LARGE DEVIATIONS FOR INTERMITTENT MAPS

Mark Pollicott and Richard Sharp

Warwick University and Manchester University

Abstract. In this note we study large deviation results for the Manneville-Pomeau map and related transformations with indifferent fixed points. In particular, we consider conditions under which the associated error term is polynomial or even exponential. For typical observables, polynomial estimates are optimal. However, under suitable conditions, the exponential error term arises from the compactness of the space of measures, despite the indifference of the fixed point.

0. Introduction

In this note we shall consider a standard example of intermittent behaviour, namely interval maps with indifferent fixed points, and the large deviation behaviour of its orbits. This complements our earlier work on the distribution of pre-images for such maps.

More precisely, let $T : I \to I$ be a map of the interval $I = [0, 1]$ which is expanding, except for an indifferent fixed point at $x = 0$ and assume that there is a unique finite absolutely continuous invariant probability measure μ. The classical example is the well-known Manneville-Pomeau map $T(x) = x + x^{1+\alpha}$, where $0 < \alpha < 1$.

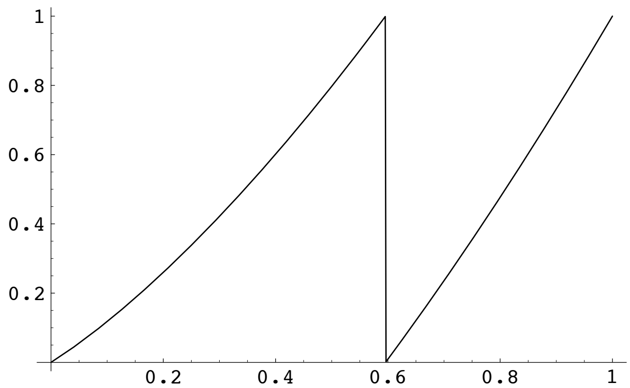

Figure 1. The Manneville Pomeau map

More generally, we can consider interval maps $T : I \to I$ such that:

(1) $T(0) = 0$;
(2) $T'(x) \geq 1$, with equality only at $x = 0$; and
(3) $T(x) = x + rx^{1+\alpha}(1 + u(x))$, where $u(x)$ is C^1 function for which $u'(x) = O(x^{t-1})$, for some $t > 0$.

These conditions are now fairly standard (cf. [5], [7], [12]).

It is well known that for such maps the decay of correlations is typically polynomial. However, as we will see, large deviation estimates may be either polynomial or exponential. We shall consider two different formulations of large deviation results [2],[3]. The first so-called *Level I* results are for functions. The second, more general but perhaps less well-known, *Level II* results are for measures.

We start by discussing Level I results. Let $f : I \to \mathbb{R}$ be a Hölder continuous function and assume without loss of generality that $\int f d\mu = 0$. We state our polynomial and exponential estimates separately.

To simplify notation, we shall write

$$E_\alpha(n) = \begin{cases} (\log n)^{2\left(\frac{1}{\alpha}-1\right)} n^{-\left(\frac{1}{\alpha}-1\right)} & \text{if } 0 < \alpha \leq \frac{1}{2} \\ n^{-\left(\frac{1}{\alpha}-1\right)} & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

Theorem 1 (Polynomial Level I result). *Suppose that $f : I \to \mathbb{R}$ is Hölder continuous and let $\epsilon > 0$.*

(a) Lower bound: If $|f(0)| \geq \epsilon$ then there exist $C > 0$ such that

$$\frac{C}{n^{\left(\frac{1}{\alpha}-1\right)}} \leq \mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\}.$$

(b) Upper bound: If $|f(0)| \geq \epsilon$ then

$$\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O(E_\alpha(n)).$$

Theorem 2 (Exponential Level I result). *Suppose that $f : I \to \mathbb{R}$ is Hölder continuous and let $\epsilon > 0$. If $|f(0)| < \epsilon$ then there exists $\beta > 0$ such that*

$$\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O(e^{-\beta n}).$$

In Theorem 1, the hypotheses preclude the function f being identically zero and give the possibility of a lower bound. However, in Theorem 2 it is still possible for f to be identically zero and thus to have no non-trivial lower bound.

Remark. Cases (a) and (b) in Theorem 1, had already been considered by Melbourne and Nicol [10], under the additional assumption that $0 < \alpha < \frac{1}{2}$. There they show that for any $\delta > 0$ there are constants $C_1, C_2 > 0$ such that

$$\frac{C_1}{n^{\left(\frac{1}{\alpha}-1\right)+\delta}} \leq \mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} \leq \frac{C_2}{n^{\left(\frac{1}{\alpha}-1\right)-\delta}}.$$

After this paper was written, we became aware that, using different methods, Melbourne [9] had independently obtained the estimate

$$\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O(n^{-(\frac{1}{\alpha}-1)}),$$

for $0 < \alpha < 1$, which is stronger than ours in the range $0 < \alpha \leq \frac{1}{2}$.

We shall now discuss the Level II results for measures. Let $\mathcal{M}$ denote the compact and convex set of Borel probability measures (with the weak* topology) and $\mathcal{M}_T$ the subset of T-invariant probability measures. We shall let δ_x denote the Dirac measure supported on the point x. In particular, δ_0 denotes the ergodic measure supported on the fixed point at 0 and we write $\mathcal{A} = \{\lambda\mu + (1-\lambda)\delta_0 : 0 \leq \lambda \leq 1\}$, the set of convex combinations of these measures.

Theorem 3 (Polynomial Level II result). *Let $\mathcal{K} \subset \mathcal{M}$ be a compact set. If $\mu \notin \mathcal{K}$ then*

$$\mu\left\{x \in I : \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i x} \in \mathcal{K}\right\} = \begin{cases} O\left((\log n)^{2(\frac{1}{\alpha}-1)} n^{-(\frac{1}{\alpha}-1)}\right) & \text{if } 0 < \alpha \leq \frac{1}{2} \\ O\left(n^{-(\frac{1}{\alpha}-1)}\right) & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

Theorem 4 (Exponential Level II result). *Let $\mathcal{K} \subset \mathcal{M}$ be a compact set. If $\mathcal{K}$ is disjoint from $\mathcal{A}$ then there exists $\beta > 0$ such that*

$$\mu\left\{x \in I : \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i x} \in \mathcal{K}\right\} = O(e^{-\beta n}).$$

The method of proof of the exponential estimates is to adapt the approach previously used in [12] for pre-images. It involves a simple argument using thermodynamic ideas, such as pressure and equilibrium states.

We now briefly outline the contents of the paper. In the next section we discuss some examples, including some numerical plots. In section 2 we prove Theorem 1, in particular using recent work of Peligrad, Utev and Wu to obtain our upper bound for $0 < \alpha < \frac{1}{2}$. In section 2 we prove Theorem 3 as a consequence of Theorem 1. In section 4, we use the thermodynamic formalism associated to T to prove Theorem 4, from which Theorem 2 follows as a simple corollary. In section 5, we discuss some higher dimensional results. We end the paper with some remarks.

1. Examples

It is easy to see the different behaviour in parts (a) and (b) of Theorem 2, say, by considering a simple class of examples. In this section we begin by empirically studying examples of Manneville-Pomeau maps for different choices of $0 < \alpha < 1$.

Example 1. Let $0 < \alpha < 1$ and consider the map $T : [0,1] \to [0,1]$ defined by $T(x) = x + x^{1+\alpha} \pmod 1$ where $0 < \alpha < 1$. T is called the Manneville Pomeau map and within this range of parameters α the map has a unique absolutely continuous invariant measure μ.

In practice, it is easier to estimate the size of the sets with respect to the Lebesgue measure λ, since we do not have an explicit expression for the density. However, if $|f(0)| \geq \epsilon$ then Theorem 1 gives that

$$\lambda\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O\left(\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\}\right)$$
$$= O\left(n^{-\frac{1}{\alpha}}\right).$$

On the other hand, if $|f(0)| < \epsilon$ then Theorem 2 gives that

$$\lambda\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O\left(\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\}\right)$$
$$= O\left(e^{-\beta n}\right).$$

Thus the same characteristic polynomial, or exponential, decay can be seen with the Lebesgue measure of the sets.

In figures 2-4 we present some numerical estimates in the three cases $\alpha = 0.5$, $\alpha = 0.25$ and $\alpha = 0.1$. In each case we takes $\epsilon = 0.5$. Furthermore, we consider the the two functions $f_1(x) = \cos(2\pi x)$ and $f_2(x) = \sin(2\pi x)$. Since $|\int f_1 d\mu| < \epsilon$ and $f_1(0) = 1 > \epsilon$ we see that case (a) applies and the first of the plots for each of parameter values exhibits the characteristic polynomial decay.

Since $|\int f_2 d\mu| < \epsilon$ and $f_2(0) = 0 < \epsilon$ we see that case (b) applies and the first of the plots for each of parameter values exhibits characteristic exponential decay.

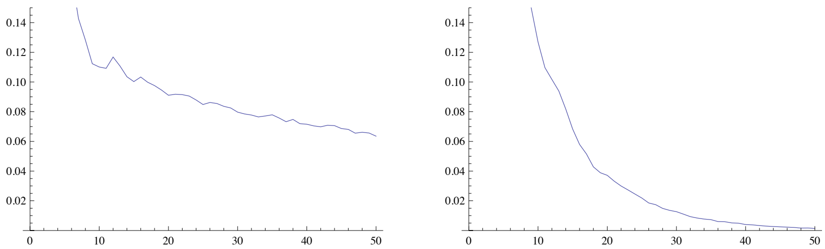

FIGURE 2. Empirical plots for $\lambda\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\}$ in the case of the Manneville-Pomeau map with $\alpha = 0.5$ and $\epsilon = 0.5$; where a) $f_1(x) = \cos(2\pi x)$; and b) $f_2(x) = \sin(2\pi x)$.

Example 2. Another related class of interval maps were studied by Yuri in [17]. These are of the general form

$$T(x) = \begin{cases} \frac{x}{(1-x^\alpha)^{1/\alpha}} & \text{if } 0 \leq x \leq 2^{-1/\alpha} \\ \frac{x}{2^{-1/\alpha}-1} - \frac{1}{1-2^{-1/\alpha}} & \text{if } 2^{-1/\alpha} < x \leq 1, \end{cases}$$

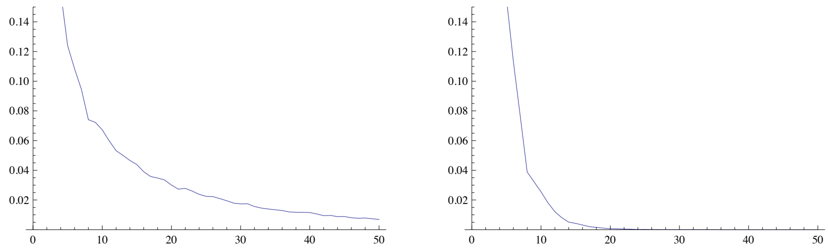

FIGURE 3. Empirical plots for $\lambda \left\{ x \in I : \left| \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \right| \geq \epsilon \right\}$ in the case of the Manneville-Pomeau map with $\alpha = 0.25$ and $\epsilon = 0.5$; where a) $f_1(x) = \cos(2\pi x)$; and b) $f_2(x) = \sin(2\pi x)$.

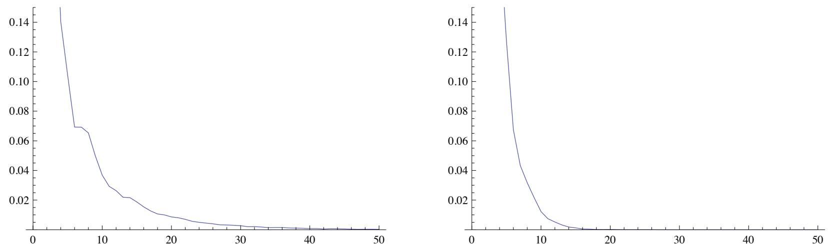

FIGURE 4. Empirical plots for $\lambda \left\{ x \in I : \left| \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \right| \geq \epsilon \right\}$ in the case of the Manneville-Pomeau map with $\alpha = 0.1$ and $\epsilon = 0.5$; where a) $f_1(x) = \cos(2\pi x)$; and b) $f_2(x) = \sin(2\pi x)$.

where $0 < \alpha < 1$. This map has a finite ergodic absolutely continuous probability measure. These satisfy the hypotheses of Theorems 1-4 and the conclusions of these results apply.

In practice, we do not have an explicit formula for the density of the measure μ. Therefore, we make numerical estimates using Lebesgue measure on the unit interval which gives some approximation to the true value using μ. The plots are then based on the behaviour of Birkhoff averages for a large number of randomly chosen points.

Finally, we mention that the same method applies to related higher dimensional examples using the methods described

2. Proof of Theorem 1

Part (a). We begin with some simple properties of the map T.

Lemma 2.1.

(i) *Let a_n denote the pre-image in $T^{-n}(1)$ which is closest to 0. Then $a_n \sim n^{-\frac{1}{\alpha}}$.*

(ii) *There exists $D_1, D_2 > 0$ such that the Radon-Nikodym derivative $d\mu/dx$ satisfies*

$$D_1 x^{-\alpha} \leq \frac{d\mu}{dx} \leq D_2 x^{-\alpha}.$$

Proof. Part (i) is proved in [7, Lemma 2.1 b)]. Part (ii) is proved in [7, Lemma 2.5] □

We shall restate the lower bound in part (a) of Theorem 1 as a proposition.

Proposition 2.1. *If $|f(0)| \geq \epsilon$ then there exists $C > 0$ such that*

$$\mu\left\{x \in I \,:\, \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} \geq \frac{C}{n^{\left(\frac{1}{\alpha}-1\right)}}, \text{ for all } n \geq 1.$$

Proof. Assume without loss of generality that $f(0) \geq \epsilon$ (the case $f(0) \leq -\epsilon$ being similar). By continuity of $f(x)$ at 0 we can choose $\delta > 0$ such that whenever $x \in [0, \delta]$ we have that $f(x) \geq \epsilon$.

We consider intervals $[0, a_n]$. By Lemma 2.1 (i) we can choose $N \geq 1$ sufficiently large that $a_N \leq \delta$. In particular, for any $n \geq 1$ we have that $T^n[0, a_{n+N}] = [0, a_N] \subset [0, \delta]$ and thus for any $x \in [0, a_{n+N}]$ we have that $\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) \notin (-\epsilon, \epsilon)$.

Moreover, by Lemma 2.1 (ii) we can find $C > 0$ such that

$$\mu([0, a_{n+N}]) \geq D_1 \int_0^{a_{n+N}} x^{-\alpha} dx \geq \frac{C}{n^{\left(\frac{1}{\alpha}-1\right)}}.$$

Combining these estimates we have that:

$$\mu\left\{x \in I \,:\, \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| > \epsilon\right\} \geq \mu([0, a_{n+N}]) \geq \frac{C_1}{n^{\left(\frac{1}{\alpha}-1\right)}}, \text{ for all } n \geq 1,$$

as required. □

Part (b). We need to consider the two ranges of values for α separately.

Case (i): $\frac{1}{2} < \alpha < 1$. Recall the following result:

Lemma 2.2 [16]**, cf. [10].** *Given a Hölder continuous function $f : I \to \mathbb{R}$ with $\int f d\mu = 0$, there exists $C(f) > 0$ such that*

$$\left|\int f(T^j x) f(x) \; d\mu(x)\right| \leq C(f) j^{-\left(\frac{1}{\alpha}-1\right)},$$

for $j \geq 1$.

We shall show the following:

Proposition 2.2. *Suppose that $|f(0)| \geq \epsilon$ and that $\frac{1}{2} < \alpha < 1$. Then*

$$\mu\left\{x \in I \,:\, \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O\left(n^{-\left(\frac{1}{\alpha}-1\right)}\right).$$

Proof. Since $\alpha > \frac{1}{2}$, we have $\frac{1}{\alpha} - 1 < 1$. Observe that

$$\begin{aligned}
&\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} \leq \frac{1}{n^2\epsilon^2}\int \left|\sum_{i=0}^{n-1} f(T^i x)\right|^2 d\mu(x) \\
&\leq \frac{1}{n^2\epsilon^2}\left(n\int |f(x)|^2 d\mu(x) + 2\sum_{0\leq i<j\leq n-1}\int f(T^{|i-j|}x)f(x)\, d\mu(x)\right) \\
&\leq \frac{1}{n^2\epsilon^2}\left(n\int |f(x)|^2 d\mu(x) + 2C(f)\sum_{r=1}^{n}\frac{n-r}{r^{\frac{1}{\alpha}-1}}\right) \qquad (2.1)\\
&= \frac{1}{n\epsilon^2}\int |f(x)|^2 d\mu(x) + \frac{C(f)}{n\epsilon^2}\sum_{r=1}^{n}\frac{1}{r^{\frac{1}{\alpha}-1}} - \frac{C(f)}{n^2\epsilon^2}\sum_{r=1}^{n} r^{2-\frac{1}{\alpha}} \\
&= O(n^{-1}) + O(n^{-(\frac{1}{\alpha}-1)}) + O(n^{-(\frac{1}{\alpha}-1)}) \\
&= O(n^{-(\frac{1}{\alpha}-1)})
\end{aligned}$$

(by bounding the finite summations using integrals). □

Case (ii): $0 < \alpha \leq \frac{1}{2}$. In order to prove the result in this case, we will need to consider the transfer operator $\mathcal{L}_{-\log|T'|} : C^0(I,\mathbb{R}) \to C^0(I,\mathbb{R})$ defined by

$$\mathcal{L}_{-\log|T'|}h(x) = \sum_{Ty=x}\frac{h(y)}{|T'(y)|}.$$

We can replace $\mathcal{L}_{-\log|T'|}$ by its normalized version using a result of Hu [5, §4]. More precisely, one can can choose a positive function u such that the operator $\mathcal{L}h = u^{-1}\mathcal{L}_{-\log|T'|}(uh)$, for $h \in C^0(I,\mathbb{R})$, satisfies $\mathcal{L}1 = 1$. Moreover, we have the following result.

Lemma 2.3. *We have the bound*

$$\int |\mathcal{L}^n f| d\mu = O(n^{1-\frac{1}{\alpha}}). \qquad (2.2)$$

Proof. This appears in [6], Proposition 5.2, (d) (iii) □

In particular, we can can write $E(f|T^{-k}\mathcal{B}) = (\mathcal{L}^k f)\circ T^k$, where $\mathcal{B}$ denotes the usual σ-algebra on $[0,1]$, cf. [8] and [15, §2]. More precisely, $E(f|T^{-k}\mathcal{B})$ is the unique integrable function which is measurable with respect to the smaller sigma algebra $T^{-k}\mathcal{B} \subset \mathcal{B}$ and satisfies

$$\int_A f d\mu(x) = \int_A E(f|T^{-k}\mathcal{B})d\mu(x)$$

for every $B \in T^{-k}\mathcal{B}$.

We shall show the following.

Proposition 2.3. *Suppose that $|f(0)| \geq \epsilon$ and that $0 < \alpha \leq \frac{1}{2}$. Then*

$$\mu\left\{x \in I \,:\, \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O\left((\log n)^{2(\frac{1}{\alpha}-1)} n^{-(\frac{1}{\alpha}-1)}\right).$$

Proof. We begin with the standard estimate, for any $\epsilon > 0$, that

$$\mu\left\{x \in I \,:\, \left|\sum_{i=0}^{n-1} f(T^i x)\right| > n\epsilon\right\} \leq \frac{1}{(n\epsilon)^{2p}} \int \left|\sum_{i=0}^{n-1} f(T^i x)\right|^{2p} d\mu(x) \tag{2.3}$$

where $p > 1$, say. Since f is bounded, we can take p as large as we please. However, we need to make sure that

$$\int \left|\sum_{i=0}^{n-1} f(T^i x)\right|^{2p} d\mu(x)$$

does not grow too fast.

We consider the natural extension of T (to obtain an invertible system) and apply Corollary 1 from [11] to get:

$$\int \left|\sum_{i=0}^{n-1} f(T^i x)\right|^{2p} d\mu(x) \leq Cn^p(\|f\|_{2p} + 240\sum_{k=1}^{n} k^{-1/2}\|E(f|T^{-k}\mathcal{B})\|_{2p})^{2p}, \tag{2.4}$$

for some $C > 0$, where $\|\cdot\|_{2p}$ denotes the L^{2p}-norm. Moreover, by invariance of μ we can write

$$\begin{aligned}
\|E(f \circ T^{-k}|\mathcal{B})\|_{2p} &= \left(\int |E(f \circ T^{-k}|\mathcal{B})|^{2p} d\mu\right)^{\frac{1}{2p}} \\
&= \left(\int |E(f|T^{-k}\mathcal{B})|^{2p} d\mu\right)^{\frac{1}{2p}} \\
&= \|E(f|T^{-k}\mathcal{B})\|_{2p}.
\end{aligned}$$

Again using the T-invariance of μ we can write

$$\begin{aligned}
\|E(f|T^{-k}\mathcal{B})\|_{2p} &= \left(\int |E(f|T^{-k}\mathcal{B})|^{2p} d\mu\right)^{\frac{1}{2p}} \\
&= \left(\int |\mathcal{L}^k f|^{2p} d\mu\right)^{\frac{1}{2p}} \\
&= \|\mathcal{L}^k f\|_{2p}.
\end{aligned}$$

Furthermore, using (2.2) we see that

$$\begin{aligned}
\|\mathcal{L}^n f\|_{2p} &\leq \left(\int |\mathcal{L}^n f|^{2p} d\mu\right)^{\frac{1}{2p}} \\
&\leq \|f\|_\infty^{(2p-1)/(2p)} \left(\int |\mathcal{L}^n f| d\mu\right)^{\frac{1}{2p}} \\
&= O\left(n^{\left(1-\frac{1}{\alpha}\right)/2p}\right).
\end{aligned} \tag{2.5}$$

Moreover, if we let $p = \frac{1}{\alpha} - 1$ then we see from (2.5) that the series on the Right Hand Side of (2.4) is $O(\log n)$. Thus the Right Hand Side of (2.4) is $O(n^p(\log n)^{2p})$ and comparing (2.3) and (2.4) gives:

$$\mu\left\{x \in I : \left|\sum_{i=0}^{n-1} f(T^i x)\right| > n\epsilon\right\} = O\left(n^{-p}(\log n)^{2p}\right)$$
$$= O\left(n^{-\left(1-\frac{1}{\alpha}\right)}(\log n)^{2\left(1-\frac{1}{\alpha}\right)}\right)$$

which gives the required bound. □

Remarks.

(1) If $\alpha = \frac{1}{2}$ then, in fact, the method used in Case (i) in gives the estimate

$$\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O\left((\log n)n^{-1}\right).$$

(2) If we chose $p < (1 - \frac{1}{\alpha})$ then the bound gives that (2.5) and the series in (2.4) are uniformly bounded in n. Therefore, comparing (2.3) and (2.4) gives that

$$\mu\left\{x \in I : \left|\sum_{i=0}^{n-1} f(T^i x)\right| > n\epsilon\right\} = O(n^{-p})$$

for $p < (1-\frac{1}{\alpha})$, which recovers the bound in [10]. This is not surprising since both approaches use Burkholder-type inequalities. However, by avoiding an explicit use of Martingale differences, and instead using the estimates in [11], we have the upper bounds in Theorem 1.

(3) There are other examples of statistical properties of the Manneville-Pomeau map where one observes a dichotomy in terms of the error terms which is a consequence of the value of the function at the indifferent fixed point. In particular, Gouëzel [4] has such results for the error term of the Central Limit Theorem when $0 < \alpha < \frac{1}{2}$.

3. Proof of Theorem 3

We shall show the following restatement of Theorem 3:

Proposition 3.1. *Assume that $\mathcal{K}$ is a compact subset of $\mathcal{M}$ such that $\mu \notin \mathcal{K}$. Then*

$$\mu\left\{x \in I : \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i x} \in \mathcal{K}\right\} = O(E_\alpha(n)).$$

Proof. We shall derive this result from Propositions 2.1 and 2.2. Let $C^\theta(I, \mathbb{R})$ denote the space of Hölder continuous functions $f : I \to \mathbb{R}$ with Hölder exponent $\theta > 0$. Let $\nu \in \mathcal{K}$. Since $\mu \notin \mathcal{K}$ and, for any fixed $\theta > 0$, $C^\theta(I, \mathbb{R})$ is uniformly dense in $C^\theta(I,\mathbb{R})$, we can find $\epsilon > 0$ and $f \in C^\theta(I,\mathbb{R})$ such that $\left|\int f d\nu - \int f d\mu\right| > \epsilon$. We can cover the arbitrary compact set of measures $\mathcal{K}$ by a union of open sets in $\mathcal{M}$:

$$\mathcal{K} \subset \bigcup_{f \in C^\theta(I,\mathbb{R}))} \bigcup_{\epsilon>0} \left\{ \nu \in \mathcal{M} : \left| \int f d\mu - \int f d\nu \right| > \epsilon \right\}.$$

By compactness, we can cover $\mathcal{K}$ by a finite union

$$\mathcal{K} \subset \bigcup_{j=1}^{N} \left\{ \nu \in \mathcal{M} : \left| \int f_j d\mu - \int f_j d\nu \right| > \epsilon_j \right\},$$

where $\epsilon_j > 0$ and $f_j \in C^\theta(I,\mathbb{R})$, $j = 1, \ldots, N$. In particular, we have that

$$\begin{aligned} \mu\left\{ x \in I : \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i x} \in \mathcal{K} \right\} &\leq \sum_{j=1}^{N} \mu \left\{ x \in I : \left| \frac{1}{n}\sum_{i=0}^{n-1} f_j(T^i x) - \int f_j d\mu \right| > \epsilon_j \right\}. \\ &= O(E_\alpha(n)). \end{aligned}$$

by applying Propositions 2.1 and 2.2 to the functions $f_j - \int f_j d\mu$. □

4. Proof of Theorems 2 and 4

We turn to the proof of the exponential estimate in Theorem 4. (Theorem 2 will follow as a corollary.) Let $P : C^0(I,\mathbb{R}) \to \mathbb{R}$ be the usual pressure function

$$P(g) = \sup_{\mu \in \mathcal{M}_T} \left\{ h(\mu) + \int g \, d\mu \right\}.$$

We can then define an auxiliary function $Q : C^0(I,\mathbb{R}) \to \mathbb{R}$ by $Q(g) = P(-\log|T'| + g)$ and, for each $\nu \in \mathcal{M}$, associate its Legendre transform

$$I(\nu) = \sup_{g \in C^0(I,\mathbb{R})} \left(\int g \, d\nu - Q(g) \right).$$

In particular, recall that $P(\cdot)$ is Lipschitz continuous and so $Q(\cdot)$ is Lipschitz continuous.

Lemma 4.1 [12, Lemma 8]**.**

(i) *If $\nu \notin \mathcal{A}$ then $I(\nu) > 0$.*
(ii) *The map $\nu \mapsto I(\nu)$ is lower semi-continuous on $\mathcal{M}_T$. Furthermore, on $\mathcal{M} - \mathcal{M}_T$, $I(\nu)$ is bounded below by the continuous function $\nu \mapsto \int \log|T'| d\nu$.*

In particular, if $\mathcal{K} \cap \mathcal{A} = \varnothing$ then $\rho = \inf_{\nu \in \mathcal{K}} I(\nu) > 0$.

For each $g \in C^0(I,\mathbb{R})$, we can generalize the definition of $\mathcal{L}_{-\log|T'|}$ is section 2 and define the transfer operator $\mathcal{L}_g : C^0(I,\mathbb{R}) \to C^0(I,\mathbb{R})$ by

$$\mathcal{L}_g h(x) = \sum_{Ty=x} e^{g(y)} h(y).$$

Lemma 4.2 [12, Lemma 1]**.** *For any $g \in C^0(I, \mathbb{R})$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \int \mathcal{L}_g^n 1(x) d\mu(x) \le P(g). \tag{4.1}$$

Proof. The map $T : I \to I$ is topologically conjugate to the doubling map, which is itself semi-conjugate to the full shift on two symbols $\sigma : \Sigma_2 \to \Sigma_2$. Lifting g to a continuous function $\bar{g} : \Sigma_2 \to \Sigma_2$, given $\eta > 0$, we may find a locally constant function $k : \Sigma_2 \to \mathbb{R}$ such that $\|\bar{g} - k\|_\infty \le \eta/4$. We shall continue to use $\mathcal{L}_{\bar{g}}$ and $P(\bar{g})$, respectively, to denote the transfer operator and pressure with respect to σ. Then

$$\frac{1}{n} \log \mathcal{L}_{\bar{g}}^n 1 \le \frac{1}{n} \log \mathcal{L}_{k+\eta/4}^n 1 \le P\left(k + \frac{\eta}{4}\right) + \frac{\eta}{4} \le P(\bar{g}) + \frac{\eta}{2},$$

for sufficiently large n. Thus

$$\lim_{n\to+\infty} \frac{1}{n} \log \int \mathcal{L}_{\bar{g}}^n 1(x)\ d\mu(x) \le \exp(n(P(\bar{g}) + \eta/2)).$$

Since $P(\bar{g}) = P(g)$ and $\eta > 0$ is arbitrary, the result is proved. $\square$

To complete the proof of Theorem 2, it suffices to show the following.

Proposition 4.1. *Let $\mathcal{K}$ be a compact subset of $\mathcal{M}$ such that $\mathcal{K} \cap \mathcal{A} = \varnothing$. Then, for any $\tau > 0$, we can choose $C > 0$ such that*

$$\mu\left\{x \in I\ :\ \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i x} \in \mathcal{K}\right\} \le Ce^{-(\rho-\tau)n}.$$

Proof. By the definition of $I(\nu)$ and ρ in Lemma 4.1, for every $\nu \in \mathcal{K}$, there exists $g \in C^0(I, \mathbb{R})$ such that

$$\int g\ d\nu - Q(g) > \rho - \tau.$$

Thus we have that

$$\mathcal{K} \subset \bigcup_{g \in C^0(I,\mathbb{R})} \left\{\nu \in \mathcal{M} :\ \int g\ d\nu - Q(g) > \rho - \tau\right\}$$

Since $\mathcal{K}$ is compact, we may choose a finite collection of functions $g_1, \dots, g_k \in C^0(I, \mathbb{R})$ such that

$$\mathcal{K} \subset \bigcup_{j=1}^{k} \left\{\nu \in \mathcal{M} :\ \int g_j\ d\nu - Q(g_j) > \rho - \tau\right\}.$$

In particular, we can bound

$$\begin{aligned} \mu\left\{x \in I :\ \frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i x} \in \mathcal{K}\right\} &\le \sum_{j=1}^{k} \mu\left\{x \in I :\ \frac{1}{n} g_j^n(x) - Q(g_j) > \rho - \tau\right\} \\ &\le \sum_{j=1}^{k} e^{-n(Q(g_j)+\rho-\tau)} \int \mathcal{L}_{\log|T'|+g_j}^n 1\ d\mu. \end{aligned} \tag{4.2}$$

Comparing (4.1) and (4.2) we have that

$$\mu\left\{x : \frac{1}{n}\sum_{i=0}^{n-1}\delta_{T^i x} \in \mathcal{K}\right\} = O(e^{-(\rho-\tau)n}). \qquad \square$$

Proof of Theorem 2. Theorem 2 follows from Theorem 4 by choosing

$$\mathcal{K} = \left\{\nu \in \mathcal{M} : \left|\int f d\nu\right| \geq \epsilon\right\}.$$

This is a compact set in $\mathcal{M}$. Furthermore, if (for $0 \leq \lambda \leq 1$) $\nu = \lambda\mu + (1-\lambda)\delta_0 \in \mathcal{A}$ then

$$\left|\int f d\nu\right| \leq (1-\lambda)|f(0)| < \epsilon,$$

so $\mathcal{K} \cap \mathcal{A} = \varnothing$. Finally,

$$\left\{x \in I : |\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)| \geq \epsilon\right\} \subset \left\{x \in I : \frac{1}{n}\sum_{i=0}^{n-1}\delta_{T^i x} \in \mathcal{K}\right\}.$$

Thus

$$\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O(e^{-\beta n}),$$

for $\beta > 0$.

5. Higher dimensional results

The results in [12] for higher dimensional maps have natural analogues in the present context.

Let $T : X \to X$ be a continuous expanding map on a compact subset of R^k and let $X = \cup_{a \in \mathcal{I}} X_a$ be a finite generating partition, where each X_a has a piecewise smooth boundary and for each a the map $T : X_a \to X$ is a C^1 diffeomorphism. Suppose also that $\inf_{x \in X} |\det(DT)(x)| > 0$. Moreover, we require the following technical hypothesis:

Finite Range Condition. There is a finite collection $\mathcal{U}$ of open subsets of X such that, if $\mathrm{int}(X_{a_1}) \cap \mathrm{int}(T^{-1}X_{a_2}) \cap \cdots \cap \mathrm{int}(T^{-(n-1)}X_{a_n}) \neq \varnothing$ then $T^n(\mathrm{int}(X_{a_1}) \cap \mathrm{int}(T^{-1}X_{a_2}) \cap \cdots \cap \mathrm{int}(T^{-(n-1)}X_{a_n})) \in \mathcal{U}$.

We write $X_{a_1 \cdots a_n}$ for the cylinder set $X_{a_1} \cap T^{-1}X_{a_2} \cap \cdots \cap T^{-(n-1)}X_{a_n}$. Fix $C > 0$. For each $n \geq 1$, let $R_n = R_n(C)$ denote the cylinders $X_{a_1 \cdots a_n}$ such that

$$\sup_{x,y \in X_{a_1 \cdots a_n}} \frac{|\det(DT^n x)|}{|\det(DT^n y)|} < C$$

and we define $D_n = D_n(C) \subset X$ to be the union of the cylinders $X_{a_1 \cdots a_n}$ for which, for each $1 \leq i \leq n$, we have that

$$\sup_{x,y \in X_{a_1 \cdots a_i}} \frac{|\det(DT^i x)|}{|\det(DT^i y)|} \geq C.$$

Lemma 5.1. *There is an ergodic absolutely continuous T-invariant measure μ provided*

(i) *there exists $C > 0$ such that $R_n \neq \varnothing$ and whenever $X_{a_1 \cdots a_n} \in R_n$ then $X_{b_1 \cdots b_m a_1 \cdots a_n} \in R_{n+m}$;*

(ii) *for each $U \in \mathcal{U}$ there exists $X_{a_1 \cdots a_n} \subset U$ such that $X_{a_n} \in R_1$ and $T^n(X_{a_1 \cdots a_n}) = X$; and*

(iii) *$\sum_{n=1}^{\infty} \lambda(D_n) < +\infty$, where λ is the usual k-dimensional Lebesgue measure*

Proof. This appears as Proposition 3 in [12]. □

Let $\mathcal{A}$ denote the convex hull of μ and the set of measures supported on $\cap_{n=1}^{\infty} D_n$. (For simplicity, we assume the latter is a single periodic orbit.)

In [12] we proved large deviation results for periodic orbits. However, a simple modification of the arguments there leads to a statement analogous to that of Theorem 4:

Proposition 5.1. *Assume that $T : X \to X$ satisfies the Finite Range Condition and satisfies (i)-(iii) above. Let $\mathcal{K} \subset \mathcal{M}$ be a compact set. If $\mathcal{K}$ is disjoint from $\mathcal{A}$ then there exists $\beta > 0$ such that*

$$\mu \left\{ x \in I \, : \, \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i x} \in \mathcal{K} \right\} = O(e^{-\beta n}).$$

Example (Brun's map). Let $X = \{(x, y) : 0 \leq y \leq x \leq 1\}$ and consider the finite partition

$$X_0 = \left\{ (x, y) \, : \, 0 \leq y \leq x \leq \frac{1}{2} \right\},$$

$$X_1 = \left\{ (x, y) \, : \, 0 \leq y \leq 1 - x, \frac{1}{2} \leq x \leq 1 \right\},$$

$$X_2 = \left\{ (x, y) \, : \, 1 - x \leq y \leq x, \frac{1}{2} \leq x \leq 1 \right\}.$$

Define a map $T : X \to X$ by

$$T(x, y) = \begin{cases} \left(\frac{x}{1-x}, \frac{y}{1-x} \right) & \text{if } (x, y) \in X_0 \\ \left(\frac{1}{x} - 1, \frac{y}{x} \right) & \text{if } (x, y) \in X_1 \\ \left(\frac{y}{x}, \frac{1}{x} - 1 \right) & \text{if } (x, y) \in X_2 \end{cases}$$

This map has a finite ergodic absolutely continuous probability measure by the Lemma 5.1, and an exponential Level II large deviation results by Proposition 5.1. (In this example, the indifferent periodic orbit is the fixed point $(0, 0)$.)

We may also consider the case of polynomial large deviation results, where there are analogues of Theorems 1 and 3. For example, if, for any Lipschitz function, the transformation mixes at a polynomial rate $O(n^{-\gamma})$, for some $\gamma > 0$, then, for a compact set $\mathcal{K} \subset \mathcal{M}$ not containing μ, we have

$$\mu \left\{ x \in X \, : \, \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i x} \in \mathcal{K} \right\} = O(n^{-\min\{1, \gamma\}}).$$

Also, assuming that an indifferent fixed point x_0 has neighbouring cylinders $X_{a_1\cdots a_n}$ whose measures have a polynomial lower bound $\mu(X_{a_1\cdots a_n}) \geq C_1 n^{-\gamma}$, for some $C_1 > 0, \gamma > 0$, and given a Hölder continuous function $f : X \to \mathbb{R}$ with $\int f d\mu = 0$, there exists $C_2 > 0$ such that

$$\mu\left\{x \in X : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} \geq C_2 n^{-\gamma}.$$

Example (Inhomogeneous diophantine approximation map). Let $X = \{(x, y) : 0 \leq y \leq 1, -y \leq x < y + 1\}$ and define $T : X \to X$ by

$$T(x, y) = \left(\frac{1}{x} - \left[\frac{1-y}{x}\right] + \left[-\frac{y}{x}\right], -\left[-\frac{y}{x}\right] - \frac{y}{x}\right).$$

By Lemma 5.1 there is an ergodic absolutely continuous T-invariant measure μ. Moreover, for any Lipschitz function it mixes at a rate $O(n^{-(1-\delta)})$, for any $\delta > 0$ and thus the upper bound above applies.

Analogous results hold in the case of Almost Anosov diffeomorphisms, in the sense of Huyi Hu. These are particular examples of surface diffeomorphisms which are hyperbolic, except at a fixed point. Hu gives conditions on the derivative at the fixed point for such maps to have a finite SRB measure and for this measure to mix at a polynomial rate [5, Theorem 3.9 i) and Theorem 4.2]. However, since they are continuous factors of finite type the preceding analysis applies.

6. Final Remarks

1. There are simple examples of interval maps which have more than one indifferent fixed points, and where the polynomial decay of correlations is controlled by the (finite) set of such points (cf. [1], for example). In this case Theorem 1 has a natural generalization in which $\mathcal{A}$ is replaced by a finite dimensional simplex.

2. The results of [11], and the method of proof in section 2, give a slightly stronger maximal result:

$$\mu\left\{x \in I : \max_{1\leq m\leq n}\left|\frac{1}{m}\sum_{i=0}^{m-1} f(T^i x)\right| \geq \epsilon\right\} = \begin{cases} O\left((\log n)^{2\left(\frac{1}{\alpha}-1\right)} n^{-\left(\frac{1}{\alpha}-1\right)}\right) & \text{if } 0 < \alpha \leq \frac{1}{2} \\ O\left(n^{-\left(\frac{1}{\alpha}-1\right)}\right) & \text{if } \frac{1}{2} < \alpha < 1. \end{cases}$$

3. There is an interesting connection between error terms for large deviations and the approximation of functions by coboundaries. We recall that the usual proof of the L^2 von Neumann ergodic theorem for an ergodic transformation $T : (X, \mu) \to X, \mu)$ exploits the fact that for any $f \in L^2(X, \mu)$ with $\int f d\mu = 0$, and any $\delta > 0$ we can choose $u \in L^2(X, \mu)$ such that $\|f - (u \circ T - u)\|_2 < \delta$. Volný and Weiss [14] showed that whenever $f \in L^\infty(X, \mu)$ with $\int f d\mu = 0$, satisfies that for any $\delta > 0$ we can choose $u \in L^p(X, \mu)$ such that $\|f - (u \circ T - u)\|_\infty < \delta$, then

$$\mu\left\{x \in I : \left|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)\right| \geq \epsilon\right\} = O\left(n^{-p}\right).$$

In particular, we see from the lower bound in Theorem 1 (a) that if $|f(0)| \geq \delta$ then

$$\inf\left\{\|f-(u\circ T-u)\|_\infty : u \in L^{\left(\frac{1}{\alpha}-1\right)}(X,\mu)\right\} > 0.$$

References

1. D. Arrowsmith and M. Barenco, *The autocorrelation of double intermittency maps and the simulation of computer packet traffic*, Dyn. Syst. **19** (2004), 61-74.
2. A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*, Springer-Verlag, New York, 1998.
3. R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, New York, 1985.
4. S. Gouëzel, *Berry-Esseen theorem and local limit theorem for non uniformly expanding maps*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), 997-1024.
5. H. Hu, *Statistical properties of some almost hyperbolic systems*, Smooth ergodic theory and its applications (Seattle, WA, 1999) (R. de la Llave, A. Katok, Ya. B. Pesin, and H. Weiss, eds.), Proc. Sympos. Pure Math., 69, Amer. Math. Soc., Providence, RI,, 2001, pp. 367-384.
6. H. Hu, *Decay of correlations for piecewise smooth maps with indifferent fixed points*, Ergod. Th. and Dynam. Sys. **24** (2004), 495-524.
7. S. Isola, *Dynamical zeta functions and correlation functions for non-uniformly hyperbolic systems*, preprint.
8. F. Ledrappier, *Principe variationnel et systèmes dynamiques symboliques*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **30** (1974), 185-202.
9. I. Melbourne, *Large and moderate deviations for slowly mixing dynamical systems*, Proc. Amer. Math. Soc. **137** (2009), 1735-1741.
10. I. Melbourne and M. Nicol, *Large deviations for nonuniformly hyperbolic systems*, Trans. Amer. Math. Soc **360** (2008), 6661-6676.
11. M. Peligrad, S. Utev and W.B. Wu, *A maximal L^p inequality for stationary sequences and its applications*, Proc. Amer. Math. Soc. **135** (2007), 541-550.
12. M. Pollicott, R. Sharp and M. Yuri, *Large deviations for maps with indifferent fixed points*, Nonlinearity **11** (1998), 1173-1184.
13. L. Rey-Bellet and L.-S. Young, *Large Deviations in nonuniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **28** (2008), 587-612.
14. D. Volný and B. Weiss, *Coboundaries in L_0^∞*, Ann. Inst. H. Poincaré Probab. Statist. **40** (2004), 771-778.
15. P. Walters, *Ruelle's operator theorem and g-measures*, Trans. Amer. Math. Soc. **214** (1975), 375-387.
16. L.-S. Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153-188.
17. M. Yuri, *On the convergence to equilibrium states for certain non-hyperbolic systems*, Ergodic Theory Dynam. Systems **17** (1997), 977-1000.

Mark Pollicott, Department of Mathematics, University of Warwick, Coventry CV4 7AL, UK

Richard Sharp, School of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, UK