

ON DISTRIBUTIONAL LIMIT LAWS FOR RECURRENCE

MARK HOLLAND AND MIKE TODD

ABSTRACT. For a probability measure preserving dynamical system (\mathcal{X}, f, μ) , the Poincaré Recurrence Theorem asserts that μ -almost every orbit is recurrent with respect to its initial condition. This motivates study of the statistics of the process $X_n(x) = d(f^n(x), x)$, and real-valued functions thereof. For a wide class of non-uniformly expanding dynamical systems, we show that the time- n counting process $R_n(x)$ associated to the number recurrences below a certain radii sequence $r_n(\tau)$ follows an *averaged* Poisson distribution $G(\tau)$. Furthermore, we obtain quantitative results on almost sure rates for the recurrence statistics of the process X_n .

1. INTRODUCTION

We consider a measure preserving dynamical system (\mathcal{X}, f, μ) , where $f : \mathcal{X} \rightarrow \mathcal{X}$ is a discrete time map preserving the measure μ . When $\mu(\mathcal{X}) = 1$, we know from the Poincaré Recurrence Theorem that for μ -almost every $x \in \mathcal{X}$, there exists a subsequence (n_k) such that $d(f^{n_k}(x), x) \rightarrow 0$, where ' d ' is the ambient metric. A fundamental problem in quantitative recurrence theory is to establish limit laws governing the statistics of the process $X_n(x) = d(f^n(x), x)$, such as those governing the frequency of close returns of $f^n(x)$ to x . We consider distributional limit laws associated to the number of visits to sets of the form $\{X_n(x) < r_n\}$, with $r_n \rightarrow 0$.

Given a sequence (r_n) we consider the sequence of balls $B_n(x)$ with $B_n(x) = B(x, r_n)$. To fix notation, let

$$R_n(r, x) := \#\{j \in [1, n] : d(x, f^j(x)) \leq r\}.$$

Equivalently,

$$R_n(r_n, x) = \sum_{j=1}^n \mathbf{1}_{B_n(x)}(f^j x),$$

Date: January 24, 2024.

2020 Mathematics Subject Classification. 37A50, 37B20, 60G55, 37E05, 60G70.

Key words and phrases. Recurrence, Return time statistics, Poisson limits, almost sure limits.

MT was partially supported by the FCT (Fundação para a Ciência e a Tecnologia) project 2022.07167.PTDC. MH acknowledges University of St Andrews for hospitality. Both authors gratefully acknowledge LMS Scheme 4 funding.

where $\mathbf{1}_A(x)$ denotes the indicator function for the set A . Thus $R_n(r_n, x)$ is a counting function of close recurrence relative to the sequence (r_n) . Related to recurrence is that of *hitting*. Let $A \subset X$ be a measurable set, and consider the counting process defined by

$$N_n(A, x) := \#\{j \in [1, n] : f^j(x) \in A\} = \sum_{j=1}^n \mathbf{1}_A(f^j x).$$

In the case A does not depend on x (or its orbit), then $N_n(A, x)$ is a hitting-time counting function relative to A . For example, A could be a ball $B(\zeta, r)$ centred at some $\zeta \in \mathcal{X}$, and r could depend on n with $r_n \rightarrow 0$.

In this paper our first main result is Theorem 2.3 where we establish distributional limit laws for the process $R_n(r_n, x)$ and contrast it with those limit laws associated to the process $N_n(A, x)$ with $A = B(\zeta, r_n)$, and ζ a generic point in \mathcal{X} . To obtain distributional limit laws, we seek a non-degenerate function $G(\tau)$, and a sequence $r_n(\tau)$ such that

$$R_n(r_n(\tau), x) \rightarrow G(\tau), \quad (n \rightarrow \infty),$$

with convergence in distribution. Here $\tau > 0$ is a real number. Since R_n is a counting random variable, we are also interested in the functional forms $G(\tau) \equiv G(\tau, k)$ in the cases $\lim_{n \rightarrow \infty} \mu\{x : R_n(r_n(\tau), x) = k\}$, for $k \in \mathbb{N}$.

To guide us on the functional form for G , consider an i.i.d processes \hat{X}_n with distribution function $F_{\hat{X}}(u) := P(\hat{X} < u)$. Suppose that for given $\tau > 0$ we choose $u_n(\tau)$ so that $n(1 - F_{\hat{X}}(u_n)) \rightarrow \tau$ as $n \rightarrow \infty$. Then

$$\mathbb{P}(\#\{j \leq n : \hat{X}_j > u_n\} = k) \rightarrow \frac{\tau^k e^{-\tau}}{k!},$$

which coincides with the Poisson distribution. For certain dynamical processes of the form $Y_n = \phi(f^n(x))$, where $\phi : \mathcal{X} \rightarrow \mathbb{R}$ is a real-valued observable, then corresponding Poisson laws are established, see [FFT]. The form of the observable is usually important in obtaining the standard Poisson limit law, and usually we require $\phi(x)$ is *distance-like* so that $\phi(x) = \psi(d(x, \zeta))$ for some monotone decreasing function $\psi : [0, \infty) \rightarrow [0, \infty]$. More importantly, the reference point ζ is usually chosen independently of the process $(Y_n)_{n \geq 0}$. In our case, as relevant for recurrence analysis, the point ζ is actually the initial orbit value. This significantly influences the limit laws that arise relative to the case where ζ is chosen independent of the orbit.

Our second main result is Theorem 2.7. This result gives almost sure bounds (both lower and upper) on $\min_{j \leq n} d(f^j x, x)$. Such results complement the distributional limit laws obtained for R_n , and are based on a more refined analysis of the event $\{R_n(r_n, x) = 0\}$. These results produce a further understanding of quantitative Poincaré recurrence

statistics compared to works of Boshernitzan [B], and more recent works such as [KKP2, P].

This paper is organised as follows. In Section 2 we introduce the main assumptions on the system (\mathcal{X}, f, μ) . These assumptions are based upon the mixing properties (such as having exponential decay of correlations), regularity of the invariant density (e.g., of bounded variation type), and quantitative estimates on the recurrence time statistics over short time scales. We then state the main results, namely Theorem 2.3 concerning a distributional limit law for R_n , and then Theorem 2.7 regarding an almost sure limit law for R_n . In Section 3 we show through specific case studies that the conclusions of the results obtained in Section 2 can also apply to systems with weaker assumptions on their mixing rates, and/or assumptions on the invariant density. This captures systems with polynomial decay of correlations, and/or systems with unbounded invariant density. The main result here is Theorem 3.2. In Sections 4, 5 and 6 we prove Theorems 2.3, 2.7 and 3.2 respectively. Finally in the appendix we give some extensions of Theorem 2.7.

2. UNIFORMLY EXPANDING INTERVAL MAPS

We consider a measure preserving system (\mathcal{X}, f, μ) , with μ an ergodic measure. The systems we consider include those that can be modelled by a Young tower [Y2]. However we do not need the precise Young tower construction in our analysis. We focus on interval maps rather than general hyperbolic systems. Applications include piecewise expanding interval maps [LY], quadratic maps at Benedicks-Carleson parameters [Y1], and maps with neutral fixed points (intermittent maps) [FFTV]

We introduce the following definitions and assumptions. We begin with the definition of decay of correlations.

Definition 2.1. We say that (\mathcal{X}, f, μ) has decay of correlations in (Banach spaces) \mathcal{B}_1 versus \mathcal{B}_2 , with rate function $\Theta(j) \rightarrow 0$ if for all $\varphi_1 \in \mathcal{B}_1$ and $\varphi_2 \in \mathcal{B}_2$ we have

$$\mathcal{C}_j(\varphi_1, \varphi_2, \mu) := \left| \int \varphi_1 \cdot \varphi_2 \circ f^j \, d\mu - \int \varphi_1 \, d\mu \int \varphi_2 \, d\mu \right| \leq \Theta(j) \|\varphi_1\|_{\mathcal{B}_1} \|\varphi_2\|_{\mathcal{B}_2},$$

where $\|\cdot\|_{\mathcal{B}_i}$ denote the corresponding norms on the Banach spaces.

In particular, we consider the L^p and BV norms of functions $\varphi: \mathcal{X} \subset \mathbb{R} \rightarrow \mathbb{R}$ defined respectively by

$$\begin{aligned} \|\varphi\|_p &:= \left(\int |\varphi|^p \, d\mu \right)^{\frac{1}{p}}, \\ \|\varphi\|_{\text{BV}} &:= \text{var}(\varphi) + \sup(|\varphi|), \end{aligned}$$

where $\text{var}(\varphi)$ denotes the total variation of φ . In particular, for an interval $J \subset \mathcal{X}$,

$$\text{var}_J(\varphi) := \sup_{\mathcal{P}} \sum_{i=1}^m |\varphi(x_i) - \varphi(x_{i-1})|,$$

where \mathcal{P} is the partition of J into intervals $[x_i, x_{i+1}]$, with $\{x_0, x_m\} = \partial J$. The supremum is then taken over all such partitions. When $J = \mathcal{X}$, we just denote this *variation* by $\text{var}(\varphi)$. Functions $\varphi: \mathcal{X} \subset \mathbb{R} \rightarrow \mathbb{R}$ with $\|\varphi\|_{\text{BV}} < \infty$ are called functions of *bounded variation*. Throughout this paper we assume that our measure μ is ergodic and absolutely continuous with respect to Lebesgue: an *acip*.

The first main assumption is the following.

- (A1) We assume that (\mathcal{X}, f, μ) has exponential decay of correlations in Banach spaces $\mathcal{B}_1 = \text{BV}$ versus $\mathcal{B}_2 = L^1(\mu)$.

We remark that within (A1), we could also have taken $\mathcal{B}_2 = L^p(\mu)$ for some $p \in (1, \infty]$. The proofs of the main results do not rely heavily on the value of p . Moreover the case $p = \infty$ can be replaced by assuming instead decay for some $p > 1$, see [HKKP, Section 15]. For certain dynamical systems, decay of correlations may only be known for \mathcal{B}_1 the space of Hölder functions (rather than BV). As discussed in Section 4, our results also apply in this case by minor modification of the proofs.

In addition to decay of correlations we make assumptions on the recurrence statistics of orbits over early time windows. We define the following sets:

$$E_j(r_n) := \{x \in \mathcal{X} : d(x, f^j x) \leq r_n\}$$

$$\tilde{E}_g(r_n) := \bigcup_{j=1}^{g(n)} E_j(r_n),$$

where (r_n) is sequence, and $g(n)$ a monotone increasing function with $g(n) \rightarrow \infty$. We make the following assumption. In the following $a_n \approx b_n$ means there is a uniform constant $c > 1$ such that $a_n/b_n \in [c^{-1}, c]$ (as $n \rightarrow \infty$).

- (A2) Suppose that $r_n \approx n^{-a}$ for some $a > 0$. There there exists $\beta_0 \in (0, 1)$, and $C > 0$ such that for all $j \leq (\log n)^2$ we have

$$\mu(E_j(r_n)) \leq C r_n^{\beta_0}.$$

with constant C independent of j .

Consequently, this implies that for rate function $g(n) = (\log n)^2$, there exists $\beta_1 \in (0, 1)$ such that

$$\mu(\tilde{E}_g(r_n)) \leq C r_n^{\beta_1}.$$

Definition 2.2. We say that a measure preserving dynamical system (\mathcal{X}, f, μ) is *hyperbolic recurrent* if it satisfies Assumptions (A1) and (A2), and has an ergodic measure μ with invariant density $\rho \in \text{BV}$.

Examples satisfying Definition 2.2 include the Gauss map $f(x) = \lfloor \frac{1}{x} \rfloor \bmod 1$, beta-transformations $f(x) = \beta x \bmod 1$, $(\beta > 1)$, Rychlik maps, and Gibbs Markov maps [HNT1] (in all these cases μ is an acip). In particular, regarding verification of (A2) see [HKKP, Section 13], and [L+]. Our first main result is the following.

Theorem 2.3. *Suppose that (\mathcal{X}, f, μ) is a measure preserving dynamical system which is hyperbolic recurrent in the sense of Definition 2.2. Consider the process*

$$X_n = d(f^n x, x), \quad R_n(r_n, x) := \#\{1 \leq j \leq n : X_j \leq r_n\},$$

with $r_n = \tau/2n$ and $\tau > 0$. Then for all $k \geq 1$,

$$\lim_{n \rightarrow \infty} \mu(x \in \mathcal{X} : R_n(r_n, x) = k) = \int_{\mathcal{X}} \frac{\tau^k \rho(x)^{k+1}}{k!} e^{-\rho(x)\tau} dx. \quad (2.1)$$

We make several remarks.

Remark 2.4. To the best of our knowledge, this is the first time equation (2.1) has appeared as a limit law for the counting function R_n . The limit law appearing here is not Poisson, but Poisson-like in the sense that the Poisson parameter τ is weighted against the density $\rho(x)$ and then averaged over \mathcal{X} w.r.t. μ . In the case of hitting time statistics where instead the process is given by $\tilde{X}_n(x) = d(f^n x, \zeta)$, (thus defined relative to a reference point $\zeta \in \mathcal{X}$), the parameter τ is governed by the density at ζ . Moreover existence of a Poisson limit-law depends on the recurrence properties of ζ , for example ζ periodic versus non-periodic. In our case, there is no fixed reference point ζ , since our time series is based on recurrence (close returns) to the initial orbit value.

Remark 2.5. For the special case $k = 0$, Collet [C] obtained a similar limit law for certain one-dimensional non-uniformly expanding dynamical systems, such as those modelled by Young towers with exponential tails. In fact Collet considered the process $X_n = -\log d(f^n x, x)$ with modified scaling sequence $r_n(u) = u + \log n$. The corresponding limit law arising is that given in equation (2.1) (in the case $k = 0$), but replacing τ by $2e^{-u}$. Note that [PS, Proposition 3.5] gives the same result in the context of billiard maps. In Section 2.2 we comment further on the limit laws that arise when looking at functions of $d(f^n x, x)$.

As an immediate corollary, we have the following standard Poisson-law for the doubling map.

Corollary 2.6. *Consider the doubling map $f(x) = 2x \bmod 1$ (preserving Lebesgue measure). Then along the sequence $r_n = \tau/2n$ we have*

$$\lim_{n \rightarrow \infty} \text{Leb}(x \in \mathcal{X} : R_n(r_n, x) = k) = \frac{\tau^k}{k!} e^{-\tau}.$$

2.1. On almost sure behaviour of recurrence statistics. In this section we consider almost sure behaviour of the recurrence statistics, that is the determination of sequences u_n, v_n such that

$$\mu \left(x \in \mathcal{X} : u_n < \min_{k \leq n} d(f^k x, x) < v_n, \text{ eventually} \right) = 1.$$

For the related hitting time process $Y_n(x) = \psi(d(f^n x, \zeta))$, with ζ independent of the orbit of x , and ψ a monotone decreasing function, see [GNO, HKKP, HNT2] concerning almost sure results. For recurrence, there is much recent interest on obtaining almost sure bounds on $\min_{k \leq n} X_n(x)$ and solving related questions on determining the probability of the event $\{x \in \mathcal{X} : d(f^j x, x) \in [0, r_n] \text{ i.o.}\}$ when $r_n \rightarrow 0$. We refer the reader to Boshernitzan [B], and more recently [KKP2, P]. The following is our second main result.

Theorem 2.7. *Suppose that (\mathcal{X}, f, μ) satisfies the hypotheses of Theorem 2.3. Then we have the following cases.*

- (a) *(Almost sure upper bounds). Consider the following alternatives:*
 - (i) *The density ρ is strictly positive, and (r_n) is a sequence satisfying $r_n > \frac{c \log \log n}{n}$, with $2c(\min_{x \in \mathcal{X}} \rho(x)) > 1$.*
 - (ii) *The density $\rho(x)$ is uniformly Hölder continuous, and (r_n) is a sequence satisfying $r_n > \frac{(\log n)^\gamma}{2n}$, with $\gamma > \gamma_0$, and*

$$\sum_{k \geq 1} \int_{\mathcal{X}} \rho(x) e^{-\epsilon k^{\gamma_0} \rho(x)} dx < \infty, \quad (2.2)$$

valid for all $\epsilon > 0$.

Then in either case (i) or (ii) we have

$$\mu \left(x \in \mathcal{X} : \min_{k \leq n} d(f^k x, x) < r_n, \text{ eventually} \right) = 1.$$

- (b) *(Almost sure lower bound). Suppose the sequence (r_n) satisfies $\sum_{n=1}^{\infty} \int_{\mathcal{X}} \mu(B(x, r_n)) d\mu < \infty$. Then*

$$\mu \left(x \in \mathcal{X} : \min_{k \leq n} d(f^k x, x) > r_n, \text{ eventually} \right) = 1.$$

We make several remarks. For case (a), we can compare this result to [KKP2, Theorem E], where they require $r_n > 1/n^\gamma$ for $\gamma < 1$. We show that case (b) follows from [KKP2, Theorem C]. At first sight, equation (2.2) looks mysterious. Analogous criteria arise in [HKKP, Section 10]. In particular, case (a)(i) implies validity of (2.2). In (a)(ii), the assumption on the regularity of the density can be relaxed to piecewise Hölder, assuming a finite number of jump discontinuity points. This is discussed in the appendix. Furthermore, case (a)(ii) is most useful if the density $\rho(x)$ equals zero for some $x \in \mathcal{X}$, as in the following example.

Example 2.1. Suppose $\mathcal{X} = [0, 1]$ and for $C > 0$, $a > 0$ we have $\rho(x) \in [C^{-1}, C]x^a$. Thus $\rho(x) \approx x^a$ as $x \rightarrow 0$. The integral in (2.2) can be estimated via

$$\int_0^1 \rho(x) e^{-\epsilon k^\gamma \rho(x)} dx \approx \int_0^1 x^a e^{-\epsilon k^\gamma x^a} dx \leq \frac{C_a}{(\epsilon k^\gamma)^{1+\frac{1}{a}}} \int_0^\infty u^{\frac{1}{a}} e^{-u} du.$$

For the latter inequality, we made a change of variable, and C_a is a constant. In particular the latter integral is also uniformly bounded. Hence, we see that (2.2) applies provided $\gamma > \frac{a}{a+1}$.

In particular, examples of dynamical systems exhibiting zeros in their invariant density are considered in [CHMV]. However, these examples, ‘intermittent cusp maps’ are non-uniformly expanding. We give details of these in Section 3, showing how the conclusion of case (a)(ii) applies. For case (a)(i), we can be explicit on the range of c in certain examples, such as those in the following Corollary.

Corollary 2.8. *Consider the doubling map $f(x) = 2x \bmod 1$ (preserving Lebesgue measure). Then case (a)(i) of Theorem 2.7 applies along the sequence $r_n = c \log \log n / n$ with $c > 1/2$.*

2.2. On extreme statistics for recurrence. In Theorem 2.3, we considered the process $X_n = d(f^n x, x)$. A more general problem is to determine corresponding limit laws for the process $X_n^\psi = \psi(d(f^n x, x))$, where $\psi : [0, \infty) \rightarrow \mathbb{R}$ is a monotone decreasing function, thus taking maximum at 0, i.e. when $f^n(x) = x$. As discussed in Remark 2.5, the special case $\psi(z) = -\log z$ was considered in [C]. Incorporation of ψ is relevant in the study of *Extreme Value Theory*, especially for determining distribution limit laws governing the maximum process

$$M_n^\psi(x) := \max_{1 \leq k \leq n} \psi(d(f^k x, x)). \quad (2.3)$$

In this case r_n is defined relative a linear scaling sequence of the form $u/a_n + b_n$ via

$$\psi(r_n(u)) = \frac{u}{a_n} + b_n, \quad a_n, b_n \in \mathbb{R}.$$

In the case of a hitting time maximum process

$$\widetilde{M}_n^\psi(x) := \max_{1 \leq k \leq n} \psi(d(f^k x, \zeta)),$$

(defined relative to a reference point $\zeta \in \mathcal{X}$), the scaling sequences a_n, b_n and limit law function $G(u)$ for $\lim_{n \rightarrow \infty} \mu\{\widetilde{M}_n^\psi(x) \leq u/a_n + b_n\}$ can be understood for a wide class of dynamical systems, see [L+]. For the related recurrence maxima process given in (2.3) a full classification of the corresponding limits $G(u)$ (relative to a given ψ) is beyond the scope of this paper. We illustrate with the following examples.

Example 2.2. Suppose that $\psi(z) = -\log z$. We choose a_n, b_n so that $\psi(r_n)$ matches $\tau/2n$, with $\tau \equiv \tau(u)$. We take $a_n = 1$ and $b_n = \log(2n)$

giving $\tau(u) = e^{-u}$. This example is consistent with the result of [C, Theorem 4.1].

Example 2.3. Suppose that $\psi(z) = z^{-\alpha}$ with $\alpha > 0$. In this case we obtain $\tau(u) = u^{-\frac{1}{\alpha}}$ in Theorem 2.3 via sequences $a_n = (2n)^{-\alpha}$ and $b_n = 0$.

Thus for scaling sequences r_n defined by $\psi(r_n) = u/a_n + b_n$, the canonical functional forms arising for $\tau(u)$ are consistent with the Type I-III distributions arising in extreme value theory (Gumbel, Fréchet and Weibull). In general, an explicit formula for $G(u)$ is not easily obtained, but corresponds to an *averaged* version of these extreme distribution types. We give an example below where $G(u)$ can be made explicit. We also give an example in Section 3.1.

Example 2.4. Consider the transformation $f(x) = \beta x \bmod 1$, with $\mathcal{X} = [0, 1]$ and $\beta^{-1} = \beta - 1$ (with $\beta \in (1, 2)$). This map is uniformly expanding and Markov. The invariant density has explicit formula

$$\rho(x) = \frac{\beta^3}{\beta^2 + 1} \mathbf{1}_{[0, \beta^{-1})}(x) + \frac{\beta^2}{\beta^2 + 1} \mathbf{1}_{[\beta^{-1}, 1]}(x),$$

and is piecewise constant. We therefore obtain for each $k \geq 0$

$$\mu(R_n(r_n, x) = k) \xrightarrow{n \rightarrow \infty} \frac{1}{k!} \left\{ \frac{1}{\beta} \left(\frac{\beta^3 \tau}{\beta^2 + 1} \right)^k e^{-\tau \frac{\beta^3}{\beta^2 + 1}} + \frac{1}{\beta^2} \left(\frac{\beta^2 \tau}{\beta^2 + 1} \right)^k e^{-\tau \frac{\beta^2}{\beta^2 + 1}} \right\}.$$

Hence, once $\tau(u)$ is specified as in the previous examples (e.g. via choice of ψ and appropriate scaling sequences a_n, b_n) then we obtain $G(u)$ accordingly.

3. DYNAMICAL SYSTEMS WITH ASYMPTOTIC HYPERBOLIC RECURRENCE

In this section we consider dynamical systems that do not satisfy Definition 2.2. A scenario that can arise includes the case where (A1) fails to hold, such as the system having sub-exponential decay of correlations. Similarly, the assumption that the invariant density $\rho(x)$ is of bounded variation type is also restrictive. However, the sources that lead to (say) sub-exponential decay of correlations or lack of regularity of ρ maybe due to local phenomena, such as the map having a neutral fixed point, and/or having a critical point. This allows us to sometimes approximate those maps using dynamical systems that are hyperbolic recurrent. This leads to the following definition.

Definition 3.1. We say that (\mathcal{X}, f, μ) is *asymptotically hyperbolic* if for each $\varepsilon > 0$ we may remove some set $U_\varepsilon = \cup_i U_{\varepsilon, i}$ for some connected sets $U_{\varepsilon, i}$ such that

- (i) $\mu(U_\varepsilon) \leq \varepsilon$
- (ii) for $Y_\varepsilon := \mathcal{X} \setminus U_\varepsilon$ and $\tau = \tau_\varepsilon$ the first return time to Y_ε , the system $(Y_\varepsilon, F_\varepsilon = f^{\tau_\varepsilon}, \mu_\varepsilon)$ is hyperbolic recurrent for $\mu_\varepsilon = \mu|_{Y_\varepsilon}/\mu(Y_\varepsilon)$.

We recall that μ is an acip, so that it is implicit in the definition of asymptotically hyperbolic that $\mathbb{E}_{\mu_\varepsilon}(\tau_\varepsilon) < \infty$. Then we may apply Theorem 2.3 to $(Y_\varepsilon, F_\varepsilon = f^{\tau_\varepsilon}, \mu_\varepsilon)$ to prove the following.

Theorem 3.2. *If (\mathcal{X}, f, μ) is asymptotically hyperbolic then the conclusions of Theorem 2.3 and Theorem 2.7 hold for (\mathcal{X}, f, μ) .*

We prove Theorem 3.2 in Section 6. In the next section we illustrate with examples for which Theorem 3.2 applies.

3.1. Dynamical systems exhibiting asymptotic hyperbolic recurrence. Given $\gamma \in (0, 1)$, define $f = f_\gamma : [0, 1] \rightarrow [0, 1]$ by

$$f_\gamma(x) := \begin{cases} x(1 + 2^\gamma x^\gamma) & \text{if } x \in [0, 1/2), \\ 2x - 1 & \text{if } x \in [1/2, 1]. \end{cases}$$

This is a standard form of Manneville-Pomeau map and it has an acip $\mu = \mu_\gamma$. Note that this is not hyperbolic recurrent, for example the density is not in BV.

Now define a Misiurewicz-Thurston unimodal map as follows. Suppose that $f : [0, 1] \rightarrow [0, 1]$ is C^2 , has negative Schwarzian derivative and has a unique critical point c , which is moreover non-flat¹. Furthermore we assume that there is some $n \geq 2$ such that $f^n(c)$ is a repelling periodic point. For details of these maps see for example [DT], a basic example being $x \mapsto 4x(1 - x)$. Again there is an acip μ , but the system is not hyperbolic recurrent, the density is not in BV.

Proposition 3.3. *If $([0, 1], f, \mu)$ is a Manneville-Pomeau map with μ an acip, or is a Misiurewicz-Thurston unimodal map, with μ an acip, then it is asymptotically hyperbolic.*

Proof. In the Manneville-Pomeau case, we can ensure that $(Y_\varepsilon, F_\varepsilon, \mu_\varepsilon)$ is Markov by removing some $[0, x_n]$ for x_n such that $f^n(x_n) = 1/2$ and $f^j(x_n) \in [0, 1/2]$ for $j = 0, \dots, n - 1$. Note that $\mu([0, x_n]) \rightarrow 0$ as $n \rightarrow \infty$. In this case the system $(Y_\varepsilon, F_\varepsilon, \mu_\varepsilon)$ is Gibbs-Markov and is hence hyperbolic recurrent.

In the Misiurewicz-Thurston case we may remove an ε -neighbourhood of the orbit of $f(c)$. This can be done in such a way that the resulting system $(Y_\varepsilon, F_\varepsilon, \mu_\varepsilon)$ is Markov by exploiting the Markov structure of the original system, see for example [DT, Section 2.1]. Again this system is then Gibbs-Markov and hence hyperbolic recurrent. \square

¹In fact we can also handle systems with flat critical points as in [Z], but we omit the details here.

The intermittent cusp map. We consider the dynamical system (\mathcal{X}, f, μ) given by $f(x) = 1 - 2\sqrt{|x|}$ for $\mathcal{X} = [-1, 1]$. This map has a neutral fixed point at $x = 0$ (quadratic tangent with diagonal), and a derivative singularity at $x = 0$. In particular we have $|f'(x)| = |x|^{-1/2}$. However, the statistical properties are well understood and this map fits the definition of being asymptotically hyperbolic, see [CHMV]. The relevant inducing set can be taken to be $Y_\epsilon = [-1 + \epsilon, 0) \cup (0, 1 - \epsilon]$. The map has an ergodic invariant (probability) density function given by $\rho(x) = \frac{1}{2}(1 - x)$. Hence, relative (say) to the Manneville-Pomeau map we can make the limit distribution function for $\mu(R_n(r_n, x) = k)$ explicit, and also determine more precisely the almost sure rates governing the growth of $\min_{j \leq n} X_j(x)$.

To determine the limit distribution, we can apply Theorem 3.2 to deduce that along $r_n = \tau/2n$ we obtain

$$\lim_{n \rightarrow \infty} \mu(R(x, r_n) = k) = \int_{-1}^1 \frac{\tau^k}{k!} \left(\frac{1-x}{2} \right)^{k+1} e^{-\frac{\tau}{2}(1-x)} dx.$$

The right-hand integral can be evaluated easily to give the limit distribution $G(\tau)$ given by

$$G(\tau) = 2(k+1) \left\{ \frac{1}{\tau^2} - e^{-\tau} \sum_{j=0}^{k+1} \frac{\tau^{k-j-1}}{(k+1-j)!} \right\}.$$

In particular when $k = 0$, we obtain

$$G(\tau) = \frac{2}{\tau^2} - 2e^{-\tau}(\tau^{-1} + \tau^{-2}).$$

Notice that for large τ , the limit distribution is heavy tailed (power law) in τ , rather than of exponential type. We can contrast this limit distribution to that of Example 2.4. The mechanism for the heavy tail in this example is due to the density $\rho(x)$ taking a value of zero (at $x = 1$).

For the almost sure bounds on $\min_{j \leq n} X_j(x)$, we apply case (a)(ii) of Theorem 2.7 by taking $\gamma_0 = 1/2$ in the summability criteria of equation (2.2). This gives for all $\eta > 0$,

$$\mu \left(\min_{j \leq n} X_j(x) < \frac{(\log n)^{\frac{1}{2} + \eta}}{n} \text{ eventually} \right) = 1.$$

For the lower bound, we can apply the summability criteria of case (b). Since $\rho(x)$ is bounded, an almost sure lower bound is given by a sequence r_n with $\sum_n r_n < \infty$. However, if we follow the proof of Theorem 3.2, then given $\epsilon > 0$ and inducing set Y_ϵ we can find $c_0 \equiv c_\epsilon$ so that case (a)(i) holds for μ -almost every $x \in Y_\epsilon$ (along the sequence $r_n = c \log \log n / n$ with $c > c_0$). In particular as $\epsilon \rightarrow 0$, $\mu(Y_\epsilon) \rightarrow 1$, but the lower bound constant c_0 then tends to infinity.

4. PROOF OF THEOREM 2.3

The idea of proof is to first remove dependence on the initial condition x . This will be achieved using decay of correlations. The second step of the proof then involves study of a hitting counting function of the form $N_n(\Delta(\zeta_k), x)$, where $\Delta(\zeta_k)$ is a small (shrinking) interval centred on ζ_k . The points ζ_k are uniformly spaced sequence in \mathcal{X} . As Δ shrinks (with increasing n), the number of such points ζ_k then increases. We are then led to average $N_n(\Delta(\zeta_k), x)$ over ζ_k . This averaging explains the integral arising in equation (2.1).

We remark that the case $\{R_n = 0\}$ matches the conclusion of [C, Theorem 4.1] in the case $\psi(z) = -\log z$. There, an *averaged* exponential law distribution is obtained. For the case $\{N_n^\psi = k\}$ a best approach here seems a variant of the *Chen-Stein method* as applied in [CC, Theorem 2.1]. Compared to the aforementioned references, the setting here is simpler since our measure admits a density $\rho(x)$ which is bounded. Moreover, decay of correlations estimates for relevant indicator functions (appearing in the proofs) have bounded BV and L^1 norms. This avoids the need to approximate indicator functions by Lipschitz (or Hölder) functions. Our proofs would still go through in these cases, but with additional minor modifications.

Notation. Regarding notation, for functions $g(x), h(x)$ defined either over \mathbb{R} or \mathbb{N} , we use the convention $f \lesssim g$ if there exists a constant $C > 0$ such that $f(x) \leq Cg(x)$ as $x \rightarrow \infty$. Equivalently we sometimes use the notation $f(x) = O(g(x))$. We say that $f \approx g$ if there exists a constant $C > 0$ such that $C^{-1} \leq f(x)/g(x) \leq C$ for all sufficiently large x . Unless stated otherwise, we use C_1, C_2, \dots to denote generic constants which do not depend on x .

4.1. Removing dependence on the initial condition. We proceed along similar lines to the proof of [C, Theorem 4.1]. Suppose \mathcal{X} is the interval $[0, 1]$, and set up a partition \mathcal{U}_m of \mathcal{X} into equally sized intervals $\Delta_k \equiv \Delta(\zeta_k)$, with $1 \leq k \leq m$, and ζ_k the centre of Δ_k . Thus each interval Δ_k has size $1/m$. We let $\delta = \frac{1}{2}|\Delta|$, where Δ is any element of \mathcal{U}_m . We let ζ denote the centre of Δ . Given Δ , and $r \geq \delta$ we introduce sets Δ^\pm and Δ_0 given by:

$$\begin{aligned}\Delta^\pm &:= \{x \in \mathcal{X} : d(x, \Delta) \leq r \pm \delta\}, \\ \Delta_0 &:= \{x \in \mathcal{X} : d(x, \Delta) \leq r\},\end{aligned}$$

see Figure 1 for a schematic picture of this.

In the later steps of the proof of Theorem 2.3, we take $r = \frac{\tau}{2n}$, and $\frac{1}{m} \ll r$, but at the moment we do not specify the dependence of r and m on n . Writing $N_{t,n}(A, x) := \#\{j \in [t, n] : f^j(x) \in A\}$, the following result compares $\{R_n(r, x) = k\}$ with $\{N_{1,n}(\Delta_0, x) = k\}$.

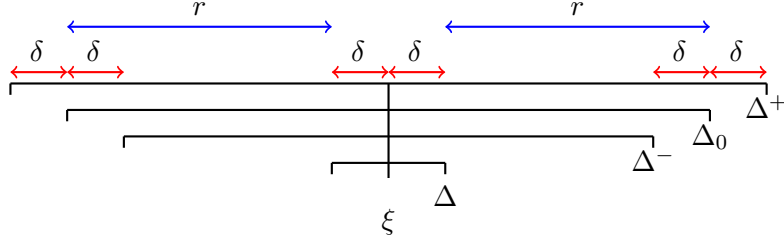


FIGURE 1. Schematic picture of $\Delta, \Delta^-, \Delta_0, \Delta^+$. We will usually take $r = \frac{\tau}{2n}$ and $\delta = \frac{1}{2m}$ for $n \ll m$, for example, later in the paper we use $m = n^{1+\beta'}$ below.

Proposition 4.1. *The following inequality holds:*

$$\left| \mu(R_n(r, x) = k) - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta \cap [N_{1,n}(\Delta_0, x) = k]) \right| \leq \mu(E_{g(n)}(r + 2\delta)) + C_1 n \delta + C_2 \delta^{\frac{1}{p}-1} \Theta(g(n)).$$

where C_1 and C_2 are positive constants.

Proof. Given \mathcal{U}_m , we have the following decomposition of $\{R_n(r, x) = k\}$ into a (disjoint) union of sets:

$$\{x \in \mathcal{X} : R_n(r, x) = k\} = \bigcup_{\Delta \in \mathcal{U}_m} \{\Delta \cap [R_n(r, x) = k]\}.$$

For each element of this union we have the following lemma.

Lemma 4.2. *In the notation above:*

$$\{\Delta \cap [R_n(r, x) = k]\} \ominus \{\Delta \cap [N_{1,n}(\Delta_0, x) = k]\} \subset \Delta \cap \left(\bigcup_{j=1}^n f^{-j}(\Delta^+ \setminus \Delta^-) \right).$$

Here \ominus denotes symmetric difference.

Proof. The proof of the lemma uses the definitions of Δ^\pm and Δ_0 . In particular, given $j \geq 1$, suppose that $x \in \Delta$ and $f^j(x) \notin \Delta^+$. Then clearly $f^j(x) \notin \Delta_0$. Furthermore we have

$$d(f^j x, x) \geq |d(f^j x, \zeta) - d(\zeta, x)| \geq r.$$

Hence, the orbit at time j gives zero count to both $R_n(r, x)$ and $N_{1,n}(\Delta_0, x)$. Similarly, when $f^j(x) \in \Delta^-$, since $r > \delta$ we obtain a unit count contribution to both $R_n(r, x)$ and $N_{1,n}(\Delta_0, x)$. In particular if the whole of the orbit of x avoids $\Delta^+ \setminus \Delta^-$, then the events $\{R_n(r, x) = k\}$ and $\{N_{1,n}(\Delta_0, x) = k\}$ have the same itinerary of times $1 \leq t_1 < t_2 < \dots < t_k \leq n$ of unit count values (while all other times give zero count contributions). This completes the proof of the lemma. \square

To prove Proposition 4.1, we first of all have

$$\begin{aligned} & \left| \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta \cap (R_n(r, x) = k)) - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta \cap (N_{1,n}(\Delta_0, x) = k)) \right| \\ & \leq \sum_{\Delta \in \mathcal{U}_m} \mu \left(\Delta \cap \left(\bigcup_{j=1}^n f^{-j}(\Delta^+ \setminus \Delta^-) \right) \right). \end{aligned} \quad (4.1)$$

On the right-hand side we make the following decomposition:

$$\begin{aligned} \bigcup_{j=1}^n f^{-j}(\Delta^+ \setminus \Delta^-) &= \left(\bigcup_{j=1}^{g(n)} f^{-j}(\Delta^+ \setminus \Delta^-) \right) \cup \left(\bigcup_{j=g(n)+1}^n f^{-j}(\Delta^+ \setminus \Delta^-) \right) \\ &:= \Delta' \cup \Delta''. \end{aligned}$$

Hence, the right-hand side of (4.1) becomes bounded by

$$\sum_{\Delta \in \mathcal{U}_m} (\mu(\Delta \cap \Delta') + \mu(\Delta \cap \Delta'')).$$

By definition of $E_{g(n)}$ we see that

$$\mu(\Delta \cap \Delta') \leq \mu(E_{g(n)}(r + 2\delta)),$$

while by decay of correlations (of BV against L^p) we have

$$\begin{aligned} \mu(\Delta \cap \Delta'') &\leq \sum_{\Delta \in \mathcal{U}_m} \sum_{j=g(n)+1}^n (\mu(\Delta) \mu(\Delta^+ \setminus \Delta^-) + \Theta(j) \|1_\Delta\|_{\text{BV}} \|1_{\Delta^+ \setminus \Delta^-}\|_p), \\ &\leq \sum_{\Delta \in \mathcal{U}_m} n \mu(\Delta) \mu(\Delta^+ \setminus \Delta^-) + C \Theta(g(n)) \delta^{\frac{1}{p}} \\ &\leq C_1 n \delta + C_2 \Theta(g(n)) \delta^{\frac{1}{p}-1}, \end{aligned}$$

where in the last step we have summed over $\Delta \in \mathcal{U}_m$, noting that the cardinality of \mathcal{U}_m is δ^{-1} . The constants C_1 and C_2 are positive (independent of n and m). This completes the proof. \square

In the next step, we consider $\mu(\Delta \cap \{N_{1,n}(\Delta_0, x) = k\})$ and compare to $\mu(\Delta) \mu(N_{1,n}(\Delta_0, x) = k)$. We do this again using the set $E_{g(n)}$ and decay of correlations. We have the following lemma

Lemma 4.3. *Given $t > 1$, the following decomposition holds (for constants $C_1, C_2 > 0$)*

$$\begin{aligned} & \left| \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta \cap (N_{1,n}(\Delta_0, x) = k)) - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \mu(N_{1,n}(\Delta_0, x) = k) \right| \\ & \leq C_1 \delta^{-1} \Theta(t) + C_2 \mu(\Delta_0) t + \mu(E_t(r + 2\delta)). \end{aligned}$$

Proof. First observe for $t, k \geq 1$ we have

$$\begin{aligned} \{N_{t,n}(\Delta_0) = k\} \setminus \{\geq 1 \text{ hit to } \Delta_0 \text{ in time } [1, t]\} &\subset \{N_{1,n}(\Delta_0) = k\} \\ &\subset \{N_{t,n}(\Delta_0) = k\} \cup \{\geq 1 \text{ hit to } \Delta_0 \text{ in time } [1, t]\} \\ \text{and for } k = 0 \text{ we have} \\ \{N_{t,n}(\Delta_0) = 0\} \setminus \{N_{1,n}(\Delta_0) = 0\} \\ &= \{0 \text{ hit to } \Delta_0 \text{ in time } [t, n], \text{ at least one in } [0, t]\}. \end{aligned}$$

Hence, putting $t = g(n)$ we have

$$\left| \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta \cap [N_{1,n}(\Delta_0, x) = k]) - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta \cap (N_{t,n}(\Delta_0, x) = k)) \right| \leq \mu(E_{g(n)}(r + \delta)).$$

Using decay of correlations, we have

$$\begin{aligned} |\mu(\Delta \cap (N_{t,n}(\Delta_0, x) = k)) - \mu(\Delta)\mu(N_{1,n-t}(\Delta_0, x) = k)| \\ \leq C\Theta(g(n))\|\mathbf{1}_\Delta\|_{\text{BV}}\|\mathbf{1}_{N_{1,n-t}(\Delta_0, k)}\|_1, \end{aligned}$$

for some $C > 0$. We can also replace $N_{1,n-t}(\Delta_0, x)$ by $N_{1,n}(\Delta_0, x)$. This introduces a further error term which can be quantified using

$$\begin{aligned} \{N_{1,n}(\Delta_0, x) = k\} \setminus \{\geq 1 \text{ hit to } \Delta_0 \text{ in } (n-t, n]\} &\subset \{N_{1,n-t}(\Delta_0, x) = k\} \\ &\subset \{N_{1,n}(\Delta_0, x) = k\} \cup \{\geq 1 \text{ hit to } \Delta_0 \text{ in } (n-t, n]\}. \end{aligned}$$

So the difference in these sets can be estimated by $\mu\left(\bigcup_{j=n-t}^n f^{-j}\Delta_0\right) \leq t\mu(\Delta_0)$. We now sum over $\Delta \in \mathcal{U}_m$ to complete the proof of the lemma. \square

Analysis of the set $N_{1,n}(\Delta_0, k)$ is now equivalent to a hitting time problem associated to the set $\Delta_0 \equiv \Delta_0(\zeta_k)$, with (ζ_k) a uniformly spaced sequence in \mathcal{X} , and with $k \in [1, m]$. Each set Δ_0 corresponds to an interval of (Lebesgue) measure $= 2r$. Moreover the centre points ζ_k are independent of the orbit of x . In the sequel we take $r = \tau/2n$ and choose $m = (2\delta)^{-1}$ accordingly. As we can see from 4.3, some balancing is required on the speed at which we allow $m \rightarrow \infty$ as $n \rightarrow \infty$ ($r \rightarrow 0$). We make this precise in the next section.

4.2. Implementing the Chen-Stein scheme. In this section we consider the process $N_n(\Delta_0, x) := N_{0,n}(\Delta_0, x)$. In conjunction with Proposition 4.1 and Lemma 4.3 we are then able to determine the limit distribution for $R_n(r, x)$ when $r \equiv \frac{\tau}{2n}$. Our analysis will also determine the error terms. The behaviour of these error terms will be useful in the proof of Theorem 2.7. We will estimate the “distance” of $N_n(A, x)$ to a Poisson distribution with parameter $n\mu(A)$. To do this, we define the

Total Variation distance $d_{TV}(U, V)$ between two integer valued random variables U, V by

$$d_{TV}(U, V) := \frac{1}{2} \sum_{k=0}^{\infty} |\mathbb{P}(U = k) - \mathbb{P}(V = k)|.$$

We have the following result, which is a re-statement of [CC, Theorem 2.1] in our context.

Proposition 4.4. *Consider the process $N_n(A, x)$, and $A \subset \mathcal{X}$. Then*

$$d_{TV}(N_n(A), \text{Poi}(n\mu(A))) \leq \mathcal{E}(A, n, p, M),$$

where

$$\mathcal{E}(A, n, p, M) := 2nM\{\mathcal{E}_1(A, n, p) + \mathcal{E}_2(A, p)\} + \mathcal{E}_3(A, n, p, M),$$

and

$$\begin{aligned} \mathcal{E}_1(A, n, p) &:= \sup_{j,q} |\mu(A \cap \{N_{p,n-j}(A) = q\}) - \mu(A)\mu(N_{p,n-j}(A) = q)|, \\ \mathcal{E}_2(A, p) &:= \sum_{j=1}^p \mu(A \cap f^{-j}(A)), \end{aligned}$$

$$\mathcal{E}_3(A, n, p, M) := 4 \left(Mp\mu(A)(1 + n\mu(A)) + \frac{(n\mu(A))^M}{M!} e^{-n\mu(A)} + n\mu(A)^2 \right),$$

where for $\mathcal{E}_1(A, n, p)$ the supremum is over $0 \leq j \leq N - p$, $0 \leq q \leq N - j - p$.

In the subsequent steps, we estimate the terms \mathcal{E}_i (for $i = 1, 2, 3$, suppressing notation where appropriate). The appropriate estimates will be based on our assumptions (A1), (A2). We remark that corresponding terms can be also estimated under different dynamical assumptions on the regularity of the measure and decay of correlations, see for example [HW, PS2]. Using Proposition 4.1 and Lemma 4.3 we then obtain the following result.

Proposition 4.5.

$$\begin{aligned} & \left| \mu\{R_n(r, x) = k\} - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \frac{(n\mu(\Delta_0))^k}{k!} e^{-n\mu(\Delta_0)} \right| \\ & \leq O(1) \left\{ \mu(E_{g(n)}(r + \delta)) + n\delta + \delta^{-1}\Theta(g(n)) + g(n)\mu(\Delta_0) \right\} \quad (4.2) \\ & \quad + \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) (2nM [\mathcal{E}_1(\Delta_0) + \mathcal{E}_2(\Delta_0)] + \mathcal{E}_3(\Delta_0)), \end{aligned}$$

where \mathcal{E}_i are corresponding error terms arising from Proposition 4.4 with $A = \Delta_0$ and appropriately chosen p and M .

We now choose suitable r, t, m, M, p . We choose $M = \log n$ and we put $p = n^\gamma$ for some $\gamma \in (0, 1)$. We take $t = (\log n)^2$ so that we can apply (A2), and we take $m = n^{1+\beta'}$ for some $\beta' > 0$ (so that

that $\delta = \frac{1}{2}n^{-1-\beta'}$). For the value of r , we take $r = \tau/2n$ for $\tau > 0$. From (4.2), we see that there is some interdependence between how fast $n \rightarrow \infty$ and how fast we choose to take $\delta \rightarrow 0$. This is due to the requiring both $n\delta \rightarrow 0$, and $\delta^{-1}\Theta(g(n)) \rightarrow 0$ within the right-hand side of (4.2). For purposes of proving Theorem 2.3, the choice of parameter rates above is appropriate.

Estimation of \mathcal{E}_3 uses information independent of the dynamics except for the regularity of the measure μ . Estimation of \mathcal{E}_1 depends on decay of correlations through (A1), while estimation of \mathcal{E}_2 utilises assumption (A2). We have the following results.

Lemma 4.6. *There exists $C > 0$ and $c_1 > 0$, such that for $M = \log n$ and $p = n^\gamma$ (for $\gamma \in (0, 1)$),*

$$\mathcal{E}_3(\Delta_0, n) \leq Cn^{-c_1}.$$

The constant C depends on τ .

Proof. Using our assumption that the invariant density ρ is of bounded variation, we have $\mu(\Delta_0) \leq C|\Delta_0| \leq C\tau/n$ for some $C > 0$. With our choice $M = \log n$, and $p = n^\gamma$ we obtain

$$\mathcal{E}_3(\Delta_0, n) \leq O(1) \left((\log n)n^{\gamma-1}\tau(1+\tau) + \frac{(C\tau)^{(\log n)}}{(\log n)!}e^{-C\tau} + \frac{C^2\tau^2}{n} \right)$$

The intermediate term on the right can be estimated using Stirling's formula:

$$\begin{aligned} \frac{(C\tau)^{(\log n)}}{(\log n)!} &\approx (C\tau)^{\log n} (\log n)^{-\log n} e^{\log n} (\log n)^{-\frac{1}{2}} \\ &\leq O(1) \exp \left\{ \left(\log(C\tau) - \frac{1}{2} \right) \log n - (\log n)(\log \log n) \right\}, \end{aligned}$$

which tends to zero at a super-polynomial speed. Hence the dominant term for \mathcal{E}_3 is the quantity of order $n^{\gamma-1} \log n$. The conclusion of the Lemma now follows. \square

Remark 4.7. For the proof of Theorem 2.7 (in Section 5) we will need to quantify this error term in case τ depends on n , with $\tau \rightarrow \infty$ permitted. By inspecting the terms within the proof Lemma 4.6 we see that the same conclusion holds for the lemma provided $\tau \equiv \tau_n = o((\log n))$.

We now consider \mathcal{E}_1 .

Lemma 4.8. *There exists $C > 0$ such that for $p = n^\gamma$ (and $\gamma \in (0, 1)$),*

$$\mathcal{E}_1(\Delta_0, n) = o(n^{-a}).$$

for all $a > 0$, (and hence superpolynomial decay).

Proof. For \mathcal{E}_1 , we just use decay of correlations, and obtain

$$\mathcal{E}_1(\Delta_0, n) \leq \Theta(p) \|1_A\|_{\text{BV}} \|1_{(N_{p,n-j}(A)=q)}\|_\infty \leq 2\Theta(n^\gamma).$$

This error term clearly goes to zero at a super-polynomial speed by assumption on exponential decay for the rate function Θ . \square

Remark 4.9. The error terms \mathcal{E}_1 and \mathcal{E}_3 do not depend on Δ , and hence can be taken uniform over $\Delta \in \mathcal{U}_m$.

Finally we consider \mathcal{E}_2 . We have the following lemma.

Lemma 4.10. *There exists $C > 0$, and $c_2 > 0$ such that for $p = n^\gamma$ (and $\gamma \in (0, 1)$) such that*

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \mathcal{E}_2(\Delta_0, p) = O(n^{-c_2-1}),$$

with implied constant within $O(\cdot)$ depending on τ .

Proof. For \mathcal{E}_2 we can write

$$\mathcal{E}_2(A, p) = \sum_{j=1}^t \mu(A \cap f^{-j}(A)) + \sum_{j=t+1}^p \mu(A \cap f^{-j}(A)), \quad (4.3)$$

where we put $t = (\log n)^2$, and note that $A = \Delta_0$ is a small interval of measure $O(\tau/2n)$, with a centre point ζ_k . Recall ζ_k are uniformly spaced points in \mathcal{X} , with $1 \leq k \leq m$. In (4.3), we have for the second term on the right hand side:

$$\begin{aligned} \sum_{j=t+1}^p \mu(A \cap f^{-j}(A)) &\leq \sum_{j=t+1}^p (\Theta(j) \|1_A\|_{\text{BV}} \|1_A\|_\infty + \mu(A)^2) \\ &\leq n^\gamma \Theta((\log n)^2) + n^\gamma (\tau/2n)^2 = O(n^{-1-c'}), \end{aligned}$$

for some $c' > 0$, where we have used (A1). We note that the implied constant in $O(\cdot)$ depends on τ . This term causes no complication, and its estimation does not depend on ζ_k .

Now the first term on the right hand side of (4.3) depends on ζ_k . Indeed if ζ_k is a periodic point then this error term will not be small. Using equation (4.2) as a guide, it is sufficient to estimate instead

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{j=1}^t \mu(\Delta_0 \cap f^{-j} \Delta_0). \quad (4.4)$$

We'll bound the sum directly using condition (A2). By assumption $|\Delta_0| \sim \tau/n$. Hence, letting ζ_k denote the centre of Δ , for $x \in \Delta_0 \cap f^{-j} \Delta_0$ we have for each j ,

$$d(x, f^j x) \leq d(x, \zeta_k) + d(f^j x, \zeta_k) < \frac{2\tau}{n}.$$

Therefore for $a < 1$ we have $x \in \Delta_0 \cap E_j(n^{-a})$, for all n sufficiently large so that $2\tau < n^{1-a}$. For $b > 0$, let

$$\mathcal{F}_n^+ := \{\Delta \in \mathcal{U}_m : \mu(\Delta_0 \cap \tilde{E}_g(n^{-a})) \geq n^{-1-b}\},$$

with $g(n) = a^2(\log n)^2$. Since the density $\rho(x)$ lies in BV we have $\mu(\Delta) \leq C|\Delta| = C/m$ for some $C > 0$ and independent of Δ . Thus (4.4) can be bounded by

$$\sum_{\Delta \notin \mathcal{F}_n^+} \frac{C}{m} \sum_{j=1}^t \mu(\Delta_0 \cap f^{-j} \Delta_0) + \sum_{\Delta \in \mathcal{F}_n^+} \frac{C}{m} \sum_{j=1}^t \mu(\Delta_0 \cap f^{-j} \Delta_0). \quad (4.5)$$

We can bound the first term by $\frac{C}{m} (\sharp(\mathcal{U}_m \setminus \mathcal{F}_n^+)) n^{-1-b} t \leq C a^2 (\log n)^2 n^{-1-b}$. For the second sum in (4.5), observe that for β_1 from (A2) we have

$$\begin{aligned} C n^{-a\beta_1} &\geq \mu(\tilde{E}_g(n^{-a})) = \mu \left(\bigcup_{\Delta} \left\{ x \in \Delta : \exists j \leq g(n^a) : d(x, f^j x) \leq \frac{1}{n^a} \right\} \right) \\ &\geq C' \frac{n}{2m\tau} \sum_{\Delta \in \mathcal{F}_n^+} \mu(\Delta_0 \cap \tilde{E}_g(n^{-a})) \geq C' \frac{n}{2m\tau} n^{-1-b} \# \mathcal{F}_n^+, \end{aligned} \quad (4.6)$$

for some $C' > 0$. Obtaining the second line above can be explained as follows. We can cover \mathcal{X} using all $\Delta \equiv \Delta(\zeta_k)$ in \mathcal{F}_n^+ , with uniformly spaced centres ζ_k , $k \in [1, m]$. For each $\Delta \in \mathcal{F}_n^+$ there is an expanded set Δ_0 centred on Δ , with $|\Delta_0| = \tau/n$. An *under-cover* of \mathcal{X} (and hence $\tilde{E}_g(n^{-a})$) can be obtained by taking *disjoint* sets of the form $\Delta_0(\zeta_{k_j})$, where the centres $\{\zeta_{k_j}\}$ are uniformly spaced, and $k_{j+1} - k_j = 2 \lfloor \tau m / n \rfloor$. (The additional multiplying factor of 2 is to ensure an under-covering using disjoint sets). For such an (under)-covering we obtain

$$\mu(\tilde{E}_g(n^{-a})) \geq \sum_{k_j \leq m} \mu(\Delta_0(\zeta_{k_j}) \cap \tilde{E}_g(n^{-a})).$$

If we range over all such $\Delta \in \mathcal{F}_n^+$ (i.e. eventually all such ζ_k for $k \in [1, m]$) then we obtain at most $2 \lfloor \tau m / n \rfloor$ such under-covers, and this leads to the estimate

$$2 \left\lfloor \frac{\tau m}{n} \right\rfloor \mu(\tilde{E}_g(n^{-a})) \geq \sum_{\Delta \in \mathcal{F}_n^+} \mu(\Delta_0(\zeta_k) \cap \tilde{E}_g(n^{-a})),$$

and hence to the estimate stated in equation (4.6). We therefore obtain

$$\# \mathcal{F}_n^+ \leq C \tau m n^{b-a\beta_1},$$

for some $C > 0$. Thus for the second term in (4.5) we obtain

$$\begin{aligned} \sum_{\Delta \in \mathcal{F}_n^+} \frac{1}{m} \sum_{j=1}^{g(n)} \mu(\Delta_0 \cap f^{-j} \Delta_0) &\lesssim \sum_{\Delta \in \mathcal{F}_n^+} \frac{1}{m} \sum_{j=1}^{g(n)} |\Delta_0| \leq \# \mathcal{F}_n^+ \frac{g(n)}{m} |\Delta_0| \\ &\lesssim \# \mathcal{F}_n^+ g(n) \frac{|\Delta_0|}{m} \leq C \tau n^{b-a\beta_1} a^2 (\log n)^2 |\Delta_0| \\ &\leq 2C \tau^2 n^{b-1-a\beta_1} a^2 (\log n)^2. \end{aligned}$$

To complete the proof of Lemma 4.10, first choose $a < 1$, and notice that $\Theta(g(n))$ still tends to 0 zero at a superpolynomial speed when adjusting

$g(n)$ from $(\log n)^2$ to $(a \log n)^2$. There is flexibility on choosing $a < 1$, and we can take this parameter arbitrarily close to 1. We then choose $b = \frac{1}{2}a\beta_1$. \square

4.3. Completing the proof of Theorem 2.3. Using Lemmas 4.8-4.6, we obtain

$$\left| \mu\{R_n(n, x) = k\} - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \frac{(n\mu(\Delta_0))^k}{k!} e^{-n\mu(\Delta_0)} \right| \leq Cn^{-\alpha'}, \quad (4.7)$$

where we take $r = \tau/2n$ and $m = n^{1+\beta'}$. In (4.7) we have $C > 0$ and $\alpha' > 0$. We now simplify the sum over \mathcal{U}_m . Lebesgue differentiation gives

$$\begin{aligned} \mu(\Delta) &= |\Delta|\rho(\zeta_k) + o_\Delta(1), \\ \mu(\Delta_0) &= |\Delta_0|\rho(\zeta_k) + o_{\Delta_0}(1), \end{aligned}$$

where for a generic (full measure) centre point ζ_k the quantity $o_A(1)$ satisfies $o_A(1)/|A| \rightarrow 0$ as $|A| \rightarrow 0$ (taking A either as Δ or Δ_0). For density $\rho \in \text{BV}$, and for non-generic centres (such as ρ having a jump discontinuity at ζ_k), the corresponding limits are uniformly bounded from above. Hence to treat the limit of the sum in equation (4.7), define the sequence of functions $g_m(x)$ given by $g_m(x) = \frac{(\tau\rho(\zeta_k))^k}{k!} e^{-\tau\rho(\zeta_k)}$, where $x \in \Delta(\zeta_k)$ and $\Delta(\zeta_k) \in \mathcal{U}_m$. Then for μ almost every $x \in \mathcal{X}$, the functions $g_m(x) \rightarrow g(x)$ pointwise as $m \rightarrow \infty$, with $g(x) = \frac{(\tau\rho(x))^k}{k!} e^{-\tau\rho(x)}$, and $g \in L^1(\text{Leb})$. The Dominated Convergence Theorem gives

$$\lim_{n \rightarrow \infty} \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \frac{(n\mu(\Delta_0))^k}{k!} e^{-n\mu(\Delta_0)} = \int_{\mathcal{X}} \frac{(\tau\rho(x))^k}{k!} e^{-\tau\rho(x)} dx.$$

This completes the proof of Theorem 2.3.

5. PROOF OF THEOREM 2.7

The proof of Theorem 2.7 is as follows. We first treat case (b). In the case that $r_n \rightarrow 0$ is a monotone sequence, an elementary observation is that

$$\mu \left\{ x \in \mathcal{X} : \min_{j \leq n} d(f^j x, x) < r_n, \text{ i.o.} \right\} = \mu \left\{ x \in \mathcal{X} : d(f^j x, x) < r_n, \text{ i.o.} \right\}.$$

(Here i.o. means infinitely often.) For the equation above, clearly the set measured on the right is contained in the set measured on the left. For the other direction see [G, HKKP]. From [KKP2, Theorem C], it can be shown that

$$\sum_n \int_{\mathcal{X}} \mu(B(x, r_n)) d\mu < \infty \implies \mu\{x \in \mathcal{X} : d(f^j x, x) < r_n, \text{ i.o.}\} = 0.$$

Hence, by considering the complement of $\{x \in \mathcal{X} : d(f^j x, x) < r_n, \text{ i.o.}\}$ the result for case (b) immediately follows.

We now turn to case (a). We follow step by step the proof of Theorem 2.3 as applied to the particular case $\{R_n(r, x) = 0\}$. Recalling that $r = \frac{\tau}{2n}$ here we obtain

$$\lim_{n \rightarrow \infty} \mu \left(x \in \mathcal{X} : \min_{k \leq n} d(f^k x, x) > \frac{\tau}{2n} \right) = \int_{\mathcal{X}} \rho(x) e^{-\tau \rho(x)} dx.$$

However, the result above is not enough on its own, as we need to estimate the error rate for the convergence to the limit. To estimate the error rate we consider the cases (a)(i) and (a)(ii) stated in the Theorem 2.7 separately. However, the assumptions relating to either of these cases will only appear in the final steps of the proof. For the sequence $r = \frac{\tau}{2n}$, we take a slower decaying sequence for $r \equiv r_n$ of the form $r = \tau(n)/(2n)$ with $\tau(n) \rightarrow \infty$. In case (a)(i) this is equivalent to taking $\tau = 2c \log \log n$, while in case (a)(ii) we take $\tau = (\log n)^\gamma$. The relevant error rates in the convergence can be obtained from (4.2). We estimate each of the error terms in turn with appropriate choice of parameter sequence. In particular we take $t = (\log n)^2$ for the correlation time-lag, $m = n^{1+\beta'}$ for the cardinality of \mathcal{U}_m . To obtain conclusions similar to those given in Lemmas 4.8-4.6 we again take $p = n^\gamma$ and $M = \log n$. Using again (A1) and (A2), we see that all the terms within the right-hand side of (4.2) all decay at a power law rate in n , thus leading to the same conclusion as given in (4.7). This leads to

$$\left| \mu(R_n(r, x) = 0) - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) e^{-n\mu(\Delta_0)} \right| \leq C n^{-\alpha'} \quad (5.1)$$

The idea now is to use a First Borel-Cantelli argument to show that for a certain subsequence $r \equiv r_{n_k}$, we have $\mu(R_{n_k}(r_{n_k}, x) = 0 \text{ i.o.}) = 0$. This requires an estimate in terms of n for the sum over \mathcal{U}_m . We therefore require assumptions on the regularity of the density $\rho(x)$ for $x \in \Delta_0$. By Lebesgue differentiation, we recall

$$\begin{aligned} \mu(\Delta) &= |\Delta| \rho(\zeta_k) + o_\Delta(1), \\ \mu(\Delta_0) &= |\Delta_0| \rho(\zeta_k) + o_{\Delta_0}(1). \end{aligned}$$

Under the assumptions of the theorem, recall that ρ is uniformly Hölder (and of bounded variation). In the following sequence of estimates, we keep the steps as general as possible, using the Hölder assumptions in the final steps. If $\rho(x)$ is of bounded variation, then

$$o_\Delta(1) \lesssim |\Delta| \text{var}_\Delta(\rho), \quad o_{\Delta_0}(1) \lesssim |\Delta_0| \text{var}_{\Delta_0}(\rho).$$

Potential problems are that the corresponding variations are not asymptotically small with $|\Delta|$ and (resp.) $|\Delta_0|$. For $r > 0$ let

$$G_r(x) := \int_{x-r}^{x+r} \rho(z) dz.$$

Clearly, when $r = \tau/2n$, $G(\zeta_k) = \mu(\Delta_0)$. The following lemma will be useful in the case where we take $\tau \rightarrow \infty$ as $n \rightarrow \infty$, such as in the case $\tau_n = c \log \log n$.

Lemma 5.1. *Suppose that for $\beta' > 0$ we have $m/n \leq n^{-\beta'}$. Furthermore suppose that $\exists \alpha_2 > 0$. Then for all $r > 0$ we have*

$$\left| \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) e^{-n\mu(\Delta_0)} - \int_{\mathcal{X}} \rho(x) e^{-nG_r(x)} dx \right| \leq C n^{-\alpha_2}.$$

Proof. To prove the lemma, first of all we have

$$\begin{aligned} \left| \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) e^{-n\mu(\Delta_0)} - \sum_{k=1}^m |\Delta(\zeta_k)| \rho(\zeta_k) e^{-nG_r(\zeta_k)} \right| \\ \lesssim \sum_{k=1}^m |\Delta_k| \text{var}_{\Delta}(\rho) e^{-nG_r(\zeta_k)}. \end{aligned} \quad (5.2)$$

This follows using (for general Δ), and any $x \in \Delta$,

$$|\mu(\Delta) - \rho(x)| |\Delta| \leq |\rho(x^+) - \rho(x^-)| |\Delta| \leq |\Delta| \text{var}_{\Delta}(\rho).$$

Here, $\rho(x^+), \rho(x^-)$ correspond to max/min values of ρ on Δ . In (5.2) above we recall that $\Delta_k \equiv \Delta(\zeta_k)$. Each Δ_k has uniform size $|\Delta| = 1/m$, and from the fact that $\rho \in \text{BV}$, then there exists $C > 0$ such that $\text{var}(\rho) \leq C$. Using the fact that $e^{-nG_r(\zeta_k)} \leq 1$, it follows that

$$\sum_{k=1}^m |\Delta_k| \text{var}_{\Delta}(\rho) e^{-nG_r(\zeta_k)} \leq |\Delta| \sum_{\Delta \in \mathcal{U}_m} \text{var}_{\Delta}(\rho) < \frac{C}{m}.$$

For $x \in \Delta$, we compare $\rho(\zeta_k) e^{-nG_r(\zeta_k)}$ to $\rho(x) e^{-nG_r(x)}$. This comparison requires only assumptions on the variation of ρ over Δ . For any $\Delta \in \mathcal{U}_m$, and all $x \in \Delta$ we observe that

$$\begin{aligned} \left| \int_{\Delta} \rho(z) e^{-nG_r(z)} dz - \rho(x) e^{-nG_r(x)} |\Delta| \right| \\ \leq |\Delta| \sup_{\Delta}(\rho) \left(e^{-nG_r(x^+)} - e^{-nG_r(x^-)} \right) + |\Delta| \sup_{\Delta} \left(e^{-nG_r(x)} \right) \text{var}_{\Delta}(\rho), \end{aligned} \quad (5.3)$$

where x^{\pm} correspond to values where $G_r(x)$ takes its max/min values in Δ . We estimate the first term on the right-hand side of (5.3). There exist $C_1, C_2 > 0$ such that

$$\sup_{\Delta}(\rho) < C_1, \quad \sup_{\Delta} \left(e^{-nG_r(z)} \right) < C_2.$$

In fact we can take $C_2 = 1$. Since $z \mapsto e^{-nz}$ is differentiable, the Mean Value Theorem implies

$$\left| e^{-nG_r(x^+)} - e^{-nG_r(x^-)} \right| \leq \sup_{\Delta} (ne^{-nG_r(x)}) |G_r(x^+) - G_r(x^-)|.$$

This leads to elementary bounds (assuming without loss of generality that $x^+ > x^-$):

$$\begin{aligned} |G_r(x^+) - G_r(x^-)| &\leq \left| \int_{x^+-r}^{x^++r} \rho(z) dz - \int_{x^--r}^{x^-+r} \rho(z) dz \right| \\ &\leq \int_{x^--r}^{x^++r} \rho(z) dz + \int_{x^--r}^{x^+-r} \rho(z) dz \\ &\leq 2 \sup_{\Delta_0}(\rho) \cdot |x^+ - x^-| \leq C|\Delta|, \end{aligned}$$

since ρ is uniformly bounded on Δ_0 (independently of r). We obtain for all $x \in \Delta$

$$\left| \int_{\Delta} \rho(z) e^{-nG_r(z)} dz - \rho(x) e^{-nG_r(x)} |\Delta| \right| \leq C'n|\Delta|^2 + C''|\Delta| \text{var}_{\Delta}(\rho). \quad (5.4)$$

Combining (5.2) and (5.4), and summing over $\Delta \in \mathcal{U}_m$ we obtain

$$\begin{aligned} \left| \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) e^{-n\mu(\Delta_0)} - \int_{\mathcal{X}} \rho(x) e^{-nG_r(x)} dx \right| \\ \leq \sum_{\Delta \in \mathcal{U}_m} (C'n|\Delta|^2 + C''|\Delta| \text{var}_{\Delta}(\rho)). \end{aligned}$$

Since $\rho \in \text{BV}$ we clearly have $\sum_{\Delta} \text{var}_{\Delta}(\rho) < \text{var}(\rho) < C$. Hence for $m/n \leq n^{-\beta'}$, we obtain the proof of the Lemma. \square

Remark 5.2. In the proof of Lemma 5.1, the dependence on r does not explicitly feature. The error terms just depend on the spacing sizes $|\Delta|$ associated to the distribution of the centre points $\{\zeta_k\}$. In the case where $\inf_{x \in \Delta} \rho(x) = 0$ the estimation of $G_r(x)$ is more delicate, and requires assumptions on how the density function ρ varies on Δ_0 , so that we can control $\text{var}_{\Delta_0} G_r$. Note that for the proof of Lemma 5.1, we did not need to study the variation of G_r over Δ_0 . It was sufficient to control the difference between the maximum and minimum of G over small intervals Δ .

We are now in a position to consider cases (a)(i) and (a)(ii). First consider case (i), and assume that the invariant density is bounded away from zero. Using the fact that $G_r(x) \geq 2r \min_{x \in \mathcal{X}} \rho(x)$, we obtain

$$\int_{\mathcal{X}} \rho(x) e^{-nG_r(x)} dx \leq C_1 \exp\{-2nr \min_{x \in \mathcal{X}} \rho(x)\}.$$

Using Lemma 5.1, (5.1), and going along the sequence $r_n = c \log \log n/n$ we obtain

$$\mu(R_n(r_n, x) = 0) \leq Cn^{-\alpha_2} + C_1 \exp\{-2c(\log \log n) \min_{x \in \mathcal{X}} \rho(x)\}, \quad (5.5)$$

for some $\alpha_2 > 0$. As it stands, the right-hand side of (5.5) is not summable. Hence we now go along the subsequence $n_k = \lfloor a^k \rfloor$ for $a > 1$. This has the property that $n_{k+1}/n_k \rightarrow a$ as $n \rightarrow \infty$. Putting $\rho_m = \min_{x \in \mathcal{X}} \rho(x)$, we obtain

$$\begin{aligned} \mu(R_{n_k}(r_{n_k}, x) = 0) &\leq C_1(\log n_k)^{2c\rho_m} + O(n_k^{-\alpha_2}), \\ &\leq C_1(k \log a)^{2c\rho_m} + O(a^{-k\alpha_2}) \end{aligned}$$

Since $a > 1$, the right-hand side is summable provided $2c\rho_m > 1$. This matches the assumption within Theorem 2.7 for the value of c . Hence, there exists k_0 such that for all $k \geq k_0$ we have

$$d(f^{n_k}x, x) < c \log \log n_k/n_k.$$

Consider $n \in [n_k, n_{k+1}]$ for $k > k_0$. Since $n_{k+1}/n_k \rightarrow a$ we have $n_k > (1 + o(1))n/a$. Therefore, for all sufficiently large n ,

$$d(f^n x, x) < \frac{c \log \log(n/a)}{(n/a)}(1 + o(1)).$$

Since a can be arbitrarily chosen close to 1 the result follows for case (a)(i).

Now consider case (a)(ii), where by assumption $\rho(x)$ is uniformly Hölder. Hence, there exist $\eta \in (0, 1]$, $L > 0$ and $\delta > 0$ such that $\forall x, y$ with $|x - y| < \delta$ we have

$$|\rho(x) - \rho(y)| \leq L|x - y|^\eta.$$

Thus, we have refined asymptotic information on $\mu(\Delta)$ and $\mu(\Delta_0)$ as follows:

$$\begin{aligned} |\mu(\Delta) - |\Delta|\rho(\zeta_k)| &\leq L|\Delta|^{1+\eta}, \\ |\mu(\Delta_0) - |\Delta_0|\rho(\zeta_k)| &\leq L|\Delta_0|^{1+\eta}. \end{aligned}$$

Thus we now take $m = n^{1+\beta'}$, with $\beta' \in (0, 1)$ and $\tau = (\log n)^\gamma$ for some $\gamma > 0$. This choice of τ corresponds to $r = (\log n)^\gamma/2n$. Hence, for all $x \in \mathcal{X}$ we have

$$\left| G_r(x) - \frac{(\log n)^\gamma}{n} \rho(x) \right| \leq L' \frac{(\log n)^{\gamma(1+\eta)}}{n^{1+\eta}},$$

and moreover

$$\left| e^{-nG_r(x)} - e^{-(\log n)^\gamma \rho(x)} \right| \leq \frac{1}{n^{\gamma'}},$$

for some $\gamma' > 0$. Therefore, using Lemma 5.1 and (5.1), we obtain

$$\mu(R_n(r_n, x) = 0) \leq Cn^{-\alpha_2} + \int_{\mathcal{X}} \rho(x) e^{-(\log n)^\gamma \rho(x)} dx. \quad (5.6)$$

In general, the right-hand side of (5.5) is not summable (for example when $\gamma < 1$). To extend to these cases, we can again go along the subsequence $n_k = \lfloor a^k \rfloor$ for $a > 1$. This leads to

$$\mu(R_{n_k}(r_{n_k}, x) = 0) \leq \int_{\mathcal{X}} \rho(x) e^{-(k \log a)^\gamma \rho(x)} dx + O(a^{-k\alpha_2}).$$

Since $a > 1$, the right-hand side is summable if

$$\sum_{k \geq 1} \int_{\mathcal{X}} \rho(x) e^{-(k \log a)^\gamma \rho(x)} dx < \infty.$$

This matches the assumption of case (a)(ii) by taking $\epsilon = (\log a)^\gamma$. The proof is then completed using the same argument as in case (a)(i).

6. PROOF OF THEOREM 3.2

For the distributional limit law, it is sufficient to get the appropriate estimates for $\mu(x \in \mathcal{X} : R_n(r_n, x) \leq k)$ consistent with Theorem 2.3. For almost sure results, we obtain almost-sure upper/lower bounds for $\min_{j \leq n} X_j(x)$ consistent with the cases described in Theorem 2.7.

For $\epsilon > 0$, consider $(Y_\epsilon, F_\epsilon = f^{\tau_\epsilon}, \mu_\epsilon)$ as in the definition of asymptotically hyperbolic. Let $R_n^{F_\epsilon}$ be the counting function for the random variable $Y_i = \text{dist}(F_\epsilon^i x, x)$. Then by assumption Theorem 2.3 applies to this counting function on $(Y_\epsilon, F_\epsilon, \mu_\epsilon)$.

Given $\delta > 0$ and $N \in \mathbb{N}$, define

$$Y_\epsilon^{\delta, N} := \left\{ x \in Y_\epsilon : \left| \frac{\tau_\epsilon^k(x)}{k} - \frac{1}{\mu(Y_\epsilon)} \right| < \delta \text{ for all } k \geq N \right\}.$$

Here, $\tau_\epsilon^k(x) = \sum_{j=0}^{k-1} \tau_\epsilon(F_\epsilon^j x)$. By the ergodic theorem, $\mu_\epsilon(Y_\epsilon^{\delta, N}) \rightarrow 1$ as $N \rightarrow \infty$.

Define $g^\pm(\epsilon, \delta) := \left(\frac{1}{\mu_\epsilon(Y)} \pm \delta \right)$. So given $x \in Y_\epsilon^{\delta, N}$, we can write

$$ng^-(\epsilon, \delta) < \tau_\epsilon^n(x) < ng^+(\epsilon, \delta).$$

So if the f -iteration time of $x \in Y_\epsilon^{\delta, N}$ is n , then the F_ϵ iteration time is in $\left(\frac{n}{g^+(\epsilon, \delta)}, \frac{n}{g^-(\epsilon, \delta)} \right)$. We also note that for every k hits of $x \in Y_\epsilon$ to an r_n -neighbourhood of x by f up to time $\tau_\epsilon^n(x)$, there can only be fewer hits by F_ϵ up to time $n/g^-(\epsilon, \delta)$.

So we deduce that

$$\begin{aligned} \mu(x \in Y_\epsilon^{\delta, N} : R_n(r_n, x) \leq k) &\leq \mu\left(x \in Y_\epsilon^{\delta, N} : R_{n/g^-(\epsilon, \delta)}^{F_\epsilon}(r_n, x) \leq k\right) \\ &\leq \mu\left(x \in Y_\epsilon : R_{n/g^-(\epsilon, \delta)}^{F_\epsilon}(r_n, x) \leq k\right). \end{aligned}$$

Moreover, for every k hits by F_ϵ of $x \in Y_\epsilon$ to an r_n -neighbourhood of x up to time $n/g^+(\epsilon, \delta)$ there can be at most k corresponding hits

by f of x to its r_n -neighbourhood, if we exclude hits at the edge of Y_ε . Taking Y_{ε, r_n} to be Y_ε with a r_n -neighbourhood removed,

$$\begin{aligned} \mu \left(x \in Y_\varepsilon^{\delta, N} : R_{n/g^+(\varepsilon, \delta)}^{F_\varepsilon}(r_n, x) \leq k \right) \\ \leq \mu(Y_\varepsilon \setminus Y_{\varepsilon, r_n}) + \mu \left(x \in Y_{\varepsilon, r_n} \cap Y_\varepsilon^{\delta, N} : R_n(r_n, x) \leq k \right), \end{aligned}$$

so

$$\begin{aligned} \mu \left(x \in Y_\varepsilon : R_{n/g^+(\varepsilon, \delta)}^{F_\varepsilon}(r_n, x) \leq k \right) - \mu(Y_\varepsilon \setminus Y_\varepsilon^{N, \delta}) \\ \leq \mu(Y_\varepsilon \setminus Y_{\varepsilon, r_n}) + \mu \left(x \in Y_\varepsilon^{\delta, N} : R_n(r_n, x) \leq k \right). \end{aligned}$$

For the upper bound above we can write

$$\mu \left(x \in Y_\varepsilon^{\delta, N} : R_n(r_n, x) \leq k \right) \leq \mu(Y_\varepsilon) \mu_\varepsilon \left(x \in Y_\varepsilon : R_{n/g^-(\varepsilon, \delta)}^{F_\varepsilon}(r_n, x) \leq k \right).$$

Since by Theorem 2.3,

$$\begin{aligned} \mu_\varepsilon \left(x \in Y_\varepsilon : R_{n/g^-(\varepsilon, \delta)}^{F_\varepsilon}(r_n, x) = j \right) \\ \rightarrow \int_{Y_\varepsilon} \frac{(\tau g^-(\varepsilon, \delta))^j \rho_\varepsilon(x)^{j+1}}{j!} e^{-\rho_\varepsilon(x) \tau g^-(\varepsilon, \delta)} dx. \end{aligned}$$

for $\rho_\varepsilon = \rho/\mu(Y_\varepsilon)$, the asymptotic estimate becomes

$$\begin{aligned} \mu(x \in \mathcal{X} : R_n(r_n, x) \leq k) \\ \lesssim \mu((Y_\varepsilon^{\delta, N})^c) + \mu(Y_\varepsilon) \sum_{j=0}^k \int_{Y_\varepsilon} \frac{(\tau g^-(\varepsilon, \delta))^j \rho_\varepsilon(x)^{j+1}}{j!} e^{-\rho_\varepsilon(x) \tau g^-(\varepsilon, \delta)} dx. \end{aligned}$$

We can take the first term to 0 by letting N get large. Then the other spoiling factors all tend to 1 as $\varepsilon \rightarrow 0$, so

$$\mu(x \in \mathcal{X} : R_n(r_n, x) \leq k) \leq \sum_{j=0}^k \int_{\mathcal{X}} \frac{\tau^j \rho(x)^{j+1}}{j!} e^{-\rho(x) \tau_\varepsilon} dx.$$

The lower bound follows similarly, thereby completing the proof of Theorem 3.2 relevant for the distributional limit law.

For the corresponding almost-sure results the same ideas apply. We consider each Y_ε , and for the various cases (a)(i), (a)(ii), and (b) in Theorem 2.7 we just replace r_n by the corresponding $r_{n/g^\pm(\varepsilon, \delta)}$, and similarly replace the density $\rho(x)$ by $\rho_\varepsilon(x)$. The corresponding range of c in case (a)(i) depends on having $\inf_{x \in \mathcal{X}} \rho(x) > 0$ (relevant to the limit $\varepsilon = 0$). If this is not the case, then we apply criteria (a)(ii) to the corresponding density. For the lower bound we apply criteria (b) to the density $\rho(x)$. This completes the proof of Theorem 3.2 for almost sure bounds on $\min_{j \leq n} X_j(x)$.

APPENDIX A. TECHNICAL ESTIMATES CONCERNING THE INVARIANT DENSITY

For the proof of Theorem 2.7, case (a)(ii) we can remove the assumption that ρ has to be uniformly Hölder, and allow for a finite number of jump discontinuities in the density, with left/right limits for ρ as points in \mathcal{C} are approached. We let $\mathcal{C} = \{y_1, \dots, y_N\}$ denote the points of discontinuity, and let $\{L_1, \dots, L_N\}$ denote the corresponding set of jump values, i.e., the differences $|\rho(y_i^+) - \rho(y_i^-)|$. We assume that ρ is uniformly Hölder on each $[y_i, y_{i+1}]$. We state the following corollary.

Corollary A.1. *Under the assumptions of Theorem 2.7, suppose instead that ρ has a finite number of points of discontinuity as described via the set \mathcal{C} above. Then statement (a)(ii) remains valid.*

Proof. In the proof of Theorem 2.7, the Hölder assumption on ρ is primarily used to obtain an asymptotic on the function $G_r(x)$ in terms of r , and in particular on the function $e^{-nG_r(x)}$. Suppose now we allow for jumps in the density. Consider $x \in \mathcal{X}$, and suppose $\Delta_0(x) = [-r + x, x + r]$ is such that $\mathcal{C} \cap \Delta_0 = \emptyset$. Then we obtain the bound

$$|G_r(x) - \rho(x) \cdot 2r| \leq Cr^{1+\eta},$$

using the Hölder assumption on ρ away from \mathcal{C} . Otherwise, suppose that r is small enough that Δ_0 contains a single point of \mathcal{C} , (recall that we take $r \equiv r_n \rightarrow 0$). Then we obtain

$$|G_r(x) - \rho(x) \cdot 2r| \leq CrL_i.$$

Considering a worst case scenario where $\rho(x)$ obtains a value of zero on Δ_0 , we obtain

$$|e^{-nG_r(x)} - e^{-2nr_n\rho(x)}| \leq |1 - e^{-CL_i(\log n)^\gamma}|$$

along the sequence $r_n = (\log n)^\gamma/2n$. This estimate is trivially bounded by 1. The measure of the set of $x \in \mathcal{X}$ such that $\Delta_0(x) \cap \mathcal{C} \neq \emptyset$ is bounded by $2r\#\mathcal{C}$. Thus, the overall (extra) contribution given to the error term as appearing in Lemma 5.1 is of the order

$$\lesssim \#\mathcal{C} \frac{(\log n)^\gamma}{n},$$

and hence of power law type bound in n . Thus, the proof of Theorem 2.7 goes through when we allow for a finite number of jumps in the density. \square

Remark A.2. When there is a countable infinite number of jumps in the density then we know that $\sum_i L_i < \infty$ under the assumption of bounded variation. However to apply the summation criteria of equation (2.2) further assumptions are required on the distribution of the jump locations, and their magnitude.

REFERENCES

- [B] M.D. Boshernitzan, *Quantitative recurrence results*, Invent. Math. **113** (1993) 617–631.
- [CC] J. Chazottes, P. Collet, *Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **33** (2013) 49–80.
- [CHMV] G. Cristadoro, N. Haydn, P. Marie, S. Vaienti, *Statistical properties of intermittent maps with unbounded derivative*, Nonlinearity, **23** (2010) 1071–1095.
- [C] P. Collet, *Statistics of closest return for some non-uniformly hyperbolic systems*, Ergodic Theory Dynam. Systems, **21** (2001) 401–420.
- [DT] M. Demers, M. Todd, *A trichotomy for hitting times and escape rates for a class of unimodal maps*, Preprint (arXiv:2309.09624).
- [FFT] J.M. Freitas, A.C. Freitas, M. Todd, *Hitting times and extreme value theory*, Probab. Theory Related Fields **147**(3) (2010) 675–710.
- [FFTV] A.C. Freitas, J.M. Freitas, M. Todd, S. Vaienti, *Rare Events for the Manneville-Pomeau map*, Stoch. Proc. App. **126** (2016) 3463–3479.
- [G] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, John Wiley and Sons, 1978.
- [GKR] A. Ghosh, M. Kirsebom, P. Roy, *Continued fractions, the Chen-Stein method and extreme value theory*, Ergodic Theory Dynam. Systems **41** (2021) 461–470.
- [GNO] C. Gupta, M. Nicol, W. Ott, *A Borel-Cantelli lemma for non-uniformly expanding dynamical systems*, Nonlinearity **23** (2010) 1991–2008.
- [HW] N. Haydn, K. Wasilewska, *Limiting distribution and error terms for the number of visits to ball in non-uniformly hyperbolic dynamical systems*, Discrete Contin. Dyn. Syst. **36** (2016) 2585–2611.
- [HKKP] M.P. Holland, M. Kirsebom, P. Kunde, T. Persson. *Dichotomy results for eventually always hitting time statistics and almost sure growth of extremes*, To appear in Trans. Amer. Math. Soc.
- [HN] M.P. Holland, M. Nicol, *Speed of convergence to an extreme value distribution for non-uniformly hyperbolic dynamical systems*, Stoch. Dyn. **15** (2015) 1550028.
- [HNT1] M.P. Holland, M. Nicol, A. Török, *Extreme value distributions for non-uniformly expanding dynamical systems*, Trans. Amer. Math. Soc. **364** (2012), 661–688.
- [HNT2] M.P. Holland, M. Nicol, A. Török, *Almost sure convergence of maxima for chaotic dynamical systems*, Stochastic Process. Appl. **126** (2016) 3145–3170.
- [KKP1] M. Kirsebom, P. Kunde, T. Persson, *Shrinking targets and eventually always hitting points for interval maps*, Nonlinearity **33** (2020) 892–914.
- [KKP2] M. Kirsebom, P. Kunde, T. Persson, *On shrinking targets and self-returning points*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **Vol. XXIV** (2023) 1499–1535.
- [LY] A. Lasota, J. A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. **186** (1973) 481–488.
- [L+] V. Lucarini, D. Faranda, A. C. M. Freitas, J. M. Freitas, M. P. Holland, T. Kuna, M. Nocol, M. Todd, S. Vaienti. *Extremes and Recurrence in Dynamical Systems*, Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2016.
- [LSV] C. Liverani, B. Saussol, S. Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory Dynam. Systems **19** (1999) 671–685.
- [PS] F. Pène, B. Saussol, *Back to balls in billiards*, Comm. Math. Phys. **293** (2010) 837–866.
- [PS2] F. Pène, B. Saussol, *Poisson law for some non-uniformly hyperbolic dynamical systems with polynomial rate of mixing*, Ergodic Theory Dynam. Systems **36** (2016) 2602–2626.

- [P] T. Persson, *A strong Borel-Cantelli lemma for recurrence*, Studia Math. **268** (2023) 75–89.
- [Y1] L.-S. Young, *Decay of Correlations for Certain Quadratic Maps*, Comm. Math. Phys. **146** (1992) 123–138.
- [Y2] L.-S. Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. Math. **147** (1998) 585–650.
- [Z] R. Zweimüller, *S-unimodal Misiurewicz maps with flat critical points*, Fund. Math. **181** (2004) 1–25.

MATHEMATICS (CEMPS), HARRISON BUILDING (327), NORTH PARK ROAD,
EXETER, EX4 4QF, ENGLAND

Email address: `m.p.holland@exeter.ac.uk`.

URL: `https://empslocal.ex.ac.uk/people/staff/mph204/`

MATHEMATICAL INSTITUTE, UNIVERSITY OF ST ANDREWS, NORTH HAUGH,
ST ANDREWS, KY16 9SS, SCOTLAND

Email address: `m.todd@st-andrews.ac.uk`

URL: `https://mtoddm.github.io`