on connected components of $M \backslash \mathcal{S}_1^T$: $\sup_{P\in\mathcal{P}_1} |g|_{\mathcal{C}^{1/3}(P)} < \infty$, where $\mathcal{P}_1$ is the partition of M into connected components of $M\backslash\mathcal{S}_1^T$.

Our final lemma of this section shows that $\mathcal{L}_{T,g}$ is well-defined as an operator from $\mathcal{B}$ to $\mathcal{B}$. Its proof is similar to [DZ1, Lemma 2.1], generalized to include potentials.

Lemma 3.3. *If g is an admissible potential for T, then $\mathcal{L}_{T,g}$ is well-defined as a continuous linear operator on both $\mathcal{B}$ and $\mathcal{B}_w$.*

---

[6] This last inclusion can be made injective by introducting a weight p' in the weak norm similar to the role of p in the strong stable norm, and requiring that $p' > \alpha$. This is carried out in [DZ3, Lemma 3.8].

[7] One can decrease the Hölder exponent 1/3 by placing another restriction on α and γ in the definition of the norms.

Proof. Let $h \in \mathcal{C}^1(M)$. The Lasota–Yorke inequalities of Proposition 4.1 show that $\mathcal{L}_{T,g}h$ has finite norm in both $\mathcal{B}$ and $\mathcal{B}_w$. In order to show that $\mathcal{L}_{T,g}h$ belongs to $\mathcal{B}$, we must approximate $\mathcal{L}_{T,g}h$ by $\mathcal{C}^1$ functions in the norm $\|\cdot\|_{\mathcal{B}}$. Note that $\mathcal{L}_{T,g}h$ has a countable number of smooth discontinuity curves given by $\mathcal{S}_{-1}^{\mathbb{H}}$ (we include the images of boundaries of the homogeneity strips). These curves define a countable partition $\mathcal{P}$ of M into open simply connected sets, and each $\mathbb{H}_k$ can intersect countably many $P \in \mathcal{P}$. In addition, the $\mathcal{C}^1$ norm of $\mathcal{L}_{T,g}h$ blows up near the curves $T\mathcal{S}_0$.

For $j \geq k_0$ let P^j denote an element of $\mathcal{P}$ such that $T^{-1}P^j \subseteq \mathbb{H}_j$. Again, the labeling is not unique, but for each j, the number of elements in $\mathcal{P}$ which are assigned the label j is finite (even in the infinite horizon case). Let $P^J = \cup_{j>J} P^j$. We claim that $\|\mathcal{L}_{T,g}h|_{P^J}\|_{\mathcal{B}}$ is arbitrarily small for J sufficiently large. On the finite set of P^j with $j \leq J$, the $\mathcal{C}^1$ norm of $\mathcal{L}_{T,g}h$ is finite and the modified partition $\mathcal{P}^* = \{P^j\}_{j \leq J} \cup \{P^J\}$ satisfies the requirements of Lemma 3.2. So we may approximate $\mathcal{L}_{T,g}h$ using Lemma 3.2 on $M \setminus P^J$ and approximate $\mathcal{L}_{T,g}h$ by 0 on P^J. Thus the lemma follows once we establish our claim.

Indeed, the claim is trivial using the estimates contained in Appendix A. For example, we must estimate $\|(\mathcal{L}_{T,g}h)|_{P^J}\|_s = \|1_{P^J}\mathcal{L}_{T,g}h\|_s$. Taking $W \in \mathcal{W}^s$ and $\psi \in \mathcal{C}^\beta(W)$ with $|W|^p|\psi|_{\mathcal{C}^\beta(W)} \leq 1$, we write

$$\int_W 1_{P^J}\mathcal{L}_{T,g}h\,\psi\,dm_W = \int_{T^{-1}(W\cap P^J)} h(J_{\mu_0}T)^{-1}e^{S_n g} J_{T^{-1}W}T\;\psi\circ T\,dm_W,$$

and the homogeneous stable components of $T^{-1}(W \cap P^J)$ correspond precisely to the tail of the series considered in (A.2) and following and so can be made arbitrarily small by choosing J large (notice that we do not need contraction here so that we may use the simpler estimate similar to Sect. A.2 applied to the strong stable norm rather than the estimate of Sect. A.3.)

Similarly, in estimating $\|\mathcal{L}_{T,a}h\|_u$, one can see that the contribution from P^J corresponds to the tail of the series from the estimates of Sect. A.4, and so this too can be made arbitrarily small by choosing J large. □

4. Proof of Theorem 2.3

The proof of Theorem 2.3 relies on the following more general proposition. Recall that an admissible potential g for $T \in \mathcal{F}$ is one that satisfies $|g|_{\mathcal{C}^{1/3}(\mathcal{P}_1)} := \sup_{P\in\mathcal{P}_1}|g|_{\mathcal{C}^{1/3}(P)} < \infty$, where $\mathcal{P}_1$ is the partition of M into connected components of $M\setminus\mathcal{S}_1^T$. For an admissible potential g, define $C_g := 1 + C_e|g|_{\mathcal{C}^\alpha(\mathcal{P}_1)}\sum_{i=0}^\infty \Lambda^{-i\beta}$.

Proposition 4.1. *There exists $C > 0$, depending only on* **(H1)–(H5)**, *such that for any $T \in \mathcal{F}$, admissible potential g, $h \in \mathcal{B}$ and $n \geq 0$,*

$$|\mathcal{L}_{T,g}^n h|_w \leq CC_g|(J_{\mu_0}T^n)^{-1}e^{S_n g}|_\infty |h|_w, \tag{4.1}$$

$$\|\mathcal{L}_{T,g}^n h\|_s \leq CC_g|(J_{\mu_0}T^n)^{-1}e^{S_n g}|_\infty \left((\theta_*^{(1-p)n} + \Lambda^{-\beta n})\|h\|_s + C\delta_0^{-p}|h|_w\right), \tag{4.2}$$

$$\|\mathcal{L}_{T,g}^n h\|_u \leq CC_g|(J_{\mu_0}T^n)^{-1}e^{S_n g}|_\infty \left(\Lambda^{-\gamma n}\|h\|_u + CC_3^n\|h\|_s\right), \tag{4.3}$$

where C_3 is from Lemma A.1(d).

The proof of this proposition is fairly technical, but has a lot of similarity with the corresponding inequalities proved in [DZ1] and [DZ2] in the case $g = 0$. We put the proof in Appendix A for completeness and to draw out the explicit dependence on the added potential.

Choose $\max\{\theta_*^{1-p}, \Lambda^{-\beta}, \Lambda^{-\gamma}\} < \sigma < 1$. Then there exists $N \geq 1$ such that

$$\begin{aligned}
\|\mathcal{L}_{T,g}^N h\|_{\mathcal{B}} &= \|\mathcal{L}_{T,g}^N h\|_s + b\|\mathcal{L}_{T,g}^N h\|_u \\
&\leq CC_g|(J_{\mu_0}T^N)^{-1}e^{S_N g}|_\infty \left(\frac{\sigma^N}{2}\|h\|_s + C\delta_0^{-p}|h|_w + b\sigma^N\|h\|_u + bCC_3^N\|h\|_s\right) \\
&\leq CC_g|(J_{\mu_0}T^N)^{-1}e^{S_N g}|_\infty \left(\sigma^N\|h\|_{\mathcal{B}} + C_{\delta_0}|h|_w\right), \qquad (4.4)
\end{aligned}$$

providing b is chosen sufficiently small so that $bCC_3^N \leq \sigma^N/2$. This is the standard Lasota–Yorke inequality for $\mathcal{L}_{T,g}$ for a general potential g. In order to specialize to the case $g = a\log J_{\mu_0}T$, we recall the following lemma about the form of the Jacobian $J_{\mu_0}T$ derived in [CZZ].

In what follows, we define $G = (G^1, G^2)$ to be the map on M induced by the twist $\mathbf{G}$, where G^1 and G^2 are C^2 functions on M. Due to (2.2) and **(A4)**, if we denote $(r_1, \varphi_1) = T_{\mathbf{F},\mathbf{0}}(r, \varphi)$, then

$$T_{\mathbf{F},\mathbf{G}}(r, \varphi) = (r_1, \varphi_1) + G(r_1, \varphi_1) = (I + G)(T_{\mathbf{F},\mathbf{0}}(r, \varphi)). \qquad (4.5)$$

Lemma 4.2 ([CZZ, Lemmas 3.2 and 4.2]). *Fix ε, τ_* and C_0 and consider $T_{\mathbf{F},\mathbf{G}} \in \mathcal{F}(\varepsilon, \tau_*, C_0)$.*

First, assume there is no twist force $\mathbf{G} = 0$ and denote $T_{\mathbf{F},\mathbf{0}} = T_{\mathbf{F}}$. Then the Jacobian of $T_{\mathbf{F}}$ with respect to μ_0 is given by

$$J_{\mu_0}T_{\mathbf{F}} = \exp\left(\int_0^{\tau_{\mathbf{F}}(\mathbf{x})} p\,\tfrac{\partial\kappa}{\partial\theta}\,dt\right), \qquad (4.6)$$

where $\tau_{\mathbf{F}}$ is the free path for the system $T_{\mathbf{F}}$ and κ is from (2.6).

*Next, assume $\mathbf{G} \neq 0$. Then by Assumption (**A4**) the Jacobian of $T_{\mathbf{E}} = T_{\mathbf{F},\mathbf{G}}$ satisfies,*

$$J_{\mu_0}T_{\mathbf{E}} = J_{\mu_0}(I + G)(T_{\mathbf{F}})J_{\mu_0}T_{\mathbf{F}}. \qquad (4.7)$$

Moreover, we may write,

$$J_{\mu_0}T_{\mathbf{E}} = 1 + \varepsilon H, \qquad \textit{where} \quad H = \tfrac{1}{\varepsilon}\left(J_{\mu_0}T_{\mathbf{E}} - 1\right), \qquad (4.8)$$

$|H|_\infty \leq C_H$ for some $C_H > 0$ independent of ε and H is C^{α_1} for some[8] *$\alpha_1 \in (1/3, 1/2]$ on each component of $\mathcal{S}_1^{T_{\mathbf{E}}}$.*

Proof. The representation of $J_{\mu_0}T_{\mathbf{F}}$ given by (4.6) is proved in [CZZ, Lemma 3.2] under precisely the same assumptions as here. Since the coordinates and assumptions used in [CZZ, Lemma 4.2] differ slightly from ours, we proceed to verify the case $\mathbf{G} \neq \mathbf{0}$ directly.

[8] The restriction on α_1 comes from the fact that H is at least C^{α_0} for some $\alpha_0 > 1/3$ by Assumption **(A3)**, but in general not smoother than $\tau_{\mathbf{F}}$, which is only $1/2$-Hölder continuous.

Due to **(A4)**, Eq. (4.7) follows immediately from the chain rule and (4.5) at any point $\mathbf{x} = (r, \varphi) \notin \mathcal{S}_1^{T_\mathbf{F}}$ (see also [DZ2, Sect. 7.2]). Then using the fact that $d\mu_0 = c \cos\varphi dm$, we have

$$J_{\mu_0}(I + G)(r_1, \varphi_1) = \frac{\cos(\varphi_1 + G^2(r_1, \varphi_1))}{\cos\varphi_1} \det(I + DG)(r_1, \varphi_1). \tag{4.9}$$

By **(A4)**, $G^2(r_1, \pm\frac{\pi}{2}) = 0$, so using **(A3)**

$$|\cos(\varphi_1 + G^2(r_1, \varphi_1)) - \cos\varphi_1| \le |G^2(r_1, \varphi_1)| \le \varepsilon|\varphi_1 - \tfrac{\pi}{2}|,$$

where without loss of generality we consider φ_1 near $\pi/2$, rather than $-\pi/2$. Thus,

$$\frac{\cos(\varphi_1 + G^2(r_1, \varphi_1))}{\cos\varphi_1} = 1 \pm \varepsilon \frac{|\varphi_1 - \pi/2|}{\cos\varphi_1},$$

and the last fraction is bounded by $\pi/2$ for $\varphi_1 \in [0, \pi/2]$. This, together with **(A3)** and the fact that $\det(I + DG) = 1 + \text{Trace}(DG) + \det(DG)$, yields (4.8) with H bounded by a uniform constant C_H independent of ε. □

4.1. A Spectral Gap for $\mathcal{L}_{T,a}$. Now we fix $a_0 > 0$ and the interval $[-a_0, 1 + a_0]$ as in the statement of Theorem 2.3. Choose $\varepsilon_0 > 0$ so small that for all $a \in [-a_0, 1 + a_0]$,

$$\frac{(1 - \text{sign}(a - 1)C_H\varepsilon_0)^{a-1}}{(1 + \text{sign}(a - 1)C_H\varepsilon_0)^{a-1}} > \sigma, \tag{4.10}$$

where σ is from (4.4).

The next lemma establishes the quasi-compactness of $\mathcal{L}_{T,a}$.

Lemma 4.3. *Let $a \in [-a_0, 1 + a_0]$ and ε_0 be as chosen in* (4.10). *Then for all $T \in \mathcal{F}(\varepsilon_0, \tau_*, C_0)$, $\mathcal{L}_{T,a}$ is quasi-compact as an operator on $\mathcal{B}$.*

Proof. When $g = a \log J_{\mu_0}T$, we have $(J_{\mu_0}T^N)^{-1}e^{S_N g} = (J_{\mu_0}T^N)^{a-1}$ and so (4.4) together with Lemma 4.2 yield the required inequalities (2.17) for Theorem 2.3. Due to the compactness of the unit ball of $\mathcal{B}$ in $\mathcal{B}_w$ [DZ1, Lemma 3.10], this implies the essential spectral radius of $\mathcal{L}_{T,a}$, $\rho_{\text{ess}}(\mathcal{L}_{T,a})$ is at most $\sigma(1 + \text{sign}(a - 1)C_H\varepsilon_0)^{a-1}$. To prove that $\mathcal{L}_{T,a}$ is quasi-compact, it remains to show that the spectral radius of $\mathcal{L}_{T,a}$, $\rho(\mathcal{L}_{T,a})$, is strictly larger than $\rho_{\text{ess}}(\mathcal{L}_{T,a})$.

To obtain a lower bound on $\rho(\mathcal{L}_{T,a})$, note that

$$\rho(\mathcal{L}_{T,a}) = \lim_{n\to\infty} \|\mathcal{L}_{T,a}^n\|_{\mathcal{B}}^{1/n} \ge \lim_{n\to\infty} \|\mathcal{L}_{T,a}^n 1\|_s^{1/n}.$$

Then we have

$$\begin{aligned}\|\mathcal{L}_{T,a}^n 1\|_s &= \sup_{W\in\mathcal{W}^s} \sup_{\substack{\psi\in\mathcal{C}^\beta(W)\\ |W|^p|\psi|_{\mathcal{C}^\beta(W)}\le 1}} \int_W \mathcal{L}_{T,a}^n 1 \cdot \psi \, dm_W \\ &\ge \sup_{W\in\mathcal{W}^s} \sup_{\substack{\psi\in\mathcal{C}^\beta(W)\\ |W|^p|\psi|_{\mathcal{C}^\beta(W)}\le 1}} \inf \mathcal{L}_{T,a}^n 1 \int_W \psi \, dm_W \;\ge\; \inf(1 + \varepsilon_0 H)^{(a-1)n} \,\|1\|_s,\end{aligned}$$

using Lemma 4.2 and the identity $\mathcal{L}_{T,a}^n 1 = (J_{\mu_0} T^n)^{a-1} \circ T^{-n}$. This implies that

$$\rho(\mathcal{L}_{T,a}) = \lim_{n\to\infty} \|\mathcal{L}_{T,a}^n\|^{\frac{1}{n}} \geq (1 - \text{sign}(a-1) C_H \varepsilon_0)^{a-1}.$$

Combining this with the upper bound on the essential spectrum of $\mathcal{L}_{T,a}$ and the choice of ε_0 from (4.10), we conclude

$$\rho_{\text{ess}}(\mathcal{L}_{T,a}) \leq \sigma(1 + \text{sign}(a-1) C_H \varepsilon_0)^{a-1} < (1 - \text{sign}(a-1) C_H \varepsilon_0)^{a-1} \leq \rho(\mathcal{L}_{T,a}).$$

□

Recall from Sect. 3.1 that a function $g : M \to \mathbb{R}$ is an admissible potential for $T \in \mathcal{F}$ if $|g|_{\mathcal{C}^{1/3}(\mathcal{P}_1)} := \sup_{P \in \mathcal{P}_1} |g|_{\mathcal{C}^{1/3}(P)} < \infty$, where $\mathcal{P}_1$ is the partition of M into connected components of $M \backslash \mathcal{S}_1^T$.

Lemma 4.4. *Suppose g is an admissible potential for $T \in \mathcal{F}(\varepsilon_0, \tau_*, C_0)$. Then the map $z \mapsto \mathcal{L}_{T,zg}$ is analytic for all $z \in \mathbb{C}$.*

Proof. Define the operator $\mathcal{A}_n h = \mathcal{L}_T(g^n h) = g^n \circ T^{-1} \mathcal{L}_T h$, for $h \in \mathcal{B}$. Notice that since g is Hölder continuous on elements of $\mathcal{P}_1$, it follows that $g \circ T^{-1}$ is Hölder continuous on elements of $\mathcal{P}_{-1}$, the partition of M into connected components of $M \backslash \mathcal{S}_{-1}^T$. Since $\mathcal{S}_{-1}^T$ consists of finitely many curves that are uniformly transverse to the stable cone, we claim that $g \circ T^{-1}$ satisfies the assumptions of Lemma 3.2. Indeed, we have the following estimate for the Hölder regularity of $g \circ T^{-1}$. For any x, y in the same component of $\mathcal{S}_{-1}^T$,

$$\frac{|g \circ T^{-1}(x) - g \circ T^{-1}(y)|}{d(x,y)^{1/6}} = \frac{|g \circ T^{-1}(x) - g \circ T^{-1}(y)|}{d(T^{-1}(x), T^{-1}(y))^{1/3}} \frac{d(T^{-1}(x), T^{-1}(y))^{1/3}}{d(x,y)^{1/6}}. \tag{4.11}$$

The first factor is bounded by $|g|_{\mathcal{C}^{1/3}(\mathcal{P}_1)}$, while the second factor is uniformly bounded due to the fact that $|T^{-1}W| \leq C|W|^{1/2}$ for any $W \in \mathcal{W}^s$ by **(H3)** (see, for example, [CM, Exercise 4.50]). Thus $g \circ T^{-1}$ is 1/6-Hölder continuous on $\mathcal{P}_1$ and $1/6 \geq \gamma/(1-\gamma)$ since $\gamma < 1/7$ so that $g \circ T^{-1}$ satisfies the conditions of Lemma 3.2.

Now Lemma 3.2 implies that $g^n \circ T^{-1} \mathcal{L}_T h \in \mathcal{B}$ and moreover,

$$\|\mathcal{A}_n h\|_{\mathcal{B}} = \|g^n \circ T^{-1} \mathcal{L}_T h\|_{\mathcal{B}} \leq C \|\mathcal{L}_T h\|_{\mathcal{B}} |g^n \circ T^{-1}|_{\mathcal{C}^{1/6}(\mathcal{P}_{-1})} \leq C \|h\|_{\mathcal{B}} |g|^n_{\mathcal{C}^{1/3}(\mathcal{P}_1)},$$

where we used (4.11) along with the simple fact that $|fg|_{\mathcal{C}^q} \leq |f|_{\mathcal{C}^q} |g|_{\mathcal{C}^q}$ to estimate $|g^n|_{\mathcal{C}^q} \leq |g|^n_{\mathcal{C}^q}$.

Therefore, the operator $\sum_{n=0}^\infty \frac{z^n}{n!} \mathcal{A}_n$ is well defined on $\mathcal{B}$ and equals $\mathcal{L}_{T,zg}$ since once we know the sum converges,

$$\sum_{n=0}^\infty \frac{z^n}{n!} \mathcal{A}_n h(\psi) = h\left(\sum_{n=0}^\infty \frac{z^n}{n!} g^n \cdot \psi \circ T\right) = h(e^{zg} \psi \circ T) = \mathcal{L}_{T,zg} h(\psi),$$

for $\psi \in \mathcal{C}^\alpha(\mathcal{W}^s)$.

□

With the analyticity of $z \mapsto \mathcal{L}_{zg}$ established, it follows from analytic perturbation theory [Ka] that both the discrete spectrum and the corresponding spectral projectors of $\mathcal{L}_{T,zg}$ vary smoothly with z. We will use the smooth dependence of the spectrum on z to prove that $\mathcal{L}_{T,a}$ has a spectral gap.

Lemma 4.5. *Fix $a_0, \tau_*, C_0 > 0$ and let ε_0 be as in (4.10). Then there exists $0 < \varepsilon_1 \le \varepsilon_0$ such that for all $T = T_{\mathbf{E}} \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$, $\mathcal{L}_{T,a}$ has a simple eigenvalue $\lambda = \rho(\mathcal{L}_{T,a})$ and all other eigenvalues have modulus strictly smaller than λ, i.e. $\mathcal{L}_{T,a}$ has a spectral gap as an operator on $\mathcal{B}$.*

Proof. Fix a uniform family $\mathcal{F}(\varepsilon_0, \tau_*, C_0)$ satisfying **(H1)–(H5)** and (4.10) such that $\mathcal{L}_T$ has a spectral gap for all $T \in \mathcal{F}(\varepsilon_0, \tau_*, C_0)$ by [DZ2].

Fixing $T = T_{\mathbf{E}} \in \mathcal{F}(\varepsilon_0, \tau_*, C_0)$ and using Lemma 4.2, we know $-s$ is an admissible potential for T. According to Lemma 4.4, the derivative, $\frac{d}{dz}\mathcal{L}_{T,-zs} = \sum_{n\ge 1} \frac{z^{n-1}}{(n-1)!}\mathcal{A}_n$ is well-defined as a bounded linear operator on $\mathcal{B}$, and

$$\|\frac{d}{dz}\mathcal{L}_{T,-zs}\|_{\mathcal{B}} \le C\|\mathcal{L}_T\|_{\mathcal{B}}|s|_{\mathcal{C}^{1/3}(\mathcal{P}_1)}e^{|z||s|_{\mathcal{C}^{1/3}(\mathcal{P}_1)}},$$

for a uniform constant C (depending only on $\mathcal{F}$). Thus for any $a \in [-a_0, 1 + a_0]$,

$$\begin{aligned}\|\mathcal{L}_T - \mathcal{L}_{T,a}\|_{\mathcal{B}} &\le C\|\mathcal{L}_T\|_{\mathcal{B}}|s|_{\mathcal{C}^{1/3}(\mathcal{P}_1)}e^{|a||s|_{\mathcal{C}^{1/3}(\mathcal{P}_1)}}|a| \\ &\le C|a|\|\mathcal{L}_T\|_{\mathcal{B}}|\log(1+\varepsilon H)|_{\mathcal{C}^{1/3}(\mathcal{P}_1)}e^{|\log(1+\varepsilon H)|^{|a|}_{\mathcal{C}^{1/3}(\mathcal{P}_1)}}, \quad (4.12)\end{aligned}$$

where we have used Lemma 4.2 and ε is the optimal ε for $\mathbf{E}$.

It follows from [DZ2] that the spectrum of $\mathcal{L}_{T_{\mathbf{E}}}$ varies continuously in $\mathbf{E}$ and converges to the spectrum $\mathcal{L}_{T_0}$ as $\mathbf{E}$ shrinks to 0 (in $\mathcal{C}^1$ norm). Thus there exists $0 < \varepsilon_2$, $\varepsilon_2 \le \varepsilon_0$ such that all $T_{\mathbf{E}} \in \mathcal{F}(\varepsilon_2, \tau_*, C_0)$ enjoy a uniform spectral gap, i.e., the distance between 1 and the second largest eigenvalue of $\mathcal{L}_{T_{\mathbf{E}}}$ is bounded below by a uniform constant; call this constant $\delta > 0$. Then by (4.12), there exists $\varepsilon_1 > 0$ such that $\mathcal{L}_{T,a}$ has a spectral gap for any $T_{\mathbf{E}} \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$ and all $a \in [-a_0, 1 + a_0]$. □

We will also find it convenient to have the following continuity in ε.

Lemma 4.6. *Fix $a \in [-a_0, 1 + a_0]$ and $T_0 = T_{\mathbf{0},\mathbf{0}} \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$. There exists $C > 0$ such that for all $\varepsilon \le \varepsilon_1$ and all $T_{\mathbf{E}} \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$ with $d_{\mathcal{F}}(T_0, T_{\mathbf{E}}) \le \varepsilon$, we have*

$$\sup\{|\mathcal{L}_{T_{\mathbf{E}},a}h - \mathcal{L}_{T_0,a}h|_w : \|h\|_{\mathcal{B}} \le 1\} \le C\varepsilon^{\gamma/2}.$$

This implies in particular that the leading eigenvalue and associated spectral projectors of $\mathcal{L}_{T_{\mathbf{E}},a}$ vary continuously with $\mathbf{E}$ in the ε_1 neighborhood of T_0.

Proof. The proof is essentially the same as the proof of [DZ1, Theorem 2.3], except with the added potential $(J_{\mu_0}T)^{a-1}$. We just sketch the proof here, noting the necessary additions.

Fixing T_0 and $T_{\mathbf{E}}$ as in the statement of the lemma, we choose $h \in \mathcal{C}^1(M)$ with $\|h\|_{\mathcal{B}} \le 1$ and $W \in \mathcal{W}^s$. Let $\psi \in \mathcal{C}^\alpha(W)$ satisfy $|\psi|_{\mathcal{C}^\alpha(W)} \le 1$. For the weak norm of the difference, we must estimate

$$\begin{aligned}\int_W (\mathcal{L}_{T_0,a}h - \mathcal{L}_{T_{\mathbf{E}},a}h)\psi\, dm_W = &\int_{T_0^{-1}W} hJ_{T_0^{-1}W}T\,\psi\circ T_0 \\ &- \int_{T_{\mathbf{E}}^{-1}W} h(J_{\mu_0}T_{\mathbf{E}})^{a-1}J_{T_{\mathbf{E}}^{-1}W}T\,\psi\circ T_{\mathbf{E}},\end{aligned}$$

where we have used the fact that $J_{\mu_0}T_0 = 1$. The required estimate is similar to the estimate for the strong unstable norm in contained in Sect. A.4, except that we have one stable curve iterated under two different maps instead of two close stable curves iterated under the same map. However, the decomposition is the same: we subdivide $T_0^{-1}W$ and $T_{\mathbf{E}}^{-1}W$ into matched and unmatched pieces. The matched pieces can be connected by a transverse foliation of unstable curves, while the unmatched pieces are short. The estimates proceed precisely as in [DZ2, Section 5], with [DZ2, Lemma 5.1] providing the bounds on all the relevant quantities. The only additional piece in the present estimate is the presence of the potential $(J_{\mu_0}T_{\mathbf{E}})^{a-1}$.

For the sum over unmatched unmatched pieces, it is bounded by $C\varepsilon^{\gamma/2}|(J_{\mu_0}T_{\mathbf{E}})^{a-1}|_{\mathcal{C}^\beta(\mathcal{P}_1)}\|h\|_s$ using the strong stable norm precisely as in (A.13), with $n=1$, since each unmatched piece has length at most $\varepsilon^{1/2}$.

Suppose U_1 and U_2 are two matched pieces of $T_0^{-1}W$ and $T_{\mathbf{E}}^{-1}W$, respectively. By construction, they are defined over a common r-interval I, i.e. they can be written as graphs of functions

$$U_j = G_{U_j}(I) = \{(r, \varphi_{U_j}(r)) : r \in I\}$$

and $d_{\mathcal{W}^s}(U_1, U_2) \le C\varepsilon^{1/2}$ ([DZ2, Lemma 5.1(a)]). The estimate over matched pieces proceeds precisely as in (A.14) and the only difference in test functions unaccounted for in (A.19) is $|(J_{\mu_0}T_{\mathbf{E}})^{a-1}-1|_{\mathcal{C}^\beta(U_2)}$. We will show that

$$|(J_{\mu_0}T_{\mathbf{E}})^{a-1}-1|_{\mathcal{C}^\beta(U_2)} \le C\varepsilon^{1-3\beta}, \tag{4.13}$$

for some uniform constant C depending on a_0. Indeed $|(J_{\mu_0}T_{\mathbf{E}})^{a-1}-1|_{\mathcal{C}^0(U_2)} \le C|a-1|\varepsilon$ follows from Lemma 4.2. For the Hölder constant, we take $x, y \in U_2$ and estimate on the one hand using **(H4)**,

$$|(J_{\mu_0}T_{\mathbf{E}})^{a-1}(x) - (J_{\mu_0}T_{\mathbf{E}})^{a-1}(y)| \le C|(J_{\mu_0}T_{\mathbf{E}})^{a-1}|_{\mathcal{C}^0(U_2)}d(x,y)^{1/3} \le C'd(x,y)^{1/3}.$$

While on the other hand,

$$|(J_{\mu_0}T_{\mathbf{E}})^{a-1}(x) - (J_{\mu_0}T_{\mathbf{E}})^{a-1}(y)| \le C''\varepsilon,$$

using Lemma 4.2 once again. So the Hölder constant is bounded by the minimum of these two expressions,

$$\min\{C'd(x,y)^{1/3-\beta}, C''\varepsilon d(x,y)^{-\beta}\}.$$

This bound can be no worse than when the two quantities are equal, i.e. $\varepsilon = (C'/C'')\,d(x,y)^{1/3}$. Thus $H_W^\beta((J_{\mu_0}T_{\mathbf{E}})^{a-1}-1) \le C\varepsilon^{1-3\beta}$. This proves (4.13).

Now gathering terms over matched and unmatched pieces as in (A.21) or [DZ2, Eq. (5.9)], we see that the least power of ε is $\varepsilon^{\gamma/2}$, from the unmatched pieces (notice that $\varepsilon^{1-3\beta} < \varepsilon^{\frac{1}{3}-\beta}$ and $\gamma \le \alpha - \beta \le \frac{1}{3} - \beta$). This completes the proof of the lemma. □

Fix $T \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$. Let $h_a \in \mathcal{B}$ be the eigenvector of $\mathcal{L}_{T,a}$ corresponding to the eigenvalue λ_a of maximum modulus, and $\nu_a \in \mathcal{B}^*$ be the corresponding eigenvector of the dual $\mathcal{L}_{T,a}^*$. That is, $\mathcal{L}_{T,a}h_a = \lambda_a h_a$, and $\mathcal{L}_a^*\nu_a = \lambda_a\nu_a$. Due to the spectral gap for $\mathcal{L}_{T,a}$, we have the following spectral decomposition,

$$\mathcal{L}_{T,a}^n h = \lambda_a^n \Pi_a h + R_a^n h, \tag{4.14}$$

where $\Pi_a R_a = R_a \Pi_a = 0$ and the spectral radius of R_a is strictly smaller than λ_a. Also, for any $h \in \mathcal{B}$, $\Pi_a h = c_a(h)h_a$, where $c_a : \mathcal{B} \to \mathbb{R}$ is a bounded linear functional.

Notice that λ_a must be real since $\mathcal{L}_{T,a}$ is a real operator and the spectral gap for $\mathcal{L}_{T,a}$ is obtained as a perturbation of $\mathcal{L}_{T,0}$, which has $\lambda_0 = 1$.

The following lemma completes the proof of Theorem 2.3.

Lemma 4.7. *Both eigenvectors h_a and ν_a are positive measures. Moreover, the pairing $\mu_a := h_a \otimes \nu_a$ defines an invariant measure for T.*

Proof. Due to (4.14), for any $\psi \in \mathcal{C}^\alpha(M)$,

$$|c_a(1)h_a(\psi)| = \lim_{n\to\infty} |\lambda_a^{-n}\mathcal{L}_{T,a}^n 1(\psi)| \le \lim_{n\to\infty} |\psi|_\infty |\lambda_a^{-n}\mathcal{L}_{T,a}^n 1(1)|$$
$$= |\psi|_\infty |c_a(1)||h_a(1)|.$$

Now $c_0(1) = 1$ and $c_a(1)$ is continuous in a by Lemma 4.4, so by (4.12), $c_a(1) > 0$ for $\varepsilon \in [0, \varepsilon_1]$. This, together with the above estimate, implies that h_a is a measure. Then it is evident that h_a is a positive measure due to the positivity of $\mathcal{L}_{T,a}$.

Similarly, one can show that ν_a is also a positive measure since

$$\lim_{n\to\infty} \lambda_a^{-n}(\mathcal{L}_{T,a}^*)^n 1(\psi) = c_a^*(1)\nu_a(\psi),$$

for some linear functional c_a^*.

By Lemma 3.2, if ψ is a piecewise Hölder continuous function on M, then $\psi h_a \in \mathcal{B}$. So we may define a measure on M via the pairing $\mu_{T,a} := h_a \otimes \nu_a$, i.e. $\mu_{T,a}(\psi) = \langle \psi h_a, \nu_a\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the pairing between $\mathcal{B}$ and its dual. Moreover, the measure $\mu_{T,a}$ is invariant under T:

$$\mu_{T,a}(\psi\circ T) = \langle \psi\circ T\cdot h_a, \lambda_a^{-1}\mathcal{L}_{T,a}^*\nu_a\rangle = \lambda_a^{-1}\langle \mathcal{L}_{T,a}(\psi\circ T\cdot h_a), \nu_a\rangle$$
$$= \lambda_a^{-1}\langle \psi\mathcal{L}_{T,a}(h_a), \nu_a\rangle = \langle \psi h_a, \nu_a\rangle = \mu_{T,a}(\psi), \quad \text{for any } \psi \in C^{1/3}(M),$$

where we have used Lemma 3.2 to conclude that $\psi\circ T\cdot h_a \in \mathcal{B}$. □

Remark 4.8. Notice that when $a = 0$, the smooth measure μ_0 is the conformal measure with respect to $\mathcal{L}_{T,0}$, i.e. $\mathcal{L}_{T,0}^*\mu_0 = \mu_0$, so that $\nu_0 = \mu_0$ and $\langle h_0, \mu_0\rangle = 1$. It then follows from Lemma 4.4 and (4.12) that that we may choose $\varepsilon_1 > 0$ sufficently small so that $\langle h_a, \mu_0\rangle > 0$ and $c_a(h_0) > 0$ for all $a \in [-a_0, 1 + a_0]$.

5. Proof of Theorem 2.4

In this section, we shall be more explicit about the dependence of the various objects on the forces $\mathbf{E} = (\mathbf{F}, \mathbf{G})$. We shall use the following notation for the map $T = T_{\mathbf{E}}$ and the potential e^{ag_0}. We have the following decomposition according to (4.14):

$$\mathcal{L}_{T_{\mathbf{E}},a} = \lambda_{\mathbf{E},a}\Pi_{\mathbf{E},a} + R_{\mathbf{E},a}.$$

Denote by $\mu_{\mathbf{E},a} = h_{\mathbf{E},a} \otimes \nu_{\mathbf{E},a}$ the $T_{\mathbf{E}}$-invariant measure constructed using the left and right eigenvectors of $\mathcal{L}_{T_{\mathbf{E}},a}$. When $a = 0$, in what follows, we will drop the subscript corresponding to a, and simply write $\mu_{\mathbf{E}} = h_{\mathbf{E}} \otimes \nu_{\mathbf{E}}$ for the SRB measure of the perturbed system $T_{\mathbf{E}} \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$. Note this notation is consistent with our use of $\mu_0 = \mu_{\mathbf{0},0}$ as both the conformal measure for $\mathcal{L}_{T,0}$ as well as the smooth invariant measure corresponding to the classical billiard map $T_{\mathbf{0},\mathbf{0}}$, with $\mathbf{F} = \mathbf{G} = \mathbf{0}$. Indeed, when $\mathbf{E} = (\mathbf{0}, \mathbf{0})$, then $h_{\mathbf{0},0} = 1$.

The moment generating function $e_{\mathbf{E}}(a)$ is defined as in (2.18),

$$e_{\mathbf{E}}(a) = \lim_{n\to\infty}\frac{1}{n}\log\mu_{\mathbf{E}}((J_{\mu_0}T_{\mathbf{E}}^n)^a).$$

Proof of Theorem 2.4. The existence and uniqueness of $\mu_{\mathbf{E}}$ for Item (1) follow from the spectral gap of $\mathcal{L}_{T_{\mathbf{E}}}$ established by Theorem 2.3.

To prove item (2), first recall that $\langle h_{\mathbf{E},a}, \mu_0\rangle > 0$ and $c_{\mathbf{E},a}(h_{\mathbf{E},0}) > 0$ by choice of ε_1 and Remark 4.8. Then

$$\begin{aligned} e_{\mathbf{E}}(a) &= \lim_{n\to\infty} \frac{1}{n} \log \mu_{\mathbf{E}}((J_{\mu_0} T^n_{\mathbf{E}})^a) = \lim_{n\to\infty} \frac{1}{n} \log\langle h_{\mathbf{E}} \cdot (J_{\mu_0} T^n_{\mathbf{E}})^a, \mu_0\rangle \\ &= \lim_{n\to\infty} \frac{1}{n} \log\langle \mathcal{L}^n_{T_{\mathbf{E}},a} h_{\mathbf{E}}, \mu_0\rangle \\ &= \lim_{n\to\infty} \frac{1}{n} \log\langle \lambda^n_{\mathbf{E},a} c_{\mathbf{E},a}(h_{\mathbf{E}}) h_{\mathbf{E},a} + R^n_{\mathbf{E},a} h_{\mathbf{E}}, \mu_0\rangle = \log \lambda_{\mathbf{E},a}. \end{aligned}$$

Thus by Lemma 4.4, since $\lambda_{\mathbf{E},a}$ is simple, $e_{\mathbf{E}}(a)$ is analytic as a function of a for $a \in [-a_0, 1 + a_0]$.

Now let $\nu \in \mathcal{B}$ be a probability measure with $c_{\mathbf{E},a}(\nu) > 0$. Then the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \nu((J_{\mu_0} T^n_{\mathbf{E}})^a)$$

exists and has the value $\log \lambda_{\mathbf{E},a}$ by precisely the same calculation as above. Thus the moment generating function can be defined using ν in place of the invariant measure $\mu_{\mathbf{E}}$. Note that since $c_{\mathbf{E},0}(\nu) = 1$ for any probability measure $\nu \in \mathcal{B}$, and due to the inequality

$$\|(\Pi_{\mathbf{E},0} - \Pi_{\mathbf{E},a})\nu\|_{\mathcal{B}} \le \|\Pi_{\mathbf{E},0} - \Pi_{\mathbf{E},a}\|_{\mathcal{B}} \|\nu\|_{\mathcal{B}},$$

Lemma 4.4 implies that if we fix a ball of radius $r > 0$ in $\mathcal{B}$, then we may choose ε_1 so that $c_{\mathbf{E},a}(\nu) > 0$ for all ν in this ball of radius r, all $a \in [-a_0, 1 + a_0]$ and all $T_{\mathbf{E}} \in \mathcal{F}(\varepsilon_1, \tau_*, C_0)$. For this range of parameters, it follows that Lebesgue measure m, the smooth measure μ_0 and the (possibly singular) SRB measure $\mu_{\mathbf{E}}$ all yield the same logarithmic moment generating function $e_{\mathbf{E}}(a)$. From this and Proposition 1.4 we conclude the symmetry $e_{\mathbf{E}}(a) = e_{\mathbf{E}}(1 - a)$ for $a \in [-a_0, 1 + a_0]$.

To prove item (3), we compute the derivatives of $e_{\mathbf{E}}(a)$ at $a = 0$, following [RY] (see also [D]). The sequence $\{\frac{1}{n} \log \mu_{\mathbf{E}}((J_{\mu_0} T^n_{\mathbf{E}})^a)\}_{n\in\mathbb{N}}$ is uniformly bounded for a in a complex neighborhood of the origin. Thus by the Vitali convergence theorem we can freely exchange derivative and limits. Thus

$$e'_{\mathbf{E}}(0) = \lim_{n\to\infty} \frac{1}{n} \mu_{\mathbf{E}}(\log J_{\mu_0} T^n_{\mathbf{E}}) = \mu_{\mathbf{E}}(\log J_{\mu_0} T_{\mathbf{E}}),$$

due to the invariance of $\mu_{\mathbf{E}}$ with respect to $T_{\mathbf{E}}$. Now using Lemma 4.2, we have $J_{\mu_0} T_{\mathbf{E}}(x) = 1 + \varepsilon H(x)$. Thus for small $|\varepsilon| < 1$,

$$e'_{\mathbf{E}}(0) = \mu_{\mathbf{E}}(\log(1 + \varepsilon H)) = \varepsilon \mu_{\mathbf{E}}(H) + \mathcal{O}(\varepsilon^2).$$

Next, using Lemma 4.6 and [KL, Corollary 1], we have

$$|\mu_{\mathbf{E}} - \mu_0|_w \le C\varepsilon^{\eta}, \tag{5.1}$$

for some $\eta > 0$. Putting these estimates together, we conclude,

$$e'_{\mathbf{E}}(0) = \varepsilon \mu_0(H) + o(\varepsilon).$$

For the second derivative, setting $s = -\log J_{\mu_0} T_{\mathbf{E}}$ as before, and $\bar{s} = s - \mu_{\mathbf{E}}(s)$, we have

$$\begin{aligned} e''_{\mathbf{E}}(0) &= \lim_{n\to\infty} \frac{1}{n}(\mu_{\mathbf{E}}((S_n s)^2) - \mu_{\mathbf{E}}(S_n s)^2) = \lim_{n\to\infty} \frac{1}{n}\mu_{\mathbf{E}}((S_n \bar{s})^2) \\ &= \mu_{\mathbf{E}}(\bar{s}^2) + 2\lim_{n\to\infty} \sum_{j=1}^{n-1} (1 - j/n)\mu_{\mathbf{E}}(\bar{s}\cdot\bar{s}\circ T_{\mathbf{E}}^j) \\ &= \mu_{\mathbf{E}}(\bar{s}^2) + 2\sum_{j=1}^{\infty} \mu_{\mathbf{E}}(\bar{s}\cdot\bar{s}\circ T_{\mathbf{E}}^j). \end{aligned}$$

The last equality follows from the exponential decay of correlations and dominated convergence.

Let $\sigma^2_{\mathbf{E}}$ denote the limit of the variance of $n^{-1/2} S_n s$ as $n \to \infty$ where $\{s \circ T^j_{\mathbf{E}}\}_{j\in\mathbb{N}}$ is distributed according to the invariant measure $\mu_{\mathbf{E}}$. (Such a $\sigma_{\mathbf{E}}$ exists and is finite whenever the auto-correlations $\mu_{\mathbf{E}}(\bar{s}\cdot\bar{s}\circ T^j_{\mathbf{E}})$ are summable.) The above Green Kubo formula then gives the diffusion coefficient:

$$e''_{\mathbf{E}}(0) = \mu_{\mathbf{E}}(\bar{s}^2) + 2\sum_{j=1}^{\infty} \mu_{\mathbf{E}}(\bar{s}\cdot\bar{s}\circ T_{\mathbf{E}}^j) = \sigma^2_{\mathbf{E}}. \tag{5.2}$$

We denote $\bar{H} = H - \mu_{\mathbf{E}}(H)$, where H is from Lemma 4.2 and $J_{\mu_0} T_{\mathbf{E}} = 1 + \varepsilon H$. Then,

$$-\bar{s} = \log J_{\mu_0} T_{\mathbf{E}} - \mu_{\mathbf{E}}(\log J_{\mu_0} T_{\mathbf{E}}) = \varepsilon(H - \mu_{\mathbf{E}}(H)) + \mathcal{O}(\varepsilon^2),$$

and also,

$$\begin{aligned} \mu_{\mathbf{E}}(\bar{s}^2) &= \mu_{\mathbf{E}}((\log(1+\varepsilon H))^2) - \mu_{\mathbf{E}}(\log(1+\varepsilon H))^2 \\ &= \varepsilon^2 \mathrm{Var}(H) + \mathcal{O}(\varepsilon^3) = \varepsilon^2 \mu_{\mathbf{E}}(\bar{H}^2) + \mathcal{O}(\varepsilon^3). \end{aligned}$$

It follows that

$$\begin{aligned} \sum_{j=1}^{\infty} \mu_{\mathbf{E}}(\bar{s}\cdot\bar{s}\circ T^j_{\mathbf{E}}) &= \sum_{j=1}^{\infty} \mu_{\mathbf{E}}(\bar{s}\cdot\bar{s}\circ T^j_{\mathbf{0}}) + \sum_{j=1}^{\infty}\Big(\mu_{\mathbf{E}}(\bar{s}\cdot\bar{s}\circ T^j_{\mathbf{E}}) - \mu_{\mathbf{E}}(\bar{s}\cdot\bar{s}\circ T^j_{\mathbf{0}})\Big) \\ &= \varepsilon^2 \sum_{j=1}^{\infty} \mu_{\mathbf{E}}(\bar{H}\cdot\bar{H}\circ T^j_{\mathbf{0}}) + o(\varepsilon^2). \end{aligned}$$

By exponential decay of correlations, the series in the last expression converges. Finally, we use (5.1) to change the measure from $\mu_{\mathbf{E}}$ to μ_0 since all the functions involved are admissible with respect to the norms we have defined. Thus

$$e''_{\mathbf{E}}(0) = \varepsilon^2 \sigma^2_H + o(\varepsilon^2),$$

where $\sigma^2_H = \mu_0(\bar{H}^2) + 2\sum_{j=1}^{\infty} \mu_0(\bar{H}\cdot\bar{H}\circ T^j_{\mathbf{0}})$.

Next, we show that in fact $e_{\mathbf{E}}(a)$ is strictly convex for $a \in [-a_0, 1 + a_0]$ whenever $\sigma^2_{\mathbf{E}} > 0$.

In order to compute $e'_{\mathbf{E}}(a)$ and $e''_{\mathbf{E}}(a)$ at $a \neq 0$, let

$$e_a(t) := \lim_{n\to\infty} \frac{1}{n} \log \mu_{\mathbf{E},a}(e^{-tS_n s}),$$

i.e. e_a is the moment generating function for $\mu_{\mathbf{E},a}$. Note that

$$\begin{aligned}\mu_{\mathbf{E},a}(e^{-tS_n s}) &= \langle e^{-tS_n s} \cdot h_{\mathbf{E},a}, \nu_{\mathbf{E},a}\rangle = \lambda_{\mathbf{E},a}^{-n} \langle e^{-tS_n s} \mathcal{L}_a^n h_{\mathbf{E},a}, \nu_{\mathbf{E},a}\rangle \\ &= \lambda_{\mathbf{E},a}^{-n} \langle \mathcal{L}_{t+a}^n h_{\mathbf{E},a}, \nu_{\mathbf{E},a}\rangle.\end{aligned}$$

Therefore,

$$\begin{aligned}e_a(t) &= \lim_{n\to\infty} \frac{1}{n} \log \lambda_{\mathbf{E},a}^{-n} \langle \mathcal{L}_{t+a}^n h_{\mathbf{E},a}, \nu_{\mathbf{E},a}\rangle = \lim_{n\to\infty} \frac{1}{n} \log \langle \mathcal{L}_{t+a}^n h_{\mathbf{E},a}, \nu_{\mathbf{E},a}\rangle - \log \lambda_{\mathbf{E},a} \\ &= \lim_{n\to\infty} \frac{1}{n} \log \langle \lambda_{\mathbf{E},a+t}^n c_{\mathbf{E},a+t}(h_{\mathbf{E},a}) h_{\mathbf{E},a+t} + R_{\mathbf{E},a+t}^n(h_{\mathbf{E},a}), \nu_{\mathbf{E},a}\rangle - \log \lambda_{\mathbf{E},a} \\ &= e_{\mathbf{E}}(a+t) - e_{\mathbf{E}}(a),\end{aligned}$$

where we have used the fact that $e_{\mathbf{E}}(a) = \log \lambda_{\mathbf{E},a}$. Differentiating with respect to t gives $e'_{\mathbf{E}}(a) = e'_a(0)$ and $e''_{\mathbf{E}}(a) = e''_a(0)$. The computation of $e'_a(0)$ and $e''_a(0)$ are the same as the case $e'_{\mathbf{E}}(0)$ and $e''_{\mathbf{E}}(0)$, with $\mu_{\mathbf{E},a}$ in place of $\mu_{\mathbf{E}}$.

Notice that $\mathcal{L}_{T_{\mathbf{0}},a} = \mathcal{L}_{T_{\mathbf{0}},0}$ for each $a \in \mathbb{R}$ since when $\mathbf{E} = (0,0)$, $s = 0$. It follows that $\mu_{\mathbf{0},a} = \mu_0$ for all $a \in \mathbb{R}$. Thus by the continuity of $\mu_{\mathbf{E},a}$ in $\mathbf{E}$ for each fixed a (Lemma 4.6), we have $e''_a(0) > 0$ for all $a \in [-a_0, 1 + a_0]$ and $\varepsilon < \varepsilon_1$ if ε_1 is chosen sufficiently small.

The positivity of the entropy production rate follows then from the symmetry and strict convexity. Indeed, suppose the entropy production rate $-e'_{\mathbf{E}}(0) = 0$. Then since $e_{\mathbf{E}}(0) = e_{\mathbf{E}}(1) = 0$ by the symmetry proved in item (2), convexity and analyticity imply that $e_{\mathbf{E}}(a) = 0$ for all $a \in [0, 1]$. This contradicts strict convexity, i.e. it contradicts that $e''_{\mathbf{E}}(a) > 0$ for all $a \in [0, 1]$.

It remains to prove that $\sigma_{\mathbf{E}}^2 > 0$ if and only if s is not a coboundary. If $s = \psi \circ T_{\mathbf{E}} - \psi + C$ for some ψ then $e_{\mathbf{E}}(a) = aC$ and so trivially $e''_{\mathbf{E}}(a) = 0$ for all a. The converse requires a more substantial proof. In order to prove it, we will invoke the following abstract version of the Central Limit Theorem for invertible systems, following the classical martingale approach of Gordin [G].

Theorem 5.1 ([V]). *Let $(X, \mathcal{A}, \mu)$ be a probability space, $\phi \in L^2(\mu)$ be such that $\int \phi \, d\mu = 0$, and $\theta : X \to X$ be an invertible map such that both θ and θ^{-1} are measurable, and μ is θ-invariant and ergodic. Let $\mathcal{A}_0 \subset \mathcal{A}$ be such that $\mathcal{A}_n = \theta^{-n}(\mathcal{A}_0)$, $n \in \mathbb{Z}$, is a non-increasing sequence of σ-algebras. Assume that*

$$\sum_{n=0}^{\infty} \|E(\phi|\mathcal{A}_n)\|_{L^2(\mu)} < \infty \quad \textit{and} \quad \sum_{n=0}^{\infty} \|\phi - E(\phi|\mathcal{A}_{-n})\|_{L^2(\mu)} < \infty, \tag{5.3}$$

and let $\sigma^2 = \int \phi^2 \, d\mu + 2 \sum_{j=1}^{\infty} \int \phi \cdot \phi \circ \theta^j \, d\mu$.

Then σ is finite, and $\sigma = 0$ if and only if $\phi = u \circ \theta - u$ for some $u \in L^2(\mu)$. Moreover, if $\sigma^2 > 0$, then $n^{-1/2} S_n \phi$ converges weakly to $\mathcal{N}(0, \sigma^2)$.

We will apply this theorem with $\theta = T_{\mathbf{E}}$, $\mu = \mu_{\mathbf{E}}$, $\phi = \bar{s}$ and $\sigma = \sigma_{\mathbf{E}}$. Let $\mathcal{A}_0$ be the sigma-algebra generated by the ($\mu_{\mathbf{E}}$-mod 0) partition of M into maximal homogeneous local stable manifolds for $T_{\mathbf{E}}$.[9] Then $\mathcal{A}_n = T_{\mathbf{E}}^{-n}(\mathcal{A}_0)$ is a decreasing sequence of sigma-algebras, as required.

With these definitions, the second condition in (5.3) is a simple consequence of the uniform contraction of stable manifolds. Denoting by $V^s(x)$ the maximal local stable manifold of x, we note that $E(\bar{s}|\mathcal{A}_{-n})$ is constant on curves of the form $T_{\mathbf{E}}^n(V^s(x))$; these are the elements of $\mathcal{A}_{-n}$ and their length is bounded by $C\Lambda^{-n}$ for some uniform $C > 0$ since stable manifolds have length uniformly bounded above due to the discontinuities of $T_{\mathbf{E}}$. In fact, since $\bar{s}$ is continuous on such curves,[10] $E(\bar{s}|\mathcal{A}_{-n})(x) = \bar{s}(y)$ for some $y \in \mathcal{A}_{-n}(x)$, the element of $\mathcal{A}_{-n}$ containing x. Thus

$$
\begin{aligned}
|\bar{s}(x) - E(\bar{s}|\mathcal{A}_{-n})(x)| &= |\bar{s}(x) - \bar{s}(y)| = |\log(1+\varepsilon H(x)) - \log(1+\varepsilon H(y))| \\
&\le \frac{\varepsilon}{1 - C_H\varepsilon} C_s^{\alpha_1}(H) d(x,y)^{\alpha_1} \le C'\Lambda^{-n\alpha_1},
\end{aligned}
$$

where we have used Lemma 4.2 and $C_s^{\alpha_1}(\cdot)$ denotes the Hölder constant along stable manifolds with exponent α_1. This estimate implies that the L^∞-norm, and therefore the L^2-norm, of $\bar{s}(x) - E(\bar{s}|\mathcal{A}_{-n})$ decays at an exponential rate and so the second sum in (5.3) converges.

The first sum in (5.3) entails a more subtle calculation. In principle, it follows from exponential decay of correlations for $\mu_{\mathbf{E}}$; however, it requires exponential decay of correlations against observables in $L^2(\mathcal{A}_0, \mu_{\mathbf{E}})$, which is a larger class than is at first available in the framework of our Banach spaces; here, $L^2(\mathcal{A}_0, \mu_{\mathbf{E}})$ is the set of L^2 functions that are measurable with respect to $\mathcal{A}_0$. To see this, we will use the dual version of the L^2-norm,

$$
\begin{aligned}
\|E(\bar{s}|\mathcal{A}_n)\|_{L^2(\mu_{\mathbf{E}})} &= \sup\left\{\int \bar{s}\,\phi\, d\mu_{\mathbf{E}} : \phi \in L^2(\mathcal{A}_n, \mu_{\mathbf{E}}) \text{ with } \|\phi\|_{L^2(\mu_{\mathbf{E}})} = 1\right\} \\
&= \sup\left\{\int \bar{s}\,\psi \circ T_{\mathbf{E}}^n\, d\mu_{\mathbf{E}} : \psi \in L^2(\mathcal{A}_0, \mu_{\mathbf{E}}) \text{ with } \|\psi\|_{L^2(\mu_{\mathbf{E}})} = 1\right\}.
\end{aligned} \tag{5.4}
$$

In order for this last integral to decay exponentially in n as a result of the spectral gap for $\mathcal{L}_{T_{\mathbf{E}}} = \mathcal{L}_{T_{\mathbf{E}},0}$, we would like $\bar{s} \in \mathcal{B}$ and $\psi \in \mathcal{B}'$, the dual to $\mathcal{B}$. Unfortunately, the first statement is false and the second statement needs some work to justify. Also, note that we can expect the correlations to decay to 0 in the above expression since $\mu_{\mathbf{E}}(\bar{s}) = 0$. It is not necessary that $\mu_{\mathbf{E}}(\psi) = 0$ as well.

As noted in the proof of Lemma 4.4, as an admissible potential $\bar{s}$ is Hölder continuous on connected components of $M \backslash \mathcal{S}_1^{T_{\mathbf{E}}}$ and so does not satisfy the assumptions of Lemma 3.2; however, $\bar{s} \circ T_{\mathbf{E}}^{-1}$ is $\alpha_1/2$-Hölder continuous on connected components of $M \backslash \mathcal{S}_{-1}^{T_{\mathbf{E}}}$ by (4.11). Thus by Lemma 3.2, since $\gamma < 1/7 \le \alpha_1/(\alpha_1 + 2)$ in the definition

[9] This partition is measurable since it is has a countable generator: $\cup_{n\ge 1}$\{connected components of $M \backslash \mathcal{S}_n^{T_{\mathbf{E}},\mathbb{H}}$\}. See for example, [CM, Section 5.1].

[10] Indeed, $\bar{s}$ is Hölder continuous on each connected component of $M \backslash \mathcal{S}_1^{T_{\mathbf{E}}}$, and local stable manifolds cannot cross $\mathcal{S}_1^{T_{\mathbf{E}}}$, otherwise they would be cut in forward time under $T_{\mathbf{E}}$, which would contradict the definition of stable manifold.

of the norms, both $\bar{s} \circ T_{\mathbf{E}}^{-1}$ and $\bar{s} \circ T_{\mathbf{E}}^{-1} h_{\mathbf{E}} \in \mathcal{B}$, where $h_{\mathbf{E}}$ is the right eigenvector of $\mathcal{L}_{T_{\mathbf{E}}}$. Since $\int \bar{s}\, \psi \circ T_{\mathbf{E}}^n \, d\mu_{\mathbf{E}} = \int \bar{s} \circ T_{\mathbf{E}}^{-1} \, \psi \circ T_{\mathbf{E}}^{n-1} \, d\mu_{\mathbf{E}}$, it suffices to work with $\bar{s} \circ T_{\mathbf{E}}^{-1}$.

Notice also that since we are in the case $a = 0$, the conformal measure is μ_0, i.e. $\mathcal{L}_{T_{\mathbf{E}}}^* \mu_0 = \mu_0$. Thus for $n \geq 0$ and $\psi \in \mathcal{C}^\alpha(T_{\mathbf{E}}^{-n} \mathcal{W}^s)$, we have $\mu_{\mathbf{E}} = h_{\mathbf{E}} \otimes \mu_0$ and $\mu_{\mathbf{E}}(\psi) = \langle h_{\mathbf{E}}, \psi \mu_0 \rangle$.

In order to estimate the expression in (5.4), we shall need two lemmas. Let $B_0(\mathcal{A}_0)$ denote the set of bounded functions on M, which are measurable with respect to $\mathcal{A}_0$.

Lemma 5.2. *Suppose there exist $C > 0$ (depending on $\bar{g}$) and $\rho < 1$ such that*

$$\mu_{\mathbf{E}}(\bar{s} \cdot \psi \circ T_{\mathbf{E}}^n) \leq C \rho^n |\psi|_\infty, \quad \text{for all } \psi \in B_0(\mathcal{A}_0), \tag{5.5}$$

where $|\psi|_\infty = \sup_{x \in M} |\psi(x)|$. Then there exists $C' > 0$ such that

$$\mu_{\mathbf{E}}(\bar{s} \cdot \psi \circ T_{\mathbf{E}}^n) \leq C' \rho^{n/2} |\psi|_{L^2(\mu_{\mathbf{E}})}^2, \quad \text{for all } \psi \in L^2(\mathcal{A}_0, \mu_{\mathbf{E}}).$$

The following lemma is a strengthening of Lemma 3.1(i). It shows that the estimate of that lemma holds true in the limit as $n \to \infty$.

Lemma 5.3. *There exists $C > 0$ such that for any $h \in \mathcal{B}_w$ and any bounded function ψ,*

$$|h(\psi)| \leq C |h|_w (|\psi|_\infty + C^\alpha_{\mathcal{A}_0}(\psi)),$$

where $C^\alpha_{\mathcal{A}_0}(\cdot)$ denotes the Hölder constant of ψ with exponent α measured along curves in $\mathcal{A}_0$.

We postpone the proofs of the lemmas and first show how they allow us to complete the proof of Theorem 2.4. For $\psi \in \mathcal{B}_0(\mathcal{A}_0)$, $C^\alpha_{\mathcal{A}_0}(\psi) = 0$ since ψ is constant on curves in $\mathcal{A}_0$. We estimate the correlations using Lemma 5.3,

$$\begin{aligned} \left| \int \bar{s}\, \psi \circ T_{\mathbf{E}}^n \, d\mu_0 \right| &= \left| \int \bar{s} \circ T_{\mathbf{E}}^{-1} \, \psi \circ T_{\mathbf{E}}^{n-1} \, d\mu_{\mathbf{E}} \right| = \left| \langle \bar{s} \circ T_{\mathbf{E}}^{-1} \, h_{\mathbf{E}}, \psi \circ T_{\mathbf{E}}^{n-1} \, \mu_0 \rangle \right| \\ &= \left| \langle \bar{s} \circ T_{\mathbf{E}}^{-1} \, h_{\mathbf{E}}, (\mathcal{L}_{T_{\mathbf{E}}}^*)^{n-1} (\psi \mu_0) \rangle \right| = \left| \langle \mathcal{L}_{T_{\mathbf{E}}}^{n-1} (\bar{s} \circ T_{\mathbf{E}}^{-1} \, h_{\mathbf{E}}), \psi \mu_0 \rangle \right| \\ &\leq C |\mathcal{L}_{T_{\mathbf{E}}}^{n-1} (\bar{s} \circ T_{\mathbf{E}}^{-1} h_{\mathbf{E}})|_w (|\psi|_\infty + H^\alpha_{\mathcal{A}_0}(\psi)) \\ &\leq C \| R_{\mathbf{E}}^{n-1} (\bar{s} \circ T_{\mathbf{E}}^{-1} h_{\mathbf{E}}) \|_{\mathcal{B}} |\psi|_\infty \\ &\leq C \rho^n \| h_{\mathbf{E}} \|_{\mathcal{B}} |\bar{s} \circ T_{\mathbf{E}}^{-1}|_{\mathcal{C}^{\alpha_1/2}(\mathcal{P}_{-1})} |\psi|_\infty, \end{aligned} \tag{5.6}$$

for some $\rho < 1$ where $|\bar{s} \circ T_{\mathbf{E}}^{-1}|_{\mathcal{C}^{\alpha_1/2}(\mathcal{P}_{-1})}$ denotes the Hölder constant of $\bar{s}$ on elements of the partition formed by the connected components of $M \backslash \mathcal{S}_{-1}^{T_{\mathbf{E}}}$, and we have used (4.14) and the fact that $\Pi_{\mathbf{E}}(\bar{s} \circ T_{\mathbf{E}}^{-1} h_{\mathbf{E}}) = 0$ since $\langle \bar{s} \circ T_{\mathbf{E}}^{-1} h_{\mathbf{E}}, \mu_0 \rangle = \mu_{\mathbf{E}}(\bar{s} \circ T_{\mathbf{E}}^{-1}) = \mu_{\mathbf{E}}(\bar{s}) = 0$.

From (5.6), we see that $\bar{s}$ has uniform exponential decay of correlations against $\psi \in B_0(\mathcal{A}_0)$ and so satisfies the hypotheses of Lemma 5.2. It follows that $\bar{s}$ enjoys a uniform exponential rate of decay of correlations against $\psi \in L^2(\mathcal{A}_0, \mu_{\mathbf{E}})$, so by (5.4), this yields an exponential decay in the L^2-norm of $E(\bar{s}|\mathcal{A}_n)$. We conclude that the first series in (5.3) converges. Since the hypotheses of Theorem 5.1 are verified, it follows that $\sigma_{\mathbf{E}}^2 = 0$ if and only if $\bar{s} = u \circ T_{\mathbf{E}} - u$ for some $u \in L^2(\mu_{\mathbf{E}})$, and the proof of Theorem 2.4 is complete. □

Proof of Lemma 5.2. Let $\psi \in L^2(\mathcal{A}_0, \mu_{\mathbf{E}})$ be arbitrary. For $L \in \mathbb{R}^+$, define $\psi_L(x) = \psi(x)$ when $|\psi(x)| \le L$ and $\psi_L(x) = 0$ otherwise. Clearly, $\psi_L \in B_0(\mathcal{A}_0)$ and $|\psi_L|_\infty \le L$. Now for $n \in \mathbb{N}$,

$$\left|\int \bar{s} \cdot \psi \circ T_{\mathbf{E}}^n \, d\mu_{\mathbf{E}}\right| \le \left|\int \bar{s} \cdot \psi_L \circ T_{\mathbf{E}}^n \, d\mu_{\mathbf{E}}\right| + |\bar{s}|_\infty \int |\psi - \psi_L| \, d\mu_{\mathbf{E}}. \tag{5.7}$$

To bound the second term on the right side of (5.7), note that

$$\int |\psi - \psi_L| \, d\mu_{\mathbf{E}} \le \int 1_{|\psi|>L} \cdot |\psi| \, d\mu_{\mathbf{E}} \le \mu_{\mathbf{E}}(|\psi| > L)^{1/2} |\psi|_{L^2(\mu_{\mathbf{E}})},$$

while $\mu(|\psi| > L) = \mu(\psi^2 > L^2) \le L^{-2}|\psi|_{L^2}^2$, by Markov's inequality. Using (5.5) for the first term on the right side of (5.7), we obtain,

$$\left|\int \bar{s} \cdot \psi \circ T_{\mathbf{E}}^n \, d\mu_{\mathbf{E}}\right| \le C\rho^n L + L^{-1}|\bar{s}|_\infty |\psi|_{L^2(\mu_{\mathbf{E}})}^2 \le \rho^{n/2}(C + |\bar{s}|_\infty |\psi|_{L^2(\mu_{\mathbf{E}})}^2), \tag{5.8}$$

if we set $L = \rho^{-n/2}$, and the lemma is proved. □

Proof of Lemma 5.3. Due to the density of $\mathcal{C}^1(M)$ in $\mathcal{B}_w$, it suffices to prove the lemma for $h \in \mathcal{C}^1(M)$.

On each component M_i of M, $i = 1, \dots, d$, we disintegrate the smooth measure μ_0 on elements of $\mathcal{A}_0$. Since elements of $\mathcal{A}_0$ are homogeneous stable manifolds, the decomposition respects the boundaries of the homogeneity strips. Let $\mathcal{A}_{0,i} = \{W_\xi\}_{\xi \in \Xi_i}$ denote the set of homogeneous local stable manifolds in M_i with index set Ξ_i. The disintegration of μ_0 on elements of $\mathcal{A}_{0,i}$ yields conditional densities η_ξ on W_ξ, normalized so that $\int_{W_\xi} \eta_\xi \, dm_{W_\xi} = 1$, and a factor measure $\hat{\mu}_0$ on the index set Ξ_i. By [CM, Corollary 5.30], η_ξ is (uniformly in ξ) log-Hölder continuous with exponent $1/3$. Now,

$$h(\psi) = \int_M h\psi \, d\mu_0 = \sum_i \int_{\Xi_i} \int_{W_\xi} h \, \psi \, \eta_\xi \, dm_{W_\xi} d\hat{\mu}_0(\xi). \tag{5.9}$$

On each W_ξ, we estimate using the weak norm.

$$\left|\int_{W_\xi} h \, \psi \, \eta_\xi \, dm_\xi\right| \le |h|_w |\psi|_{\mathcal{C}^\alpha(W_\xi)} |\eta_\xi|_{\mathcal{C}^\alpha(W_\xi)}.$$

Due to the log-Hölder regularity of η_ξ, there exists a constant $C_\eta > 0$ such that

$$|\eta_\xi(x) - \eta_\xi(y)| \le C_\eta |\eta_\xi(x)| d(x,y)^{1/3} \le C' |W_\xi|^{-1} d(x,y)^{1/3},$$

where the bound on the sup-norm of η_ξ comes from the normalization of the conditional measures. Putting these estimates into (5.9) we obtain,

$$|h(\psi)| \le C' |h|_w (|\psi|_\infty + H^\alpha_{\mathcal{A}_0}(\psi)) \sum_i \int_{\Xi_i} |W_\xi|^{-1} \, d\hat{\mu}_0(\xi).$$

This last integral is precisely the integral that characterizes the Z-function for a standard family [CM, Section 7.4] which measures the prevalence of short curves in that family. Since the disintegration of μ_0 on maximal homogeneous stable manifolds creates a proper family,[11] this integral is finite (see [CM, Exercise 7.22] for the decomposition using stable manifolds for the unperturbed billiard T_0 and [CZ,CZZ] for the decomposition using stable manifolds for the perturbed billiard $T_{\mathbf{E}}$). □

Appendix A. Lasota–Yorke Estimates

A.1. Preliminary estimates. Before proving the Lasota–Yorke inequalities, we show how **(H1)–(H5)** imply several other uniform properties for our class of maps $\mathcal{F}$. In particular, we will be interested in iterating the one-step expansion given by **(H3)**. We recall the estimates we need from [DZ1, Section 3.2].

Let $T \in \mathcal{F}$ and $W \in \mathcal{W}^s$. Let V_i denote the maximal connected components of $T^{-1}W$ after cutting due to singularities and the boundaries of the homogeneity strips. To ensure that each component of $T^{-1}W$ is in $\mathcal{W}^s$, we subdivide any of the long pieces V_i whose length is $> \delta_0$, where δ_0 is chosen in (2.15). This process is then iterated so that given $W \in \mathcal{W}^s$, we construct the components of $T^{-n}W$, which we call the n^{th} generation $\mathcal{G}_n(W)$, inductively as follows. Let $\mathcal{G}_0(W) = \{W\}$ and suppose we have defined $\mathcal{G}_{n-1}(W) \subset \mathcal{W}^s$. First, for any $W' \in \mathcal{G}_{n-1}(W)$, we partition $T^{-1}W'$ into at most countably many pieces W'_i so that T is smooth on each W'_i and each W'_i is a homogeneous stable curve. If any W'_i have length greater than δ_0, we subdivide those pieces into pieces of length between $\delta_0/2$ and δ_0. We define $\mathcal{G}_n(W)$ to be the collection of all pieces $W^n_i \subset T^{-n}W$ obtained in this way. Note that each W^n_i is in $\mathcal{W}^s$ by **(H2)**.

At each iterate of T^{-1}, typical curves in $\mathcal{G}_n(W)$ grow in size, but there exist a portion of curves which are trapped in tiny homogeneity strips and in the infinite horizon case, stay too close to the infinite horizon points. In Lemma A.1, we make precise the sense in which the proportion of curves that never grow to a fixed length decays exponentially fast.

For $W \in \mathcal{W}^s$, $n \geq 0$, and $0 \leq k \leq n$, let $\mathcal{G}_k(W) = \{W^k_i\}$ denote the k^{th} generation pieces in $T^{-k}W$. Let $B_k(W) = \{i : |W^k_i| < \delta_0/3\}$ and $L_k(W) = \{i : |W^k_i| \geq \delta_0/3\}$ denote the index of the short and long elements of $\mathcal{G}_k(W)$, respectively. We consider $\{\mathcal{G}_k\}_{k=0}^n$ as a tree with W as its root and $\mathcal{G}_k$ as the k^{th} level.

At level n, we group the pieces as follows. Let $W^n_{i_0} \in \mathcal{G}_n(W)$ and let $W^k_j \in L_k(W)$ denote the most recent long "ancestor" of $W^n_{i_0}$, i.e. $k = \max\{0 \leq \ell \leq n : T^{n-\ell}(W^n_{i_0}) \subset W^\ell_j \text{ and } j \in L_\ell\}$. If no such ancestor exists, set $k = 0$ and $W^k_j = W$. Note that if $W^n_{i_0}$ is long, then $W^k_j = W^n_{i_0}$. Let

$$\mathcal{I}_n(W^k_j) = \{i : W^k_j \in L_k(W) \text{ is the most recent long ancestor of } W^n_i \in \mathcal{G}_n(W)\}.$$

The set $\mathcal{I}_n(W)$ represents those curves W^n_i that belong to short pieces in $\mathcal{G}_k(W)$ at each time step $1 \leq k \leq n$, i.e. such W^n_i are never part of a piece that has grown to length $\geq \delta_0/3$.

We collect the necessary complexity estimates in the following lemma.

Lemma A.1. *Let $W \in \mathcal{W}^s$, $T \in \mathcal{F}$ and for $n \geq 0$, let $\mathcal{I}_n(W)$ and $\mathcal{G}_n(W)$ be defined as above. There exist constants $C_1, C_2, C_3 > 0$, independent of W and T, such that for any $n \geq 0$,*

[11] In fact, Example 7.21 and Exercise 7.22 of [CM] are stated in terms of the disintegration of μ_0 on maximal homogeneous unstable manifolds. Using the reversibility of $T_{\mathbf{E}}$, the analogous properties hold for maximal homogeneous stable manifolds.

(a) $\sum_{i \in \mathcal{I}_n(W)} |J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)} \le C_1 \theta_*^n$;

(b) $\sum_{W_i^n \in \mathcal{G}_n(W)} |J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)} \le C_2$;

(c) *for any* $0 \le \varsigma \le 1$, $\sum_{W_i^n \in \mathcal{G}_n(W)} \frac{|W_i^n|^\varsigma}{|W|^\varsigma} |J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)} \le C_2^{1-\varsigma}$.

(d) *for* $\varsigma > 1/2$, $\sum_{W_i^n \in \mathcal{G}_n(W)} |J_{W_i^n} T^n|^\varsigma_{\mathcal{C}^0(W_i^n)} \le C_3^n$, *where* C_3 *depends on* ς.

Proof. Item (a) follows from the one-step expansion **(H3)** by induction as in [DZ1, Lemma 3.1]. Items (b) and (c) are precisley [DZ1, Lemmas 3.2 and 3.3].

For item (d), we first prove that the claimed estimate holds for $n = 1$. Indeed, due to **(H1)**, the expansion for each stable curve landing in a homogeneity strip $\mathbb{H}_k$ under T^{-1} is of the order of k^{-2}. If $T^{-1}W$ crosses a countable number of singularity curves, the sum of the expansion factors is uniformly bounded as long as $\varsigma > 1/2$. Since there are only finitely many genuine singularity curves in $\mathcal{S}_{-1}^T$ (not counting homogeneity strips), the required sum is uniformly bounded for all $W \in \mathcal{W}^s$ with $|W| \le \delta_0$. The estimate for general n follows by induction as in [DZ1, Lemma 3.4]. □

Next we state a distortion bound for the stable Jacobian of T along different stable curves in the following context. Let $W^1, W^2 \in \mathcal{W}^s$ and suppose there exist $U^k \subset T^{-n}W^k$, $k = 1, 2$, such that for $0 \le i \le n$,

(i) $T^iU^k \in \mathcal{W}^s$ and the curves T^iU^1 and T^iU^2 lie in the same homogeneity strip;
(ii) U^1 and U^2 can be put into a 1-1 correspondence by a smooth foliation $\{\gamma_x\}_{x \in U^1}$ of curves $\gamma_x \in \widehat{\mathcal{W}}^u$ such that $\{T^n\gamma_x\} \subset \widehat{\mathcal{W}}^u$ creates a 1-1 correspondence between T^nU^1 and T^nU^2;
(iii) $|T^i\gamma_x| \le 2\max\{|T^iU^1|, |T^iU^2|\}$, for all $x \in U^1$.

Let $J_{U^k}T^n$ denote the stable Jacobian of T^n along the curve U^k with respect to arclength. The following lemma was proved in [DZ2].

Lemma A.2. *In the setting above, for* $x \in U^1$, *define* $x^* \in \gamma_x \cap U^2$. *There exists* $C_0 > 0$, *independent of* $T \in \mathcal{F}$, $W \in \mathcal{W}^s$ *and* $n \ge 0$ *such that*

(a) $d_{\mathcal{W}^s}(U^1, U^2) \le C_0 \Lambda^{-n} d_{\mathcal{W}^s}(W^1, W^2)$;

(b) $\left| \frac{J_{U^1}T^n(x)}{J_{U^2}T^n(x^*)} - 1 \right| \le C_0[d(T^nx, T^nx^*)^{1/3} + \theta(T^nx, T^nx^*)]$,

where $\theta(T^nx, T^nx^*)$ *is the angle formed by the tangent lines of* T^nU^1 *and* T^nU_2 *at* T^nx *and* T^nx^*, *respectively.*

To prove Proposition 4.1, we fix $T \in \mathcal{F}$ and prove the required Lasota–Yorke inequalities (4.1)–(4.3). It is shown in Lemma 3.3 that $\mathcal{L}_{T,g}$ is a continuous operator on both $\mathcal{B}$ and $\mathcal{B}_w$ so that it suffices to prove the inequalities for $h \in \mathcal{C}^1(M)$. They extend to the completions by continuity. Our purpose now is to show how they depend explicitly on the uniform constants given by **(H1)–(H5)** and do not require additional information.

A.2. Estimating the weak norm. Let $h \in \mathcal{C}^1(M)$, $W \in \mathcal{W}^s$ and $\psi \in \mathcal{C}^\alpha(W)$ such that $|\psi|_{\mathcal{C}^\alpha(W)} \le 1$. For brevity, we define

$$\hat{g} = g - \log J_{\mu_0}T, \quad \text{so that} \quad e^{S_n\hat{g}} = e^{S_ng}(J_{\mu_0}T^n)^{-1}.$$

For $n \geq 0$, we write,

$$\int_W \mathcal{L}^n_{T,g} h\, \psi\, dm_W = \sum_{W^n_i \in \mathcal{G}_n(W)} \int_{W^n_i} h e^{S_n \hat{g}} J_{W^n_i} T^n \psi \circ T^n dm_{W^n_i} \tag{A.1}$$

where $J_{W^n_i} T^n$ denotes the Jacobian of T^n along W^n_i.

Using the definition of the weak norm on each W^n_i, we estimate (A.1) by

$$\int_W \mathcal{L}^n_{T,g} h\, \psi\, dm_W \;\leq \sum_{W^n_i \in \mathcal{G}_n} |h|_w |J_{W^n_i} T^n|_{\mathcal{C}^\alpha(W^n_i)} |e^{S_n \hat{g}}|_{\mathcal{C}^\alpha(W^n_i)} |\psi \circ T^n|_{\mathcal{C}^\alpha(W^n_i)}. \tag{A.2}$$

Using the bounded distortion property **(H4)**, we estimate,

$$|J_{W^n_i} T^n|_{\mathcal{C}^\alpha(W^n_i)} \leq (1 + C_d) |J_{W^n_i} T^n|_{\mathcal{C}^0(W^n_i)}, \tag{A.3}$$

and similarly for $|(J_{\mu_0} T^n)^{-1}|_{\mathcal{C}^\alpha(W^n_i)}$. Next, for the potential g and $x, y \in W^n_i$,

$$\begin{aligned} |e^{S_n g(x)} - e^{S_n g(y)}| &\leq |e^{S_n g}|_{\mathcal{C}^0(W^n_i)} |S_n g(x) - S_n g(y)| \\ &\leq |e^{S_n g}|_{\mathcal{C}^0(W^n_i)} |g|_{\mathcal{C}^\alpha(\mathcal{P}_1)} \sum_{i=0}^{n-1} C_e \Lambda^{-i\alpha} |x - y|^\alpha, \end{aligned}$$

so that $|e^{S_n g}|_{\mathcal{C}^\alpha(W^n_i)} \leq C_g |e^{S_n g}|_{\mathcal{C}^0(W^n_i)}$, where $C_g := 1 + C_e |g|_{\mathcal{C}^\alpha(\mathcal{P}_1)} \sum_{i=0}^\infty \Lambda^{-i\beta}$. This, together with (A.3) applied to $(J_{\mu_0} T^n)^{-1}$, yields,

$$|e^{S_n \hat{g}}|_{\mathcal{C}^\alpha(W^n_i)} \leq C_g |e^{S_n g}|_{\mathcal{C}^0(W^n_i)} (1 + C_d) |(J_{\mu_0} T^n)^{-1}|_{\mathcal{C}^0(W^n_i)} \leq C_g (1 + C_d)^2 |e^{S_n \hat{g}}|_{\mathcal{C}^0(W^n_i)}, \tag{A.4}$$

where we have used again the bounded distortion of $J_{\mu_0} T^n$ to combine the two $\mathcal{C}^0$-norms.

Finally, we esimate the norm of $\psi \circ T^n$, again using **(H1)**. For $x, y \in W^n_i$,

$$\frac{|\psi(T^n x) - \psi(T^n y)|}{d_W(T^n x, T^n y)^\alpha} \cdot \frac{d_W(T^n x, T^n y)^\alpha}{d_W(x, y)^\alpha} \leq |\psi|_{\mathcal{C}^\alpha(W)} |J_{W^n_i} T^n|^\alpha_{\mathcal{C}^0(W^n_i)} \leq C_e \Lambda^{-\alpha n} |\psi|_{\mathcal{C}^\alpha(W)}, \tag{A.5}$$

so that $|\psi \circ T^n|_{\mathcal{C}^\alpha(W^n_i)} \leq C_e |\psi|_{\mathcal{C}^\alpha(W)} \leq C_e$. We use this estimate together with (A.3) and (A.4) to bound (A.2) by

$$\begin{aligned} \int_W \mathcal{L}^n_{T,g} h\, \psi\, dm_W &\leq C_e (1 + C_d)^3 C_g |e^{S_n \hat{g}}|_\infty |h|_w \sum_{W^n_i \in \mathcal{G}_n} |J_{W^n_i} T^n|_{\mathcal{C}^0(W^n_i)} \\ &\leq C' C_g |e^{S_n \hat{g}}|_\infty |h|_w, \end{aligned}$$

where $C' = C_e (1 + C_d)^3 C_2$ and we have used Lemma A.1(b) for the last inequality. Taking the supremum over all $W \in \mathcal{W}^s$ and $\psi \in \mathcal{C}^\alpha(W)$ with $|\psi|_{\mathcal{C}^\alpha(W)} \leq 1$ yields (4.1) expressed with uniform constants given by **(H1)–(H5)**.

A.3. Estimating the strong stable norm. Let $W \in \mathcal{W}^s$ and let W_i^n denote the elements of $\mathcal{G}_n(W)$ as defined above. For $\psi \in \mathcal{C}^\beta(W)$, $|\psi|_{\mathcal{C}^\beta(W)} \le |W|^{-p}$, define $\overline{\psi}_i = |W_i^n|^{-1} \int_{W_i^n} \psi \circ T^n \, dm_W$. Using equation (A.1), we write

$$\int_W \mathcal{L}_{T,g}^n h \, \psi \, dm_W = \sum_i \int_{W_i^n} e^{S_n \hat{g}} h \cdot J_{W_i^n} T^n \cdot (\psi \circ T^n - \overline{\psi}_i) \, dm_W + \overline{\psi}_i \int_{W_i^n} h e^{S_n \hat{g}} \cdot J_{W_i^n} T^n \, dm_W. \tag{A.6}$$

To estimate the first term of (A.6), we first estimate $|\psi \circ T^n - \overline{\psi}_i|_{\mathcal{C}^\beta(W_i^n)}$. If $H_W^\beta(\psi)$ denotes the Hölder constant of ψ along W, then Eq. (A.5) implies

$$\frac{|\psi(T^n x) - \psi(T^n y)|}{d_W(x, y)^\beta} \le C_e \Lambda^{-n\beta} H_W^\beta(\psi) \tag{A.7}$$

for any $x, y \in W_i^n$. Since $\overline{\psi}_i$ is constant on W_i^n, we have $H_{W_i^n}^\beta(\psi \circ T^n - \overline{\psi}_i) \le C_e \Lambda^{-\beta n} H_W^\beta(\psi)$. To estimate the $\mathcal{C}^0$ norm, note that $\overline{\psi}_i = \psi \circ T^n(y_i)$ for some $y_i \in W_i^n$. Thus for each $x \in W_i^n$,

$$\begin{aligned} |\psi \circ T^n(x) - \overline{\psi}_i| = |\psi \circ T^n(x) - \psi \circ T^n(y_i)| &\le H_{W_i^n}^\beta(\psi \circ T^n) |W_i^n|^\beta \\ &\le C_e H_W^\beta(\psi) \Lambda^{-\beta n}. \end{aligned}$$

This estimate together with (A.7) and the fact that $|W|^p |\psi|_{\mathcal{C}^\beta(W)} \le 1$, implies

$$|\psi \circ T^n - \overline{\psi}_i|_{\mathcal{C}^\beta(W_i^n)} \le C_e \Lambda^{-\beta n} |\psi|_{\mathcal{C}^\beta(W)} \le C_e \Lambda^{-\beta n} |W|^{-p}. \tag{A.8}$$

We apply (A.3), (A.4) and (A.8) and the definition of the strong stable norm to the first term of (A.6),

$$\begin{aligned} &\sum_i \int_{W_i^n} h e^{S_n \hat{g}} J_{W_i^n} T^n \, (\psi \circ T^n - \overline{\psi}_i) \, dm_W \\ &\le C_e \sum_i \|h\|_s \frac{|W_i^n|^p}{|W|^p} \left| e^{S_n \hat{g}} J_{W_i^n} T^n \right|_{\mathcal{C}^\beta(W_i^n)} \Lambda^{-\beta n} \\ &\le |e^{S_n \hat{g}}|_\infty C_g (1 + C_d)^3 C_e C_g \Lambda^{-\beta n} \|h\|_s \sum_i \frac{|W_i^n|^p}{|W|^p} |J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)} \\ &\le C_4 C_g |e^{S_n \hat{g}}|_\infty \Lambda^{-\beta n} \|h\|_s, \end{aligned} \tag{A.9}$$

where $C_4 = (1 + C_d)^3 C_e C_2^{1-p}$ and in the second line we have used Lemma A.1(c) with $\varsigma = p$.

For the second term of (A.6), we use the fact that $|\overline{\psi}_i| \le |W|^{-p}$ since $|W|^p |\psi|_{\mathcal{C}^\beta(W)} \le 1$. Recall the notation introduced before the statement of Lemma A.1. Grouping the pieces $W_i^n \in \mathcal{G}_n(W)$ according to most recent long ancestors $W_j^k \in L_k(W)$, we have

$$\sum_i |W|^{-p} \int_{W_i^n} h e^{S_n \hat{g}} \cdot J_{W_i^n} T^n \, dm_W$$
$$= \sum_{k=1}^{n} \sum_{j \in L_k(W)} \sum_{i \in \mathcal{I}_n(W_j^k)} |W|^{-p} \int_{W_i^n} h e^{S_n \hat{g}} \cdot J_{W_i^n} T^n \, dm_W$$
$$+ \sum_{i \in \mathcal{I}_n(W)} |W|^{-p} \int_{W_i^n} h e^{S_n \hat{g}} J_{W_i^n} T^n \, dm_W$$

where we have split up the terms involving $k = 0$ and $k \geq 1$. We estimate the terms with $k \geq 1$ by the weak norm and the terms with $k = 0$ by the strong stable norm. Using again (A.3) and (A.4),

$$\sum_i |W|^{-p} \int_{W_i^n} h e^{S_n \hat{g}} \cdot J_{W_i^n} T^n \, dm_W$$
$$\leq |e^{S_n \hat{g}}|_\infty C_g (1 + C_d)^3 \sum_{k=1}^{n} \sum_{j \in L_k(W)} \sum_{i \in \mathcal{I}_n(W_j^k)} |W|^{-p} |h|_w |J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)}$$
$$+ |e^{S_n \hat{g}}|_\infty C_g (1 + C_d)^3 \sum_{i \in \mathcal{I}_n(W)} \frac{|W_i^n|^p}{|W|^p} \|h\|_s |J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)}.$$

In the first sum above corresponding to $k \geq 1$, we write

$$|J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)} \leq |J_{W_i^n} T^{n-k}|_{\mathcal{C}^0(W_i^n)} |J_{W_j^k} T^k|_{\mathcal{C}^0(W_j^k)}.$$

Thus using Lemma A.1(a) from time k to time n,

$$\sum_{k=1}^{n} \sum_{j \in L_k} \sum_{i \in \mathcal{I}_n(W_j^k)} |W|^{-p} |J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)}$$
$$\leq \sum_{k=1}^{n} \sum_{j \in L_k(W)} |J_{W_j^k} T^k|_{\mathcal{C}^0(W_j^k)} |W|^{-p} \sum_{i \in \mathcal{I}_n(W_j^k)} |J_{W_i^n} T^{n-k}|_{\mathcal{C}^0(W_i^n)}$$
$$\leq 3^p \delta_0^{-p} \sum_{k=1}^{n} \sum_{j \in L_k(W)} |J_{W_j^k} T^k|_{\mathcal{C}^0(W_j^k)} \frac{|W_j^k|^p}{|W|^p} C_1 \theta_*^{n-k},$$

since $|W_j^k| \geq \delta_0/3$. The inner sum is bounded by C_2^{1-p} for each k by Lemma A.1(c) while the outer sum is bounded by $C_1/(1 - \theta_*)$ independently of n.

Finally, for the sum corresponding to $k = 0$, since

$$|J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)} \leq (1 + C_d)|T^n W_i^n| |W_i^n|^{-1} \leq (1 + C_d)|J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)},$$

we use Jensen's inequality and Lemma A.1(a) to estimate,

$$\sum_{i \in \mathcal{I}_n(W)} \frac{|W_i^n|^p}{|W|^p} |J_{W_i^n} T^n|_{\mathcal{C}^0(W_i^n)} \leq (1 + C_d) \left(\sum_{i \in \mathcal{I}_n(W)} \frac{|T^n W_i^n|}{|W_i^n|} \right)^{1-p} \leq (1 + C_d) C_1 \theta_*^{n(1-p)}.$$

Gathering these estimates together, we have

$$\sum_i |W|^{-p} \left| \int_{W_i^n} he^{S_n \hat{g}} J_{W_i^n} T^n \, dm_W \right| \leq C_g |e^{S_n \hat{g}}|_\infty \Big(C_5 \delta_0^{-p} |h|_w + C_6 \|h\|_s \theta_*^{n(1-p)} \Big), \tag{A.10}$$

where $C_5 = 3^p(1+C_d)^3 C_1 C_2^{1-p}/(1-\theta_*)$ and $C_6 = (1+C_d)^4 C_1$. Putting together (A.9) and (A.10) proves (4.2),

$$\|\mathcal{L}_{T,g}^n h\|_s \leq C' C_g |e^{S_n \hat{g}}|_\infty \left(\Lambda^{-\beta n} + \theta_*^{n(1-p)} \right) \|h\|_s + C' C_g |e^{S_n \hat{g}}|_\infty \delta_0^{-p} |h|_w,$$

with $C' = \max\{C_4, C_5, C_6\}$, a uniform constant depending only on **(H1)**–**(H5)**.

A.4. Estimating the strong unstable norm. Fix $\varepsilon \leq \varepsilon_0$ and consider two curves $W^1, W^2 \in \mathcal{W}^s$ with $d_{\mathcal{W}^s}(W^1, W^2) \leq \varepsilon$. For $n \geq 1$, we describe how to partition $T^{-n}W^\ell$ into "matched" pieces U_j^ℓ and "unmatched" pieces V_k^ℓ, $\ell = 1, 2$. In the what follows, we use C_t to denote a transversality constant which depends only on the minimum angle between various transverse directions: the minimum angle between $C^s(x)$ and $C^u(x)$, between S_{-n}^T and $C^s(x)$, and between $C^s(x)$ and the vertical and horizontal directions.

Let ω be a connected component of $W^1 \backslash \mathcal{S}_{-n}^T$ such that $T^{-n}\omega \in \mathcal{G}_n(W)$. We define a smooth local foliation $\{\gamma_x\}_{x \in T^{-n}\omega}$ about $T^{-n}\omega$ such that for each $x \in T^{-n}\omega$: (1) γ_x is centered at x, (2) $\gamma_x \in \widehat{W}^u$; (3) $|\gamma_x| \leq 2BC_tC_e\Lambda^{-n}\varepsilon$ such that its image $T^n\gamma_x$, if not cut by a singularity or the boundary of a homogeneity strip, will have a projection on the vertical direction of length 2ε. By item (3) and the definition of $d_{\mathcal{W}^s}(W^1, W^2)$, it follows that any curve $T^n\gamma_x$ that is not cut by a singularity or the boundary of a homogeneity strip must necessarily intersect W^2, except possibly if $T^n\gamma_x$ lies near the endpoints of W^1. By **(H2)**, $T^i\gamma_x \in \widehat{\mathcal{W}}^u$ for each $i \geq 0$.

Doing this for each connected component of $W^1 \backslash \mathcal{S}_{-n}^T$, we subdivide $W^1 \backslash \mathcal{S}_{-n}^T$ into a countable collection of subintervals of points for which $T^n\gamma_x$ intersects $W^2 \backslash \mathcal{S}_{-n}^T$ and subintervals for which this is not the case. This in turn induces a corresponding partition on $W^2 \backslash \mathcal{S}_{-n}^T$.

We denote by V_k^ℓ the pieces in $T^{-n}W^\ell$ which are not matched up by this process and note that the images $T^nV_k^\ell$ occur either at the endpoints of W^ℓ or because the curve γ_x has been cut by a singularity or the boundary of a homogeneity strip. In both cases, the length of the curves $T^nV_k^\ell$ can be at most $C_t\varepsilon$ due to the uniform transversality of $\mathcal{S}_{-n}^T$ with $C^s(x)$, of $C^s(x)$ with $C^u(x)$ and of $C^s(x)$ with the horizontal.

In the remaining pieces the foliation $\{T^n\gamma_x\}_{x \in T^{-n}W^1}$ provides a one to one correspondence between points in W^1 and W^2. We partition these pieces in such a way that the lengths of their images under T^{-i} are less than δ_0 for each $0 \leq i \leq n$ and the pieces are pairwise matched by the foliation $\{\gamma_x\}$. We call these matched pieces $\widetilde{U}_j^\ell$ and note that $T^i\widetilde{U}_j^\ell \in \mathcal{G}_{n-i}(W^\ell)$ for each $i = 0, 1, \ldots n$. For convenience, we further trim the $\widetilde{U}_j^\ell$ to pieces U_j^ℓ so that U_j^1 and U_j^2 are both defined on the same arclength interval I_j. The at most two components of $T^n(\widetilde{U}_j^\ell \backslash U_j^\ell)$ have length less than $C_t\varepsilon$ due to the uniform transversality of $C^s(x)$ with the vertical direction. We attach these trimmed pieces to the adjacent U_i^ℓ or V_k^ℓ as appropriate so as not to create any additional components in the partition.

We further relabel any pieces U_j^ℓ as V_j^ℓ and consider them unmatched if for some i, $0 \le i \le n$, and for some $x \in U_j^\ell$, $|T^i\gamma_x| > 2|T^iU_j^\ell|$. i.e. we only consider pieces matched if at each intermediate step, the distance between them is at most of the same order as their length. We do this in order to be able to apply Lemma A.2 to the matched pieces. Notice that since the distance between the curves at each intermediate step is at most $C_tC_e\varepsilon$ and due to the uniform contraction of stable curves going forward, we have $|T^nV_k^\ell| \le C_tC_e^2\varepsilon$ for all such pieces considered unmatched by this last criterion.

In this way we write $W^\ell = (\cup_j T^nU_j^\ell) \cup (\cup_k T^nV_k^\ell)$. Note that the images $T^nV_k^\ell$ of the unmatched pieces must have length $\le C_v\varepsilon$ for some uniform constant C_v while the images of the matched pieces U_j^ℓ may be long or short.

Recalling the notation of Sect. 3, we have arranged a pairing of the pieces U_j^ℓ with the following property:

$$\text{If } U_j^1 = G_{U_j^1}(I_j) = \{(r, \varphi_{U_j^1}(r)) : r \in I_j\}, \text{ then } U_j^2 = G_{U_j^2}(I_j) = \{(r, \varphi_{U_j^2}(r)) : r \in I_j\}, \tag{A.11}$$

so that the point $x = (r, \varphi_{U_j^1}(r)) \in U_j^1$ can be associated with the point $\bar{x} = (r, \varphi_{U_j^2}(r)) \in U_j^2$ by the vertical line $\{(r, s)\}_{s\in[-\pi/2,\pi/2]}$, for each $r \in I_j$. In addition, the U_j^ℓ satisfy the assumptions of Lemma A.2.

Given ψ_ℓ on W^ℓ with $|\psi_\ell|_{\mathcal{C}^\alpha(W^\ell)} \le 1$ and $d_\beta(\psi_1, \psi_2) \le \varepsilon$, with the above construction we must estimate

$$\left|\int_{W^1} \mathcal{L}_{T,g}^n h\, \psi_1\, dm_W - \int_{W^2} \mathcal{L}_{T,g}^n h\, \psi_2\, dm_W\right| \le \sum_{\ell,k}\left|\int_{V_k^\ell} he^{S_n\hat{g}} J_{V_k^\ell}T^n \psi_\ell \circ T^n\, dm_W\right|$$
$$+\sum_j \left|\int_{U_j^1} he^{S_n\hat{g}} J_{U_j^1}T^n\psi_1 \circ T^n\, dm_W - \int_{U_j^2} he^{S_n\hat{g}} J_{U_j^2}T^n\psi_2\circ T^n\, dm_W\right|. \tag{A.12}$$

First we estimate the unmatched pieces V_k^ℓ using the strong stable norm. Note that by (A.5), $|\psi_\ell \circ T^n|_{\mathcal{C}^\beta(V_k^\ell)} \le C_e|\psi_\ell|_{\mathcal{C}^\alpha(W^\ell)} \le C_e$. We estimate as in Sect. A.3, using the fact that $|T^nV_k^\ell| \le C_v\varepsilon$, as noted above,

$$\begin{aligned}
&\sum_{\ell,k}\left|\int_{V_k^\ell} he^{S_n\hat{g}} J_{V_k^\ell}T^n\psi_\ell\circ T^n\, dm_W\right| \\
&\quad\le C_e\sum_{\ell,k}\|h\|_s|V_k^\ell|^p|e^{S_n\hat{g}}|_{\mathcal{C}^\beta(V^\ell,k)}|J_{V_k^\ell}T^n|_{\mathcal{C}^\beta(V^\ell,k)} \\
&\quad\le C_e(1+C_d)^3C_g|e^{S_n\hat{g}}|_\infty\|h\|_s\sum_{\ell,k}|V_k^\ell|^p|J_{V_k^\ell}T^n|_{\mathcal{C}^0(V_k^\ell)} \\
&\quad\le C'\varepsilon^pC_g|e^{S_n\hat{g}}|_\infty\|h\|_s\sum_{\ell,k}|J_{V_k^\ell}T^n|^{1-p}_{\mathcal{C}^0(V_k^\ell)} \le 2C'\varepsilon^pC_g|e^{S_n\hat{g}}|_\infty\|h\|_sC_3^n,
\end{aligned} \tag{A.13}$$

with $C' = C_e(1+C_d)^4C_v^p$, where we have applied Lemma A.1(d) with $\varsigma = 1-p > 1/2$ since there are at most two V_k^ℓ corresponding to each element $W_i^{\ell,n} \in \mathcal{G}_n(W^\ell)$ as defined in Sect. A.1 and $|J_{V_k^\ell}T^n|_{\mathcal{C}^0(V_k^\ell)} \le |J_{W_i^{\ell,n}}T^n|_{\mathcal{C}^0(W_i^{\ell,n})}$ whenever $V_k^\ell \subseteq W_i^{\ell,n}$.

Next, we must estimate

$$\sum_j \left| \int_{U_j^1} h e^{S_n \hat{g}} J_{U_j^1} T^n \, \psi_1 \circ T^n \, dm_W - \int_{U_j^2} h e^{S_n \hat{g}}_{U_j^2} T^n \, \psi_2 \circ T^n \, dm_W \right|.$$

We fix j and estimate the difference. Define

$$\phi_j = (e^{S_n \hat{g}} J_{U_j^1} T^n \, \psi_1 \circ T^n) \circ G_{U_j^1} \circ G_{U_j^2}^{-1}.$$

The function ϕ_j is well-defined on U_j^2 and we can write,

$$\begin{aligned} &\left| \int_{U_j^1} h e^{S_n \hat{g}} J_{U_j^1} T^n \, \psi_1 \circ T^n - \int_{U_j^2} h e^{S_n \hat{g}} J_{U_j^2} T^n \, \psi_2 \circ T^n \right| \\ &\quad \le \left| \int_{U_j^1} h e^{S_n \hat{g}} J_{U_j^1} T^n \, \psi_1 \circ T^n - \int_{U_j^2} h \, \phi_j \right| + \left| \int_{U_j^2} h (\phi_j - e^{S_n \hat{g}} J_{U_j^2} T^n \, \psi_2 \circ T^n) \right|. \end{aligned} \tag{A.14}$$

We estimate the first term on the right hand side of (A.14) using the strong unstable norm. Using (A.3), (A.4) and (A.5),

$$|e^{S_n \hat{g}} J_{U_j^1} T^n \cdot \psi_1 \circ T^n|_{\mathcal{C}^\alpha(U_j^1)} \le C_e (1 + C_d)^3 C_g |e^{S_n \hat{g}}|_\infty |J_{U_j^1} T^n|_{\mathcal{C}^0(U_j^1)}. \tag{A.15}$$

Notice that

$$|G_{U_j^1} \circ G_{U_j^2}^{-1}|_{\mathcal{C}^1(U_j^2)} \le \sup_{r \in U_j^2} \frac{\sqrt{1 + (d\varphi_{U_j^1}/dr)^2}}{\sqrt{1 + (d\varphi_{U_j^2}/dr)^2}} \le \sqrt{1 + \Gamma^2} =: C_a, \tag{A.16}$$

where Γ is the maximum slope of curves in $\mathcal{W}^s$ given by **(H1)**. Using this, we estimate as in (A.15),

$$|\phi_j|_{\mathcal{C}^\alpha(U_j^2)} \le C_a C_e (1 + C_d)^3 C_g |e^{S_n \hat{g}}|_\infty |J_{U_j^1} T^n|_{\mathcal{C}^0(U_j^1)}.$$

By the definition of ϕ_j and $d_\beta(\cdot, \cdot)$,

$$d_\beta(e^{S_n \hat{g}} J_{U_j^1} T^n \psi_1 \circ T^n, \phi_j) = \left| \left[e^{S_n \hat{g}} J_{U_j^1} T^n \psi_1 \circ T^n \right] \circ G_{U_j^1} - \phi_j \circ G_{U_j^2} \right|_{\mathcal{C}^\beta(I_j)} = 0.$$

By Lemma A.2(a), we have $d_{\mathcal{W}^s}(U_j^1, U_j^2) \le C_0 \Lambda^{-n} \varepsilon =: \varepsilon_1$. In view of (A.15) and following, we renormalize the test functions by $R_j = C_7 C_g |e^{S_n \hat{g}}|_\infty |J_{U_j^1} T^n|_{\mathcal{C}^0(U_j^1)}$ where $C_7 = C_a C_e (1 + C_d)^3$. Then we apply the definition of the strong unstable norm with ε_1 in place of ε. Thus,

$$\begin{aligned} &\sum_j \left| \int_{U_j^1} h e^{S_n \hat{g}} J_{U_j^1} T^n \, \psi_1 \circ T^n - \int_{U_j^2} h \, \phi_j \right| \\ &\quad \le C_7 C_0^\gamma \varepsilon^\gamma \Lambda^{-\gamma n} C_g |e^{S_n \hat{g}}|_\infty \|h\|_u \sum_j |J_{U_j^1} T^n|_{\mathcal{C}^0(U_j^1)} \end{aligned} \tag{A.17}$$

where the sum is $\leq C_2$ by Lemma A.1(b) since there is at most one matched piece U_j^1 corresponding to each element $W_i^{1,n} \in \mathcal{G}_n(W^1)$ and $|J_{U_j^1}T^n|_{\mathcal{C}^0(U_j^1)} \leq |J_{W_i^{1,n}}T^n|_{\mathcal{C}^0(W_i^{1,n})}$ whenever $U_j^1 \subseteq W_i^{1,n}$.

It remains to estimate the second term in (A.14) using the strong stable norm.

$$\left| \int_{U_j^2} h(\phi_j - e^{S_n \hat{g}} J_{U_j^2}T^n \psi_2 \circ T^n) \right| \leq \|h\|_s |U_j^2|^p \left| \phi_j - e^{S_n \hat{g}} J_{U_j^2}T^n \psi_2 \circ T^n \right|_{\mathcal{C}^\beta(U_j^2)}. \tag{A.18}$$

In order to estimate the $\mathcal{C}^\beta$-norm of the function in (A.18), we split it up into two differences. Since $|G_{U_j^\ell}|_{\mathcal{C}^1} \leq C_a$ and $|G_{U_j^\ell}^{-1}|_{\mathcal{C}^1} \leq 1$, $\ell = 1, 2$, we write

$$\begin{aligned}
&|\phi_j - (e^{S_n \hat{g}} J_{U_j^2}T^n) \cdot \psi_2 \circ T^n|_{\mathcal{C}^\beta(U_j^2)} \\
&\leq \left| \left[(e^{S_n \hat{g}} J_{U_j^1}T^n) \cdot \psi_1 \circ T^n \right] \circ G_{U_j^1} - \left[(e^{S_n \hat{g}} J_{U_j^2}T^n) \cdot \psi_2 \circ T^n \right] \circ G_{U_j^2} \right|_{\mathcal{C}^\beta(I_j)} \\
&\leq \left| (e^{S_n \hat{g}} J_{U_j^1}T^n) \circ G_{U_j^1} \left[\psi_1 \circ T^n \circ G_{U_j^1} - \psi_2 \circ T^n \circ G_{U_j^2} \right] \right|_{\mathcal{C}^\beta(I_j)} \\
&\quad + \left| \left[(e^{S_n \hat{g}} J_{U_j^1}T^n) \circ G_{U_j^1} - (e^{S_n \hat{g}} J_{U_j^2}T^n) \circ G_{U_j^2} \right] \psi_2 \circ T^n \circ G_{U_j^2} \right|_{\mathcal{C}^\beta(I_j)} \\
&\leq C_a(1 + C_d)^3 C_g |e^{S_n \hat{g}} J_{U_j^1}T^n|_{\mathcal{C}^0(U_j^1)} \left| \psi_1 \circ T^n \circ G_{U_j^1} - \psi_2 \circ T^n \circ G_{U_j^2} \right|_{\mathcal{C}^\beta(I_j)} \\
&\quad + C_a C_e \left| (e^{S_n \hat{g}} J_{U_j^1}T^n) \circ G_{U_j^1} - (e^{S_n \hat{g}} J_{U_j^2}T^n) \circ G_{U_j^2} \right|_{\mathcal{C}^\beta(I_j)}.
\end{aligned} \tag{A.19}$$

To bound the two differences above, we need the following lemma, which was proved in [DZ2] Lemma 4.2. The only difference is the factor $e^{S_n \hat{g}}$ which does not play any significant role in the proof, so we omit the proof here.

Lemma A.3. *There exist constants $C_8, C_9 > 0$, depending only on* **(H1)–(H5)**, *such that,*

(a) $|(e^{S_n \hat{g}} J_{U_j^1}T^n) \circ G_{U_j^1} - e^{S_n \hat{g}} J_{U_j^2}T^n) \circ G_{U_j^2}|_{\mathcal{C}^\beta(I_j)} \leq C_8 C_g |e^{S_n \hat{g}} J_{U_j^2}T^n|_{\mathcal{C}^0(U_j^2)} \varepsilon^{1/3-\beta}$;

(b) $|\psi_1 \circ T^n \circ G_{U_j^1} - \psi_2 \circ T^n \circ G_{U_j^2}|_{\mathcal{C}^q(I_{r_j})} \leq C_9 C_g \varepsilon^{\alpha-\beta}$.

It follows from Lemma A.3(a) that

$$|e^{S_n \hat{g}} J_{U_j^1}T^n|_{\mathcal{C}^0(U_j^1)} \leq (1 + C_8 C_g \varepsilon^{1/3-\beta}) |e^{S_n \hat{g}} J_{U_j^2}T^n|_{\mathcal{C}^0(U_j^2)}$$

which we will use to simplify (A.19). Starting from (A.18), we apply Lemma A.3 to (A.19) to obtain,

$$\begin{aligned}
&\sum_j \left| \int_{U_j^2} h(\phi_j - e^{S_n \hat{g}} J_{U_j^2}T^n \psi_2 \circ T^n) \, dm_W \right| \\
&\leq \bar{C} C_g \|h\|_s \sum_j |U_j^2|^p |e^{S_n \hat{g}} J_{U_j^2}T^n|_{\mathcal{C}^0(U_j^2)} \, \varepsilon^{\alpha-\beta} \\
&\leq \bar{C} C_g |e^{S_n \hat{g}}|_\infty \|h\|_s \varepsilon^{\alpha-\beta} \sum_j |J_{U_j^2}T^n|_{\mathcal{C}^0(U_j^2)},
\end{aligned} \tag{A.20}$$

for some uniform constant $\bar{C}$ where again the sum is finite as in (A.17). This completes the estimate on the second term in (A.14). Now we use this bound, together with (A.13) and (A.17) to estimate (A.12)

$$\begin{aligned}\left|\int_{W^1} \mathcal{L}^n_{T,g} h\, \psi_1\, dm_W - \int_{W^2} \mathcal{L}^n_{T,g} h\, \psi_2\, dm_W\right| &\leq C C_3^n C_g |e^{S_n \hat{g}}|_\infty \|h\|_s \varepsilon^p \\ &\quad + C\|h\|_u \Lambda^{-\gamma n} C_g |e^{S_n \hat{g}}|_\infty \varepsilon^\gamma \\ &\quad + C C_g |e^{S_n \hat{g}}|_\infty \|h\|_s \varepsilon^{\alpha-\beta},\end{aligned} \tag{A.21}$$

where again C depends only on **(H1)–(H5)** through the estimates above. Since $\alpha - \beta \geq \gamma$ and $p \geq \gamma$, we divide through by ε^γ and take the appropriate suprema to complete the proof of (4.3).

References

[A] Andrey, L.: The rate of entropy change in non-Hamiltonian systems. Phys. Lett. A **11**, 45–46 (1985).

[BDL] Baladi, V., Demers, M., Liverani, C.: Exponential decay of correlations for finite horizon Sinai billiard flows. Inventiones Mathematicae **211**, 39–177 (2018)

[BT] Bálint, P., Tóth, I.P.: Correlation decay in certain soft billiards. Commun. Math. Phys. **243**, 55–91 (2003)

[BS] Bunimovich, L., Sinai, Y.G.: Statistical properties of Lorentz gas with periodic configuration of scatterers. Commun. Math. Phys. **78**, 479–497 (1980/1981)

[BSC] Bunimovich, L., Sinai, Y.G., Chernov, N.: Statistical properties of two-dimensional hyperbolic billiards. Russ. Math. Surv. **46**, 47–106 (1991)

[Ch1] Chernov, N.: Decay of correlations in dispersing billiards. J. Stat. Phys. **94**, 513–556 (1999)

[Ch2] Chernov, N.: Sinai billiards under small external forces. Ann. Henri Poincaré **2**, 197–236 (2001)

[Ch4] Chernov, N.: Sinai billiards under small external forces II. Ann. Henri Poincaré **9**, 91–107 (2008)

[CELS1] Chernov, N., Eyink, G., Lebowitz, J., Sinai, Y.G.: Derivation of Ohm's law in a deterministic mechanical model. Phys. Rev. Lett. **70**, 2209–2212 (1993)

[CELS2] Chernov, N., Eyink, G., Lebowitz, J., Sinai, Y.G.: Steady-state electrical conduction in the periodic Lorentz gas. Commun. Math. Phys. **154**, 569–601 (1993)

[CM] Chernov, N., Markarian, R.: Chaotic Billiards. Mathematical Surveys and Monographs, vol. 127. AMS, Providence (2006)

[CZ] Chernov, N., Zhang, H.-K.: Statistical properties of hyperbolic systems with general singularities. J. Stat. Phys **136**, 615–642 (2009)

[CZZ] Chernov, N., Zhang, H.-K., Zhang, P.: Electrical current for Sinai billiards under general small forces. J. Stat. Phys. **153**, 1065–1083 (2013)

[ChGa] Chetrite, R., Gawedzki, K.: Fluctuation relations for diffusion processes. Commun. Math. Phys. **282**, 469–518 (2008)

[CG] Cohen, E.G.D., Gallavotti, G.: Note on two theorems in nonequilibrium statistical mechanics. J. Stat. Phys. **96**, 1343–1349 (1999)

[DZe] Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications (Applications of Mathematics, 38), 2nd edn. Springer, New York (1998)

[D] Demers, M.F.: Functional norms for Young towers. Ergod. Theory Dyn. Syst. **30**, 1371–1398 (2010)

[DZ1] Demers, M.F., Zhang, H.-K.: Spectral analysis of the transfer operator for the Lorentz gas. J. Mod. Dyn. **5**, 665–709 (2011)

[DZ2] Demers, M.F., Zhang, H.-K.: A functional analytic approach to perturbations of the Lorentz gas. Commun. Math. Phys. **324**(3), 767–830 (2013)

[DZ3] Demers, M.F., Zhang, H.-K.: Spectral analysis of hyperbolic systems with singularities. Nonlinearity **27**, 379–433 (2014)

[ECM] Evans, D.J., Cohen, E.G.D., Morriss, G.P.: Probability of second law violations in shearing steady states. Phys. Rev. Lett. **71**, 2401–2404 and 3616 (1993)

[EM] Evans, D.J., Morriss, G.P.: Statistical Mechanics of Non Equilibrium Liquids. Academic Press, London (1990)

[ES1] Evans, D.J., Searles, D.J.: Equilibrium microstates which generate the second law violating steady states. Phys. Rev. E **50**, 1645–1648 (1994)
[ES2] Evans, D.J., Searles, D.J.: The fluctuation theorem. Adv. Phys. **51**, 1529–1585 (2002)
[GC1] Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in non-equilibrium statistical mechanics. Phys. Rev. Lett. **74**, 2694–2697 (1995)
[GC2] Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in stationary states. J. Stat. Phys. **80**, 931970 (1995)
[G] Gordin, M.I.: The central limit theorem for stationary processes. Sov. Math. Dokl. **10**, 1174–1176 (1969)
[Ja] Jarzynski, C.: A nonequilibrium equality for free energy differences. Phys. Rev. Lett. **78**, 2690–2693 (1997)
[JPR] Jaksic, V., Pillet, C.-A., Rey-Bellet, L.: Entropic fluctuations in statistical mechanics I. Classical dynamical systems. Nonlinearity **24**, 699–763 (2011)
[Ka] Kato, T.: Perturbation Theory for Linear Operators (Grundlehren der mathematischen Wissenchaften, 132), 2nd edn. Springer, Berlin (1984)
[KL] Keller, G., Liverani, C.: Stability of the spectrum for transfer operators. Annali Della Scuola Normale Superiore di Pisa, Scienze Fisiche E Matematiche, (4) **28**, 141–152 (1999)
[Ku] Kurchan, J.: Fluctuation theorem for stochastic dynamics. J. Phys. A Math. Gen. **31**, 3719–3729 (1998)
[LS] Lebowitz, J., Spohn, H.: A Gallavotti–Cohen type symmetry in the large deviation functional for stochastic dynamics. J. Stat. Phys. **95**, 333–365 (1995)
[M1] Maes, C.: The fluctuation theorem as a Gibbs property. J. Stat. Phys. **95**, 367–392 (1999)
[M2] Maes, C.: On the origin and the use of fluctuation relations for the entropy. Séminaire Poincaré **2**, 29–62 (2003)
[MN] Maes, C., Netockny, K.: Time reversal and entropy. J. Stat. Phys. **110**, 269–310 (2003)
[MN1] Melbourne, I., Nicol, M.: Almost sure invariance principle for nonuniformly hyperbolic systems. Commun. Math. Phys. **260**, 393–401 (2005)
[MN2] Melbourne, I., Nicol, M.: Large deviations for nonuniformly hyperbolic systems. Trans. Am. Math. Soc. **360**, 6661–6676 (2008)
[RY] Rey-Bellet, L., Young, L.-S.: Large deviations in nonuniformly hyperbolic dynamical systems. Ergod. Theory Dyn. Syst. **28**, 587–612 (2008)
[R1] Ruelle, D.: Positivity of entropy production in nonequilibrium statistical mechanics. J. Stat. Phys. **85**, 1–23 (1996)
[R2] Ruelle, D.: Entropy production in nonequilibrium statistical mechanics. Commun. Math. Phys. **189**, 365–371 (1997)
[R3] Ruelle, D.: Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics. J. Stat. Phys. **94**, 393–468 (1999)
[S] Sinai, Ya.: Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. Russ. Math. Surv. **25**, 137–189 (1970)
[V] Viana, M.: Stochastic Dynamics of Deterministic Systems. Brazilian Mathematics Colloquium, IMPA (1997)
[Y] Young, L.-S.: Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. **147**, 585–650 (1998)
[Z] Zhang, H.-K.: Current in periodic Lorentz gases with twists. Commun. Math. Phys. **306**, 747–776 (2011)

Communicated by C. Liverani