# THE DIFFUSION COEFFICIENT FOR PIECEWISE EXPANDING MAPS OF THE INTERVAL WITH METASTABLE STATES

DMITRY DOLGOPYAT AND PAUL WRIGHT

ABSTRACT. Consider a piecewise smooth expanding map of the interval possessing several invariant subintervals and the same number of ergodic absolutely continuous invariant probability measures (ACIMs). After this system is perturbed to make the subintervals lose their invariance in such a way that there is a unique ACIM, we show how to approximate the diffusion coefficient for an observable of bounded variation by the diffusion coefficient of a related continuous time Markov chain.

## 1. INTRODUCTION

Metastable dynamics arise in a number of physical systems. In such systems, the phase space can be divided into a finite number of components, called metastable states, that are nearly invariant under the dynamics. A typical trajectory will remain in one metastable state for an extended period of time before escaping to another metastable state and repeating this behavior. Rigourous results about metastability in dynamical systems perturbed by noise can be found in [6, 9].

In this paper we are concerned with reducing the description of purely deterministic chaotic systems to cooresponding finite state Markov chains. In particular, we continue the study of the dynamics of hyperbolic interval maps with metastable states initiated in [7]. These systems arise from perturbing an initial system $T_0$ with $m$ disjoint invariant intervals $I_1, I_2, \ldots, I_m$. The initial map has $m$ mutually singular ergodic absolutely continuous invariant measures (ACIMs), $\mu_1, \mu_2, \ldots, \mu_m$. $T_0$ is perturbed in such a way that the $\mu_j$ lose their invariance, and the perturbed map $T_\varepsilon$ has only one ACIM, $\mu_\varepsilon$. See Figure 1.

Such metastable systems can be understood in the context of deterministic dynamical systems with holes (see [4]) as follows. As the invariance of initially invariant intervals is destroyed by the perturbation, we think of the small set of points $I_i \cap T_\varepsilon^{-1} I_j$ that switch from $I_j$ to $I_i$ after the application of $T_\varepsilon$, as being holes in the initially invariant sets. Therefore the techniques developed to study systems with holes are useful in our analysis.

As was shown in [7], we are able to approximate $\mu_\varepsilon$, for small $\varepsilon$, by a convex combination $\sum_{j=1}^m p_j \mu_j$ of the initially invariant measures, where $(p_1, p_2, \ldots, p_m)$ is the invariant measure for the continuous time Markov chain on $m$ states with transition rates proportional to the asymptotic sizes of the holes. Intuitively, since our map is chaotic on each interval, we expect the transition times between the $I_j$'s to be almost independent, which explains the appearence of the above mentioned Markov chain. In this paper, we show that the Markov approximation also extends to the diffusion matrix for smooth observables. We believe the methods developed in our paper can be used to describe the transport coefficients in other chaotic systems, for example, billiards with narrow tunnels [12]. However, in order to present the ideas of our proof in the simplest possible setting, we restrict our attention here to the setup of [7].
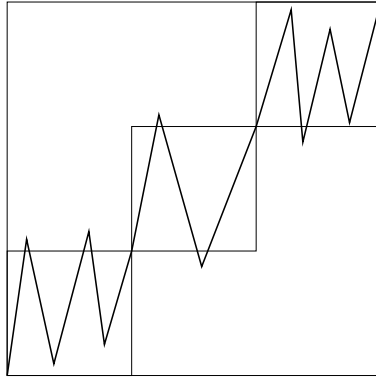
1

FIGURE 1. A map with three almost invariant intervals

## 2. STATEMENT OF THE MAIN RESULTS

In this section, we define a family of dynamical systems with $m$ nearly invariant (metastable) subsets. They are perturbations of a one-dimensional piecewise smooth expanding map with $m$ invariant subintervals $I_1, I_2, \ldots, I_m$ of positive Lebesgue measure. On each of these intervals, the unperturbed system has a unique ACIM. The perturbations break this invariance in such a way that each perturbed system will have only one ACIM. Our main result is an asymptotic formula for the diffusion coefficent of a smooth observable as the size of the perturbation tends to zero. We also show that the sequence of jump times in between different intervals asymptotically approach exponential random variables.

Let $I = [0, 1]$. In this paper, a map $T : I \circlearrowleft$ is called a piecewise $C^2$ map with $\mathcal{C} = \{0 = c_0 < c_1 < \cdots < c_d = 1\}$ as a critical set if for each $i$, $T|_{(c_i, c_{i+1})}$ extends to a $C^2$ function on a neighborhood of $[c_i, c_{i+1}]$. We call $T$ uniformly expanding if its minimum expansion, $\inf_{x \in I \setminus \mathcal{C}_0} |T_0'(x)|$, is greater than 1. As is customary for piecewise smooth maps, we consider $T$ to be bi-valued at points $c_i \in \mathcal{C}$ where it is discontinuous. In such cases we let $T(c_i)$ be both values obtained as $x$ approaches $c_i$ from either side.

2.1. **The initial system and its perturbations.** The unperturbed system is a piecewise $C^2$ uniformly expanding map $T_0 : I \circlearrowleft$ with $\mathcal{C}_0 = \{0 = c_{0,0} < c_{1,0} < \cdots < c_{d,0} = 1\}$ as a critical set. There are boundary points $\mathcal{B} = \{b_j\} \subset (0, 1)$ such that $I_j = [b_{j-1}, b_j]$ ($b_0 = 0, b_m = 1$) i.e. $T_0(I_j) \subset I_j$. The existence of an ACIM of bounded variation for $T_0|_{I_j}$ is guaranteed by [10]. We assume in addition:
**(I)** *Unique mixing ACIMs on the initially invariant subsets:* $T_0|_{I_j}$, $j \in \{1 \ldots m\}$, has only one ACIM $\mu_j$, whose density is denoted by $\phi_j = d\mu_j/dx$. $(T_0, \mu_j)$ is mixing.

From **(I)**, it follows that all ACIMs of $T_0$ are convex combinations of the ergodic ones, $\{\mu_j\}$.

We define the points in $H_0 = (T_0^{-1}\mathcal{B}) \setminus \mathcal{B}$ to be *infinitesimal holes*[1].

**(II)** *No return of the critical set to the infinitesimal holes:* For every $k > 0$, $(T_0^k \mathcal{C}_0) \cap H_0 = \emptyset$.

Since $\phi_j$ are of bounded variation, they can be suitably defined so that they are continuous except on at most a countable set of points where it has jump discontinuities. Moreover (see [7, Section 4.2]), **(II)** implies that after $\phi_j$ have been so defined, they are continuous at each of the infinitesimal holes.

**(III)** *Positive ACIMs at infinitesimal holes:* $\phi_j$ is positive at each of the points in $H_0 \cap I_j$.

**(II)** and **(III)** are generic conditions, although **(II)** may be difficult to verify for specific examples. Similar assumptions are made elsewhere, including in [9, Section 3.2] and [7].

**(IV)** *Restriction on periodic critical points:* Either

    (a) $\inf_{x \in I \setminus \mathcal{C}_0} |T_0'(x)| > 2$, or

    (b) $T_0$ has no periodic critical points, except possibly that 0 or 1 may be fixed points.

Because $T_0$ may be bi-valued at points in $\mathcal{C}_0$, a critical point $c_{i,0}$ is considered periodic if there exists $n > 0$ such that $c_{i,0} \in T_0^n\{c_{i,0}\}$. Condition **(IV)** is necessary in order to ensure that the transfer operators cooresponding to the perturbed systems defined below satisfy uniform Lasota-Yorke inequalities. These uniform inequalities are essential for establishing the perturbative spectral results, Propositions 2 and 3, which are a key ingredient of our proof. Since we cannot exclude the possibility of the forward orbit of a critical point containing other critical points, these uniform inequalities do not follow directly from the original paper [10], but rather from later works [2, 1], see [7, Section 4.2].

In what follows, we consider $C^2$-small perturbations $T_\varepsilon : I \circlearrowleft$ of $T_0$ for $\varepsilon > 0$. A critical set for $T_\varepsilon$ may be chosen as $\mathcal{C}_\varepsilon = \{0 = c_{0,\varepsilon} < c_{1,\varepsilon} < \cdots < c_{d,\varepsilon} = 1\}$, where for each $i$, $\varepsilon \mapsto c_{i,\varepsilon}$ is a $C^2$ function for $\varepsilon \geq 0$. Furthermore, there exists $\delta > 0$ such that for all sufficiently small $\varepsilon$, there exists a $C^2$ extension $\hat{T}_{i,\varepsilon} : [c_{i,0} - \delta, c_{i+1,0} + \delta] \to \mathbb{R}$ of $T_\varepsilon|_{[c_{i,\varepsilon}, c_{i+1,\varepsilon}]}$, and $\hat{T}_{i,\varepsilon} \to \hat{T}_{i,0}$ in the $C^2$ topology.

We also assume:

**(V)** *The boundary points do not move, and no holes are created near them:* Precisely, for each $b \in \mathcal{B}$ we have the following

    (a) If $b \notin \mathcal{C}_0$, then necessarily $T_0(b) = b$. We assume further that for all $\varepsilon > 0$, $T_\varepsilon(b) = b$.

    (b) If $b \in \mathcal{C}_0$, we assume that $T_0(b_-) < b < T_0(b_+)$, and also that $b \in \mathcal{C}_\varepsilon$ for all $\varepsilon$.

This boundary condition can be considerably relaxed by suitably redefining the holes discussed below, as was explained in [7, Section 2.4]. For simplicity of presentation, we do not detail these generalizations here.

Set $H_{ij,\varepsilon} = I_i \cap T_\varepsilon^{-1}(I_j)$. We refer to these sets as *holes*. Once a $T_\varepsilon$-orbit enters a hole, it leaves one of the invariant sets for $T_0$ and continues in another. As $\varepsilon \to 0$, the holes converge (in the Hausdorff metric) to the infinitesimal holes from which they arise. Our assumptions imply that there exist numbers $\beta_{ij} \geq 0$ such that

$$(1) \qquad \qquad \mu_i(H_{ij,\varepsilon}) = \varepsilon\beta_{ij} + o(\varepsilon).$$

Conisider a continuous time Markov chain with states $1, 2, \ldots, m$ and jump rates from state $i$ to state $j$ equal to $\beta_{ij}$.

---

[1]The exclusion of boundary points from the set of infinitesimal holes is not essential, although it does simplify our presentation. See assumption **(V)** and the discussion thereafter.

**(VI)** *Irreducibility:* The Markov chain defined above is irreducible.

Condition **(VI)** implies that for small $\varepsilon > 0$, $T_\varepsilon$ has only one ACIM $\mu_\varepsilon$, with density $\phi_\varepsilon = d\mu_\varepsilon/dx$.

Examples of families $T_\varepsilon$ that satisfy **(I)** through **(VI)** can be found in [7, Section 2.4].

It is natural to inquire if the family of ACIMs $\mu_\varepsilon$ has a unique limit as $\varepsilon \to 0$, and if this limit exists, how to express it as a convex combination of $\mu_1, \ldots, \mu_m$. This problem was successfuly addressed in [7].

**Proposition 1** (Theorem 1 in [7]). *As $\varepsilon \to 0$,*

$$\phi_\varepsilon \xrightarrow{L^1} \phi_0 = \sum_j p_j \phi_j$$

*where $(p_1, p_2, \ldots, p_m)$ is the invariant measure for the Markov chain.*

We define $\mu_0$ to be the measure whose density is $\phi_0$.

2.2. **The main results.** Proposition 1 shows that the limiting Markov chain provides useful information about our metastable system. In this section we prove two additional results making the connection between the dynamical system and the Markov chain more precise.

To state our first result we need to define, both for the $T_\varepsilon$-dynamics and for the Markov chain, a sequence of times that indicate when a transition occurs between different $I_j$'s.

**For the Markov chain:** Set $t_0^M = 0$, and for $i > 0$, let $t_i^M$ be the $i^{\text{th}}$ time that the Markov dynamics have changed states. Then for $i \geq 1$, set $\mathcal{T}_i^M = t_i^M - t_{i-1}^M$. Let $z_i^M$ be the state of the chain after the $i^{\text{th}}$ transition. Let $\mathbb{P}^r$ denote the probability measure constructed on the space $\{1, \ldots, m\}^{[0,\infty)}$ for the Markov chain by starting at $I_r$ at $t = 0$ and then evolving forward by the Markov dynamics. Then, given the present state of the chain, $\mathcal{T}_i^M$ are exponential random variables. That is, for $t \geq 0$, the densities are given by

$$d\mathbb{P}^r(\mathcal{T}_i^M = t | z_{i-1}^M = j) = \beta_j e^{-\beta_j t} dt$$

where $\beta_j = \sum_k \beta_{jk}$. Also given that $z_{i-1} = j$, $z_i$ is independent of $\mathcal{T}_i^M$ and

$$\mathbb{P}^r(z_i^M = k | z_{i-1}^M = j) = \frac{\beta_{jk}}{\beta_j}.$$

**For $T_\varepsilon$:** Let $z(x) = k$ if $x \in I_k$. Set $t_0^\varepsilon = 0$, and for $i > 0$, let $t_i^\varepsilon = \inf\{n > t_{i-1}^\varepsilon : z(T_\varepsilon^n x) \neq z(T_\varepsilon^{t_{i-1}^\varepsilon} x)\}$. Then for $i \geq 1$, set $\mathcal{T}_i^\varepsilon = t_i^\varepsilon - t_{i-1}^\varepsilon$.

Our first main result states that the finite dimensional distributions of jumps of the deterministic systems converge to the finite dimensional distributions of jumps of the Markov chain.

**Theorem 1.** *Fix $j, p$ and $S$. For any intervals $\Delta_k = [a_k, b_k]$, and numbers $r_k \in \{1, \ldots, m\}$, $k = 1, \ldots, p$*

$$\mu_j(\varepsilon \mathcal{T}_k^\varepsilon \in \Delta_k, z(t_k^\varepsilon) = r_k, \text{ for } k = 1, \ldots, p) \to \mathbb{P}^r(\mathcal{T}_k \in \Delta_k, z_k^M = r_k, \text{ for } k = 1, \ldots, p)$$

*and the convergence is uniform for $\max b_k \leq S$.*

Consider an observable $A : I \to \mathbb{R}$. For each fixed small $\varepsilon > 0$, the Ergodic Theorem provides a law of large numbers, i.e. $N^{-1} \sum_{k=0}^N A \circ T_\varepsilon^k \to \mu_\varepsilon(A)$ a.e. and in $L^1$ as $N \to \infty$. If $A$ is of bounded variation, the Central Limit Theorem applies [11] and states that $N^{-1/2} \sum_{k=0}^N A \circ T_\varepsilon^k$ approaches a normal distribution as $N \to \infty$. We let $D_\varepsilon(A)$ be the diffusion coeffient, which is the variance of the limiting normal distribution.

**Theorem 2.** For any observable $A$ of bounded variation,
$$\varepsilon D_\varepsilon(A) \to \mathbf{D}^M(\mathbf{A})$$
as $\varepsilon \to 0$, where $\mathbf{A}$ is the observable on the state space of our Markov chain such that $\mathbf{A}(j) = \int_{I_j} A\phi_j dx$ and $\mathbf{D}^M$ stands for the diffusion coefficient.

**Remark.** $\mathbf{D}^M$ can be computed efficiently. Assume for simplicity that $\mathbf{A}$ has zero mean (otherwise we subtract a constant from $\mathbf{A}$). Let $G$ denote the generator matrix of our Markov process, that is $G_{jk} = \beta_{jk}$ if $j \neq k$ and $G_{jj} = -\beta_j$. Then
$$\mathbf{D}^M = \sum_{jk} p_j \mathbf{A}(j) \int_0^\infty p_{jk}(t)\mathbf{A}(k)dt = \langle p\mathbf{A}, \int_0^\infty e^{tG}\mathbf{A}dt \rangle = \langle p\mathbf{A}, G^{-1}\mathbf{A} \rangle.$$

The proof of Theorems 1 and 2 depend on two perturbation results, Propositions 2 and 3. Proposition 3 is a variation on a result of Keller and Liverani [8, 9], see Appendix A for details, while Proposition 2 appears to be new. Our paper is organized as follows. In Section 3 we recall the transfer operator approach to the study of piecewise expanding interval maps and state Propositions 2 and 3. In Section 4 we derive Theorem 1 from Proposition 3. Section 5 contains the derivation of Theorem 2 from Theorem 1 and Proposition 2. Finally in Section 6 we explain how Proposition 2 follows from Theorem 1.

## 3. Preparatory material

3.1. **Function spaces and norms.** We use Leb to denote normalized Lebesgue measure on $I$ and $L^1$ to denote the space of Lebesgue integrable functions on $I$. For $f : I \to \mathbb{R}$, let $|f|_{L^1} = \int_I |f(x)|\, dx$, $|f|_{L^\infty} = \sup_{x \in I} |f(x)|$, and $\mathrm{var}(f)$ be the total variation of $f$ over $I$; that is,
$$\mathrm{var}(f) = \sup\{\sum_{i=1}^n |f(x_i) - f(x_{i-1})| : n \geq 1, 0 \leq x_0 < x_1 < \cdots < x_n \leq 1\}.$$
$\mathrm{BV} = \mathrm{BV}(I)$ is the Banach space of functions $f : I \to \mathbb{R}$ with norm
$$\|f\| = \inf_{g=f \text{ except on a countable set}} \mathrm{var}(g) + |g|_{L^\infty}.$$
An element of BV is technically an equivalence class of functions, any two of which agree except on a countable set; we generally ignore this distinction.

3.2. **Transfer operators and the spectral setting.** For $\varepsilon \geq 0$, let $\mathcal{L}_\varepsilon$ be the transfer operator associated with $T_\varepsilon$ acting on BV, i.e.
$$\mathcal{L}_\varepsilon A(x) = \sum_{y \in T_\varepsilon^{-1}\{x\}} \frac{A(y)}{|T_\varepsilon'(y)|}.$$
Note that $\mathcal{L}_0$ has 1 as an isolated eigenvalue of multiplicity $m$, and then a spectral gap. Let $P$ denote the associate spectral projection, that is
$$(PA)(x) = \left(\int_{I_j} A(y)dy\right)\phi_j(x) \text{ if } x \in I_j.$$
For small $\varepsilon > 0$, 1 becomes an isolated eigenvalue creating a small spectral gap. Note that $\mathcal{L}_\varepsilon$ preserves the space $\mathrm{BV}_0$ of BV-functions with zero mean.

**Proposition 2.** There exists $\eta < 1$ and $\kappa > 0$ such that

$$||\mathcal{L}_\varepsilon^{\kappa n/\varepsilon}||_{\mathrm{BV}_0} \leq \eta^n.$$

The proof of Proposition 2 is given is Section 6.

**Remark.** In case of two intervals much more precise asymptotics of the spectral gap can be deduced from the results of [9]. It is likely that a similar statement holds for an arbitrary number of intervals, however we do not pursue this question here since the weaker version stated above is sufficient for our purposes. In fact we hope that the argument used to prove Proposition 2 can be extended also to the case of piecewise hyperbolic systems such as billiards consisting of several regions connected by small holes.

For $j \in \{1, \ldots, m\}$, let $\mathcal{L}_{j,\varepsilon}$ be the transfer operator acting on $\mathrm{BV}(I_j)$ where we consider $H_{j,\varepsilon}$ as a hole, i.e.

$$\mathcal{L}_{j,\varepsilon}A(x) = \sum_{y \in (I_j \cap T_\varepsilon^{-1}\{x\})} \frac{A(y)}{|T'_\varepsilon(y)|}.$$

From [8] we see that, for small $\varepsilon > 0$, $\mathcal{L}_{j,\varepsilon}$ has an isolated simple eigenvalue $\lambda_{j,\varepsilon} < 1$ of multiplicity one with $\lambda_{j,\varepsilon} \to 1$ as $\varepsilon \to 0$, and otherwise the spectrum has a uniform spectral gap. More precisely, if $\mathcal{Q}_{j,\varepsilon}$ is the spectral projection corresponding to $\lambda_{j,\varepsilon}$, the following statement holds.

**Proposition 3.** There exists a measure $\nu_{j,\varepsilon}$ and a function $\phi_{j,\varepsilon} \in \mathrm{BV}(I_j)$ such that

$$\nu_{j,\varepsilon}(\phi_{j,\varepsilon}) = 1 \quad , \mathcal{Q}_{j,\varepsilon} = \nu_{j,\varepsilon}(\cdot)\phi_{j,\varepsilon}$$

and

(a) $\lambda_{j,\varepsilon} = 1 - \beta_j\varepsilon + o(\varepsilon)$.
(b) As $\varepsilon \to 0$, $\nu_{j,\varepsilon}(A) \to \int A(y)dy$ where the convergence holds in strong topology in $\mathrm{BV}^*$ and $\phi_{j,\varepsilon} \to \phi_j$ in $L^1$. Moreover

(2)
$$\limsup_{\varepsilon \to 0} \sup_{H_{j,\varepsilon}} |\phi_{j,\varepsilon} - \phi_j| = 0.$$

(c) There exists $C > 0$ such that, for all small $\varepsilon$, $A \in \mathrm{BV}(I_j)$, and all $n \geq 0$,

$$\left\|\lambda_{j,\varepsilon}^{-n}\mathcal{L}_{j,\varepsilon}^n A - \mathcal{Q}_{j,\varepsilon}A\right\| \leq C\theta^n \|A\|.$$

This proposition follows from the work of Keller and Liverani [8, 9], see Appendix A for details.

## 4. Convergence of the jump process

4.1. **Tightness.** We need to know that the distribution of jumps is tight so we begin with the following result.

**Lemma 1.** Given $S, \delta, p$ there exists $\sigma$ such that, for all $\varepsilon$ sufficiently small,

$$\mu_p(\exists k \text{ with } t_k^\varepsilon \leq S/\varepsilon \text{ and } \mathcal{T}_{k+1}^\varepsilon \leq \sigma/\varepsilon) \leq \delta.$$

The following estimate plays a key role in our analysis. Given a segment $J \subset I$ let $r_n(x)$ be the distance of $T_\varepsilon^n x$ to the boundary of the component of $T_\varepsilon^n J$ containing it. Recall that Leb denotes Lebesgue measure on $I = [0, 1]$.

**Lemma 2** (Growth Lemma)**.** (see [3, Section 5.10]) There exists $\Lambda > 1$, $c > 0$ such that, for all $\varepsilon$ small enough, and all $J$, and all $n \geq 0$,

$$\mathrm{Leb}(x : r_n(x) \leq \varepsilon) \leq \mathrm{Leb}(x : r_0(x) \leq \varepsilon/\Lambda^n) + c\,\mathrm{Leb}(J)\varepsilon.$$

*Proof of Lemma 1.* First, $t_0^\varepsilon = 0$, and if $N = N(\sigma, \varepsilon) = \lceil \sigma/\varepsilon \rceil$, then

$$\mu_p(\mathcal{T}_1^\varepsilon \leq \kappa/\varepsilon) \leq 1 - \mu_p(\mathcal{T}_1^\varepsilon > N) = 1 - \int_L \mathcal{L}_{L,\varepsilon}^N(\phi_L)\,dx = 1 - \lambda_{p,\varepsilon}^N \int Q_{L,\varepsilon}(\phi_L)\,dx + O(\theta^N),$$

where we have used Proposition 3 (c). But from parts (a) and (b) of the same proposition, we see that this can be made arbitrarily small for all small $\varepsilon$ by taking $\sigma$ small enough.

Because $\mu_l \ll \mathrm{Leb}$ it suffices to prove the statement for Leb.

We follow [5], Section 18. Let $S_{n,m} = \sum_{j=n+1}^{n+m} 1_{T_\varepsilon^j x \in H_\varepsilon}$. We have to show that

$$\sum_{n \leq S/\varepsilon} \int 1_{T_\varepsilon^n x \in H_\varepsilon} 1_{S_{n,\sigma/\varepsilon}(x)>0} dx = o(1), \quad \varepsilon \to 0, \sigma \to 0.$$

Take a small $r$. We say that a visit of $x$ to the hole at time $n$ is ($r$-)inessential if the length of the smoothness component of $T_\varepsilon^n L \cap H_\varepsilon$ containing $T_\varepsilon^n x$ is less than $r\varepsilon$. By Lemma 2 the probability that $x$ will have an inessential visit to the hole before time $S/\varepsilon$ is less than $CrS$ which can be made as small as we wish by taking $r$ small. Therefore it suffices to show that for any fixed $r$

$$\sum_{n \leq S/\varepsilon} \int 1_{\mathcal{E}_{n,\varepsilon}(x)} 1_{S_{n,\sigma/\varepsilon}(x)>0} dx = o(1)$$

where $\mathcal{E}_{n,\varepsilon}(x) = \{x \text{ has essential visit to the hole at time } n\}$.

$$\int 1_{\mathcal{E}_{n,\varepsilon}(x)} 1_{S_{n,\sigma/\varepsilon}(x)>0} dx = \mathrm{Leb}(\mathcal{E}_{n,\varepsilon}(x))\mathbb{P}(S_{n,\sigma/\varepsilon}(x) > 0|\mathcal{E}_{n,\varepsilon}(x)).$$

Since $\mathrm{Leb}(\mathcal{E}_{n,\varepsilon}(x)) \leq \mathrm{Leb}(x : T_\varepsilon^n x \in H_\varepsilon) \leq C\varepsilon$ it suffices to check that

$$(3) \qquad \max_{n \leq S/\varepsilon} \mathrm{Leb}(S_{n,\sigma/\varepsilon} > 0|\mathcal{E}_{n,\varepsilon}(x)) = o(1), \quad \varepsilon \to 0, \sigma \to 0.$$

In order to prove this we observe that due to assumption (II) for any fixed $M_0$, $S_{n,M_0}(x) = 0$ for all $x \in \mathcal{E}_{n,\varepsilon}$ provided that $\varepsilon$ is small enough. On the other hand if $k \geq M_0$ then by Lemma 2 applied to $H_\varepsilon$

$$\mathrm{Leb}(1_{H_\varepsilon} T_\varepsilon^{n+k} x) > 0|(\mathcal{E}_{n,\varepsilon}(x)) \leq C\left(\varepsilon + \frac{1}{\Lambda^k}\right).$$

Summing over $k \in [M_0 + 1, \kappa/\varepsilon]$ we obtain (3). $\qquad\qquad\square$

## 4.2. **Convergence of the finite dimensional distributions.**

*Proof of Theorem 1.* The proof is by induction on $p$. First, for $p = 1$ assume that the initial distribution of $x$ is chosen according to some density $\rho \in BV(\mu_j)$. Then by Proposition 3(c)

$$\mu_j(\mathcal{T}_1^\varepsilon = n, z(T_\varepsilon^n x) = r) = \int_{H_{jr}} \mathcal{L}_{j,\varepsilon}^n(\rho)dx = \lambda_{j,\varepsilon}^n \int_{H_{jr}} (\mathcal{Q}_{j,\varepsilon}\rho)dx + O(\theta^n).$$

Recall that

$$\mathcal{Q}_{j,\varepsilon}\rho = \nu_{j,\varepsilon}(\rho)\phi_{j,\varepsilon}.$$

By Proposition 3 $\nu_{j,\varepsilon}(\rho) \to \int \rho(x)dx$, $\int_{H_{jr}} \phi_{j,\varepsilon}/\varepsilon \to \beta_{jr}$ and for $\varepsilon n \approx t$, $\lambda_{j,\varepsilon}^n \to e^{-\beta_j t}$. Thus summation over $n \in \Delta/\varepsilon$ concludes the proof for $p = 1$.

Next suppose that the statement is known for some $p$. Denote
$$\Omega = \{\mathcal{T}_k^\varepsilon \in \Delta_k, z(t_k^\varepsilon) = r_k, \text{ for } k = 1, \ldots, p\}.$$
To carry the induction step it is enough to prove that

(4)
$$\mu_j(\varepsilon \mathcal{T}_{p+1}^\varepsilon \in \Delta_{p+1} z(t_{p+1}^\varepsilon) = r_{p+1}, \Omega) \to \mu_j(\Omega)\beta_{z_p z_{p+1}} \int_{\Delta_{p+1}} e^{-\beta_{z_p} t} dt.$$

Set
$$(\mathcal{L}_\Omega A)(x) = \sum_{T_\varepsilon^{t_\varepsilon^p} y = x} \frac{A(y)}{(T_\varepsilon^{t_\varepsilon^p})'(y)}$$
where the sum is taken over $y \in \Omega$. Then

(5)
$$\mu_j((\varepsilon \mathcal{T}_{p+1}^\varepsilon = n z(t_{p+1}^\varepsilon) = r_{p+1}, \Omega) = \int_{H_{z_p z_{p+1}}} \mathcal{L}_{j,\varepsilon}^n (\mathcal{L}_\Omega(1_{I_j} \phi_j)) dx.$$

Take a small $\sigma$ and rewrite
$$\mathcal{L}_{j,\varepsilon}^n (\mathcal{L}_\Omega(1_{I_j} \phi_j)) = \mathcal{L}_{j,\varepsilon}^{n-\sigma/\varepsilon} \left[ \mathcal{L}_{j,\varepsilon}^{\sigma/\varepsilon} (\mathcal{L}_\Omega(1_{I_j} \phi_j)) \right].$$

Due to Lasota-Yorke inequality
$$\left[ \mathcal{L}_{j,\varepsilon}^{\sigma/\varepsilon} (\mathcal{L}_\Omega(1_{I_j} \phi_j)) \right]$$
has bounded BV-norm. Hence arguing as in the $p = 1$ case we see that (5) is asymptotic to
$$\beta_{z_p z_{p+1}} \int_{\Delta_{p+1}} e^{-\beta_{z_p} t} dt \int_I \left[ \mathcal{L}_{j,\varepsilon}^{\sigma/\varepsilon} (\mathcal{L}_\Omega(1_{I_j} \phi_j)) \right] dx(1 + o_{\sigma \to 0}(1)).$$

Due to Lemma 1 the last integral here equals to
$$\mu_j(\Omega, \mathcal{T}_{p+1}^\varepsilon > \sigma/\varepsilon) = \mu_j(\Omega) + o_{\sigma \to 0}(1).$$
Since $\sigma$ is arbitrary this proves (4) completing the proof of Theorem 1. $\qquad \square$

## 5. Proof of Theorem 2

In this section we deduce Theorem 2 from Theorem 1 and Proposition 2.

*Proof.* Let $\bar{A} = A - \mu_\varepsilon(A)$. By Proposition 2 for each $\delta$ we can find $S$ such that
$$\left| \sum_{n \geq S/\varepsilon} \mu_\varepsilon(\bar{A}\bar{A} \circ T_\varepsilon^n) \right| < \delta$$
so it suffices to get the asymptotics of $\mu_\varepsilon(\bar{A}\bar{A} \circ T_\varepsilon^n)$ for $n \approx t/\varepsilon$ where $t \leq S$. We need to estimate $D_n = \mu_\varepsilon(\bar{A}\bar{A} \circ T_\varepsilon^n)$. Take a natural number $n_0$. We have
$$D_n = \int \bar{A}(x)\phi_\varepsilon(x)\bar{A}(T_\varepsilon^n x)dx = \int \bar{A}(T_\varepsilon^{n_0} y) \left[ \mathcal{L}_\varepsilon^{n-n_0} (\mathcal{L}_\varepsilon^{n_0} (\bar{A}\phi_\varepsilon)) \right](y)dy.$$

Since $L_\varepsilon$ depends continuously on $\varepsilon$ as a map BV $\to L^1$ we can rewrite the last expression as
$$\int \bar{A}(T_\varepsilon^{n_0} y) \left[ \mathcal{L}_\varepsilon^{n-n_0} (P(\bar{A}\phi_\varepsilon)) \right](y)dy + O(\theta^{n_0}) + o_{\varepsilon \to 0}(1).$$

Since $T_0$ is mixing we have

$$\int \bar{A}(T_\varepsilon^{n_0} y) B(y) dy = \sum_k \left( \int_{I_k} \bar{A}(y) \phi_k dy \int_{I_k} B(y) dy \right) + [O(\theta^{n_0}) + o_{\varepsilon \to 0}(1)] \, ||A||_{\mathrm{BV}} ||B||_{\mathrm{BV}}.$$

As $\varepsilon \to 0$ the first factor converges to $\mathbf{A}(k) - \sum_l p_l \mathbf{A}(l)$ while the second term equals

$$\int_{I_k} \mathcal{L}_\varepsilon^{n-2n_0} P(\bar{A} \phi_e p) = \sum_j 1_{I_k} p_j \left( \mathbf{A}(j) - \sum_l p_l \mathbf{A}(l) \right) \mathcal{L}_\varepsilon^{n-2n_0} \phi_j dy + o_{\varepsilon \to 0}(1)$$

$$= \sum_j p_j \left( \mathbf{A}(j) - \sum_l p_l \mathbf{A}(l) \right) \mu_j (T_\varepsilon^{n-2n_0} x \in I_k).$$

Since $n_0$ was arbitrary we conclude that

$$\lim_{\varepsilon \to 0} \varepsilon \sum_{n \leq S/\varepsilon} \mu_\varepsilon (\bar{A} \bar{A} \circ T_\varepsilon^n) = \sum_{jk} \int_0^S \left[ p_j \mathbf{A}(j) \mathbf{A}(k) p_{jk}(t) - \left( \sum_j p_j \mathbf{A}(j) \right)^2 \right] dt.$$

Since $S$ is arbitrary we can let $S \to \infty$ and obtain Theorem 2. $\qquad\square$

## 6. Proof of Proposition 2

*Proof.* Recall that $\mathcal{L}_\varepsilon$ satisfy uniform Lasota-Yorke Inequality, that is there is a constant $K$ such that

(6)   $$\mathrm{var}(\mathcal{L}_\varepsilon^n A) \leq K \left[ ||A||_{L^1} + \theta^n \mathrm{var}(A) \right]$$

(7)   $$||\mathcal{L}_\varepsilon^n A||_{L^1} \leq K ||A||_{L^1}$$

Consequently it suffices to show that if $A \in \mathrm{BV}_0$ then

(8)   $$||\mathcal{L}_\varepsilon^{\kappa/2\varepsilon} A||_{L^1} \leq \frac{||A||_{\mathrm{BV}}}{4K}$$

since combining (6), (7) and (8) gives

$$||\mathcal{L}_\varepsilon^{\kappa/\varepsilon} A||_{\mathrm{BV}} \leq \frac{||A||_{\mathrm{BV}}}{2}.$$

Next mixing of $T_0$ on $I_k$ implies that if $\int_{I_k} A(x) dx = 0$ for each $k$ then

$$||\mathcal{L}_0^{n_0} A||_{L^1} \leq C \theta^{n_0} ||A||_{\mathrm{BV}}.$$

Consequently given $\delta$ we can find $n_0, \varepsilon_0$ and $\delta_1$ such that if $\varepsilon \leq \varepsilon_0$ and $|\int_{I_k} A(x) dx| \leq \delta_1$ for each $k$ then

$$||\mathcal{L}_\varepsilon^{n_0}||_{L^1} \leq \delta.$$

Let $\bar{n} = \frac{\kappa}{2\varepsilon} - n_0$. By the foregoing discussion it remains to show that for each $k$

$$\left| \int_{I_k} (\mathcal{L}_\varepsilon^{\bar{n}} A)(x) dx \right| \leq \delta_1$$

provided that $\kappa$ is large enough. Since

$$\mathcal{L}_\varepsilon^{\bar{n}} A = \mathcal{L}_\varepsilon^{\bar{n}-n_1} \mathcal{L}_\varepsilon^{n_1} A = \mathcal{L}_\varepsilon^{\bar{n}-n_1} (PA) + O(\theta^{n_1}) + o_{\varepsilon \to 0}(1)$$

we need to show that $\int_{I_k}(\mathcal{L}_\varepsilon^{\bar{n}-n_1}PA)(x)dx$ is small. By Theorem 1 this integral is asymptotic to $\sum_j \int_{I_j} A(y)dy p_{jk}(\kappa)$. As $\kappa \to \infty$ this expression converges to $p_k \int_I A(y)dy = 0$ so the result follows. $\qquad\square$

## Appendix A. Proof of Proposition 3

*Proof.* Part (a) is proven using arguments similar to the ones in [9, Section 3.2]. Namely we apply [9, Theorem 2.1] with $\mathcal{P}_0 = \mathcal{L}_0$, $\mathcal{P}_\varepsilon = \mathcal{L}_{j,\varepsilon}$, and the Banach space $\mathrm{BV}(I_j)$. This theorem says that

$$\lim_{\varepsilon \to 0} \frac{1 - \lambda_{j,\varepsilon}}{\Delta_\varepsilon} = 1 - \sum_{k \geq 0} q_k$$

with

$$\Delta_\varepsilon = \mu_j(H_{j,\varepsilon}) = \beta_j \varepsilon + o(\varepsilon),$$

$$q_k = \lim_{\varepsilon \to 0} \frac{\mu_j(T_0^{-1}T_\varepsilon^{-k}H_{j,\varepsilon}) - \mu_j(T_\varepsilon^{-(k+1)}H_{j,\varepsilon})}{\Delta_\varepsilon}.$$

Now for fixed $k$ and all small $\varepsilon$ the set $T_\varepsilon^{-k}H_{j,\varepsilon}$ consists of a finite number of intervals of size $O(\varepsilon)$ so the sizes of their preimages by $T_0$ and $T_\varepsilon$ differ by $O(\varepsilon^2)$ (here we are using assumption **(II)**, which implies that $\phi_0$ is continuous at points in $T_0^{-(k+1)}H_\varepsilon$).

This completes the proof of part (a). Parts (b) and (c) follow from [8] except for (2) which is proven in [7]. $\qquad\square$

## References

[1] V. Baladi. *Positive transfer operators and decay of correlations*, volume 16 of *Advanced Series in Nonlinear Dynamics.* World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[2] V. Baladi and L.-S. Young. On the spectra of randomly perturbed expanding maps. *Comm. Math. Phys.*, 156(2):355–385, 1993.

[3] Chernov N., Markarian R. *Chaotic billiards,* Math. Surv. & Monogr. **127** AMS, Providence, RI, 2006.

[4] Demers M. F., Young Lai-Sang *Escape rates and conditionally invariant measures* Nonlinearity **19** (2006) 377–397.

[5] Dolgopyat D. *Limit Theorems for partially hyperbolic systems,* Trans. AMS **356** (2004) 1637–1689.

[6] Freidlin M. I., Wentzell A. D. *Random perturbations of dynamical systems,* 2d ed. Grundlehren der Mathematischen Wissenschaften **260** (1998) Springer, New York.

[7] Gonzalez Tokman C., Hunt B., Wright P. *Approximating invariant densities of metastable systems,* preprint.

[8] Keller G., Liverani C. *Stability of the spectrum for transfer operators,* Annali della Scuola Normale Superiore di Pisa, Classe di Scienze **XXVIII** (1999) 141–152.

[9] Keller G., Liverani C. *Rare events, escape rates and quasistationarity: some exact formulae,* Journal of Statistical Physics **135** (2009), 519–534.

[10] A. Lasota and J.A. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.*, 186:481–488, 1973.

[11] Liverani C. *Central limit theorem for deterministic systems,* in International Conference on Dynamical Systems (Montevideo, 1995), 56–75, Pitman Res. Notes Math. Ser., **362** Longman, Harlow, 1996.

[12] Machta J., Zwanzig R. *Diffusion in a periodic Lorentz gas,* Phys. Rev. Lett. **50** (1983) 1959–1962.

Department of Mathematics, University of Maryland, College Park, MD 20742
*E-mail address*: dmitry@math.umd.edu  and  paulrite@math.umd.edu