Probab. Theory Relat. Fields (2010) 147:675–710
DOI 10.1007/s00440-009-0221-y

# Hitting time statistics and extreme value theory

**Ana Cristina Moreira Freitas · Jorge Milhazes Freitas · Mike Todd**

Received: 18 June 2008 / Revised: 6 April 2009 / Published online: 29 April 2009


**Abstract** We consider discrete time dynamical systems and show the link between Hitting Time Statistics (the distribution of the first time points land in asymptotically small sets) and Extreme Value Theory (distribution properties of the partial maximum of stochastic processes). This relation allows to study Hitting Time Statistics with tools from Extreme Value Theory, and vice versa. We apply these results to non-uniformly hyperbolic systems and prove that a multimodal map with an absolutely continuous invariant measure must satisfy the classical extreme value laws (with no extra condition on the speed of mixing, for example). We also give applications of our theory to higher dimensional examples, for which we also obtain classical extreme value laws and exponential hitting time statistics (for balls). We extend these ideas to the subsequent returns to asymptotically small sets, linking the Poisson statistics of both processes.

J. M. Freitas is partially supported by POCI/MAT/61237/2004 and M. Todd is supported by FCT grant SFRH/BPD/26521/2006. All three authors are supported by FCT through CMUP.

A. C. M. Freitas
Faculdade de Economia, Centro de Matemática, Universidade do Porto,
Rua Dr. Roberto Frias, 4200-464 Porto, Portugal
e-mail: amoreira@fep.up.pt

J. M. Freitas (✉) · M. Todd
Centro de Matemática da Universidade do Porto,
Rua do Campo Alegre 687, 4169-007 Porto, Portugal
e-mail: jmfreita@fc.up.pt
URL: http://www.fc.up.pt/pessoas/jmfreita

M. Todd
e-mail: mtodd@fc.up.pt
URL: http://www.fc.up.pt/pessoas/mtodd

**Keywords** Return time statistics · Extreme value theory · Non-uniform hyperbolicity · Interval maps

**Mathematics Subject Classification (2000)** 37A50 · 37C40 · 60G10 · 60G70 · 37B20 · 37D25 · 37E05

## 1 Introduction

In this paper we demonstrate and exploit the link between Extreme Value Laws (EVL) and the laws for the Hitting Time Statistics (HTS) for discrete time non-uniformly hyperbolic dynamical systems with an absolutely continuous invariant measure.

The setting is a discrete time dynamical system $(\mathcal{X}, \mathcal{B}, \mu, f)$, where $\mathcal{X}$ is a $d$-dimensional Riemannian manifold, $\mathcal{B}$ is the Borel $\sigma$-algebra, $f : \mathcal{X} \to \mathcal{X}$ is a measurable map and $\mu$ an $f$-invariant probability measure (for all $A \in \mathcal{B}$ we have $\mu(f^{-1}(A)) = \mu(A)$). We consider a Riemannian metric on $\mathcal{X}$ that we denote by 'dist' and for any $\zeta \in \mathcal{X}$ and $\delta > 0$, we define $B_\delta(\zeta) = \{x \in \mathcal{X} : \mathrm{dist}(x, \zeta) < \delta\}$. Also let Leb denote Lebesgue measure on $\mathcal{X}$ and for every $A \in \mathcal{B}$ we will write $|A| := \mathrm{Leb}(A)$. The measure $\mu$ will be an absolutely continuous invariant probability measure (acip) with density denoted by $\rho = \frac{d\mu}{d\mathrm{Leb}}$. We will denote $\mathbb{R}^+ := (0, \infty)$ and $\mathbb{R}_0^+ := [0, \infty)$.

### 1.1 Extreme value laws

In this context, by EVL we mean the study of the asymptotic distribution of the partial maximum of observable random variables evaluated along the orbits of the system. To be more precise, take an observable $\varphi : \mathcal{X} \to \mathbb{R} \cup \{\pm\infty\}$ achieving a global maximum at $\zeta \in \mathcal{X}$ (we allow $\varphi(\zeta) = +\infty$) and consider the stationary stochastic process $X_0, X_1, \ldots$ given by

$$X_n = \varphi \circ f^n, \quad \text{for each } n \in \mathbb{N}. \tag{1.1}$$

Define the partial maximum

$$M_n := \max\{X_0, \ldots, X_{n-1}\}. \tag{1.2}$$

If $\mu$ is ergodic then Birkhoff's law of large numbers says that $M_n \to \varphi(\zeta)$ almost surely. Similarly to central limit laws for partial sums, we are interested in knowing if there are normalising sequences $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ and $\{b_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ such that

$$\mu\left(\{x : a_n(M_n - b_n) \le y\}\right) = \mu\left(\{x : M_n \le u_n\}\right) \to H(y), \tag{1.3}$$

for some non-degenerate distribution function (d.f.) $H$, as $n \to \infty$. Here

$$u_n := u_n(y) = \frac{y}{a_n} + b_n \tag{1.4}$$

is such that

$$n\mu(X_0 > u_n) \to \tau, \quad \text{as } n \to \infty, \tag{1.5}$$

for some $\tau = \tau(y) \geq 0$ and in fact $H(y) = \tilde{H}(\tau(y))$. (For more details of the existence of such sequences $u_n$, see Lemma 2.1.) When this happens we say that we have an EVL for $M_n$. Note that, clearly, we must have $u_n \to \varphi(\zeta)$, as $n \to \infty$. We refer to an event $\{X_j > u_n\}$ as an *exceedance*, at time $j$, of level $u_n$.

Classical Extreme Value Theory asserts that there are only three types of non-degenerate asymptotic distributions for the maximum of an independent and identically distributed (i.i.d.) sample under linear normalisation. They will be referred to as *classical* EVLs and we denote them by:

Type 1: $EV_1(y) = \mathrm{e}^{-\mathrm{e}^{-y}}$ for $y \in \mathbb{R}$; this is also known as the *Gumbel* extreme value distribution (e.v.d.).

Type 2: $EV_2(y) = \mathrm{e}^{-y^{-\alpha}}$, for $y > 0$, $EV_2(y) = 0$, otherwise, where $\alpha > 0$ is a parameter; this family of d.f.s is known as the *Fréchet* e.v.d.

Type 3: $EV_3(y) = \mathrm{e}^{-(-y)^{\alpha}}$, for $y \leq 0$, $EV_3(y) = 1$, otherwise, where $\alpha > 0$ is a parameter; this family of d.f.s is known as the *Weibull* e.v.d.

The same limit laws apply to stationary stochastic processes, under certain conditions on the dependence structure, which allow the reduction to the independent case. With this in mind, to a given stochastic process $X_0, X_1, \ldots$ we associate an i.i.d. sequence $Y_0, Y_1, \ldots$ whose d.f. is the same as that of $X_0$, and whose partial maximum we define as

$$\hat{M}_n := \max\{Y_0, \ldots, Y_{n-1}\}. \tag{1.6}$$

We want to compare the asymptotic distribution of $M_n$ with that of $\hat{M}_n$, when properly normalised. We recall that because $Y_0, Y_1, \ldots$ are i.i.d., by [31, Theorem 1.5.1], the convergence in (1.5) is equivalent to

$$\mu(\hat{M}_n \leq u_n) = (\mu(X_0 \leq u_n))^n \to \mathrm{e}^{-\tau}, \quad \text{as } n \to \infty. \tag{1.7}$$

Depending on the type of limit law that applies, we have that $\tau = \tau(y)$ is of one of the following three types: $\tau_1(y) = \mathrm{e}^{-y}$ for $y \in \mathbb{R}$, $\tau_2(y) = y^{-\alpha}$ for $y > 0$, and $\tau_3(y) = (-y)^{\alpha}$ for $y \leq 0$.

In the dependent context, the general strategy is to prove that if $X_0, X_1, \ldots$ satisfies some conditions, then the same limit law for $\hat{M}_n$ applies to $M_n$ with the same normalising sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$. Following [31] we refer to these conditions as $D(u_n)$ and $D'(u_n)$, where $u_n$ is the sequence of thresholds appearing in (1.3). Both conditions impose some sort of independence but while $D(u_n)$ acts on the long range, $D'(u_n)$ is a short range requirement.

The original condition $D(u_n)$ from [31], which we will denote by $D_1(u_n)$, is a type of uniform mixing requirement specially adapted to Extreme Value Theory. Let $F_{i_1,\ldots,i_n}(x_1, \ldots, x_n)$ denote the joint d.f. of $X_{i_1}, \ldots, X_{i_n}$, and set $F_{i_1,\ldots,i_n}(u) = F_{i_1,\ldots,i_n}(u, \ldots, u)$.

**Condition** ($D_1(u_n)$) We say that $D_1(u_n)$ holds for the sequence $X_0, X_1, \ldots$ if for any integers $i_1 < \cdots < i_p$ and $j_1 < \cdots < j_k$ for which $j_1 - i_p > m$, and any large $n \in \mathbb{N}$,

$$\left|F_{i_1,\ldots,i_p,j_1,\ldots,j_k}(u_n) - F_{i_1,\ldots,i_p}(u_n)F_{j_1,\ldots,j_k}(u_n)\right| \leq \gamma(n,m),$$

where $\gamma(n, m_n) \xrightarrow[n\to\infty]{} 0$, for some sequence $m_n = o(n)$.

Since usually the information concerning mixing rates of the systems is known through decay of correlations, in [21] we proposed a weaker version, which we will denote by $D_2(u_n)$, which still allows us to relate the distributions of $\hat{M}_n$ and $M_n$. The advantage is that it follows immediately from sufficiently fast decay of correlations for observables which are of bounded variation or Hölder continuous (see [21, Section 2] and Lemma 6.1).

**Condition** ($D_2(u_n)$) We say that $D_2(u_n)$ holds for the sequence $X_0, X_1, \ldots$ if for any integers $\ell$, $t$ and $n$

$$\begin{aligned}|\mu\left(\{X_0 > u_n\} \cap \{\max\{X_t, \ldots, X_{t+\ell-1}\} \leq u_n\}\right) - \mu(\{X_0 > u_n\}) \\ \times \mu(\{M_\ell \leq u_n\})| \leq \gamma(n,t),\end{aligned}$$

where $\gamma(n, t)$ is nonincreasing in $t$ for each $n$ and $n\gamma(n, t_n) \to 0$ as $n \to \infty$ for some sequence $t_n = o(n)$.

By (1.5), the sequence $u_n$ is such that the average number of exceedances in the time interval $\{0, \ldots, \lfloor n/k \rfloor\}$ is approximately $\tau/k$, which goes to zero as $k \to \infty$. However, the exceedances may have a tendency to be concentrated in the time period following the first exceedance at time 0. To avoid this we introduce:

**Condition** ($D'(u_n)$) We say that $D'(u_n)$ holds for the sequence $X_0, X_1, \ldots$ if

$$\lim_{k\to\infty} \limsup_{n\to\infty} n \sum_{j=1}^{\lfloor n/k \rfloor} \mu(\{X_0 > u_n\} \cap \{X_j > u_n\}) = 0. \tag{1.8}$$

This guarantees that the exceedances should appear scattered through the time period $\{0, \ldots, n-1\}$.

The main result in [21, Theorem 1] states that if $D_2(u_n)$ and $D'(u_n)$ hold for the process $X_0, X_1, \ldots$ and for a sequence of levels satisfying (1.5), then the following limits exist, and

$$\lim_{n\to\infty} \mu(\hat{M}_n \leq u_n) = \lim_{n\to\infty} \mu(M_n \leq u_n). \tag{1.9}$$

The above statement remains true if we replace $D_2(u_n)$ by $D_1(u_n)$ (see [31, Theorem 3.5.2]).

We assume that the observable $\varphi : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ is of the form

$$\varphi(x) = g(\text{dist}(x, \zeta)), \tag{1.10}$$

where $\zeta$ is a chosen point in the phase space $\mathcal{X}$ and the function $g : [0, +\infty) \to \mathbb{R} \cup \{+\infty\}$ is such that 0 is a global maximum ($g(0)$ may be $+\infty$); $g$ is a strictly decreasing bijection $g : V \to W$ in a neighbourhood $V$ of 0; and has one of the following three types of behaviour:

Type 1: there exists some strictly positive function $p : W \to \mathbb{R}$ such that for all $y \in \mathbb{R}$

$$\lim_{s \to g_1(0)} \frac{g_1^{-1}(s + yp(s))}{g_1^{-1}(s)} = \mathrm{e}^{-y}; \tag{1.11}$$

Type 2: $g_2(0) = +\infty$ and there exists $\beta > 0$ such that for all $y > 0$

$$\lim_{s \to +\infty} \frac{g_2^{-1}(sy)}{g_2^{-1}(s)} = y^{-\beta}; \tag{1.12}$$

Type 3: $g_3(0) = D < +\infty$ and there exists $\gamma > 0$ such that for all $y > 0$

$$\lim_{s \to 0} \frac{g_3^{-1}(D - sy)}{g_3^{-1}(D - s)} = y^{\gamma}. \tag{1.13}$$

Examples of each one of the three types are as follows: $g_1(x) = -\log x$ (in this case (1.11) is easily verified with $p \equiv 1$), $g_2(x) = x^{-1/\alpha}$ for some $\alpha > 0$ (condition (1.12) is verified with $\beta = \alpha$) and $g_3(x) = D - x^{1/\alpha}$ for some $D \in \mathbb{R}$ and $\alpha > 0$ (condition (1.13) is verified with $\gamma = \alpha$).

*Remark 1* Let the d.f. $F$ be given by $F(u) = \mu(X_0 \leq u)$ and set $u_F = \sup\{y : F(y) < 1\}$. Observe that if at time $j \in \mathbb{N}$ we have an exceedance of the level $u$ (sufficiently large), i.e., $X_j(x) > u$, then we have an entrance of the orbit of $x$ into the ball $B_{g^{-1}(u)}(\zeta)$ of radius $g^{-1}(u)$ around $\zeta$, at time $j$. This means that the behaviour of the tail of $F$, i.e., the behaviour of $1 - F(u)$ as $u \to u_F$ is determined by $g^{-1}$, if we assume that Lebesgue's Differentiation Theorem holds for $\zeta$, since in that case $1 - F(u) \sim \rho(\zeta)|B_{g^{-1}(u)}(\zeta)|$, where $\rho(\zeta) = \frac{d\mu}{d\mathrm{Leb}}(\zeta)$. From classical Extreme Value Theory we know that the behaviour of the tail determines the limit law for partial maximums of i.i.d. sequences and vice-versa. The above conditions are just the translation in terms of the shape of $g^{-1}$, of the sufficient and necessary conditions on the tail of $F$ of [31, Theorem 1.6.2], in order to exist a non-degenerate limit distribution for $\hat{M}_n$. In fact, if some $EV_i$ applies to $\hat{M}_n$, for some $i \in \{1, 2, 3\}$, then $g$ must be of type $g_i$.

As can be seen from the definitions of $D_2(u_n)$ and $D'(u_n)$, proving EVLs for absolutely continuous invariant measures for uniformly expanding dynamical systems is straightforward. The study of EVLs for non-uniformly hyperbolic dynamical systems were previously addressed in the papers [14,20].

In [14], Collet considered non-uniformly hyperbolic $C^2$ maps of the interval which admit an acip $\mu$, with exponential decay of correlations and obtained a Gumbel EVL for observables of type $g_1$ (actually he took $g_1(x) = -\log x$), achieving a global maximum at $\mu$-a.e. $\zeta$ in the phase space. We remark that neither the critical points nor its orbits were included in this full $\mu$-measure set of points $\zeta$.

In [20] the quadratic maps $f_a(x) = 1 - ax^2$ on $I = [-1, 1]$ were considered, with $a \in \mathcal{BC}$, where $\mathcal{BC}$ is the Benedicks-Carleson parameter set introduced in [6]. For each map $f_a$ with $a \in \mathcal{BC}$, a Weibull EVL was obtained for observables of type $g_3$ achieving a maximum either at the critical point or at the critical value.

## 1.2 Hitting time statistics

We next turn to Hitting Time Statistics for the dynamical system $(\mathcal{X}, \mathcal{B}, f, \mu)$. For a set $A \subset \mathcal{X}$ we let $r_A(y)$ denote the *first hitting time to A* of the point $y$. That is, the first time $j \geq 1$ so that $f^j(y) \in A$. We will be interested in the fluctuations of this function as the set $A$ shrinks. Firstly we consider the Return Time Statistics (RTS) of this system. Let $\mu_A$ denote the conditional measure on $A$, i.e., $\mu_A := \frac{\mu|_A}{\mu(A)}$. By Kac's Lemma, the expected value of $r_A$ with respect to $\mu$ is $\int_A r_A \, d\mu_A = 1/\mu(A)$. So in studying the fluctuations of $r_A$ on $A$, the relevant normalising factor is $1/\mu(A)$. Given a sequence of sets $\{U_n\}_{n\in\mathbb{N}}$ so that $\mu(U_n) \to 0$, the system has *Return Time Statistics* $G(t)$ for $\{U_n\}_{n\in\mathbb{N}}$ if for all $t \geq 0$ the following limit exists and equals $G(t)$:

$$\lim_{n\to\infty} \mu_{U_n}\left(r_{U_n} \geq \frac{t}{\mu(U_n)}\right). \tag{1.14}$$

We say that $(\mathcal{X}, f, \mu)$ has *Return Time Statistics* $G(t)$ *to balls at* $\zeta$ if for any sequence $\{\delta_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ such that $\delta_n \to 0$ as $n \to \infty$ we have RTS $G(t)$ for $U_n = B_{\delta_n}(\zeta)$.

If we study $r_A$ defined on the whole of $\mathcal{X}$, i.e., not simply restricted to $A$, we are studying the Hitting Time Statistics. Note that we will use the same normalising factor $1/\mu(A)$ in this case. Analogously to the above, given a sequence of sets $\{U_n\}_{n\in\mathbb{N}}$ so that $\mu(U_n) \to 0$, the system has *Hitting Time Statistics* $G(t)$ for $\{U_n\}_{n\in\mathbb{N}}$ if for all $t \geq 0$ the following limit is defined and equals $G(t)$:

$$\lim_{n\to\infty} \mu\left(r_{U_n} \geq \frac{t}{\mu(U_n)}\right). \tag{1.15}$$

HTS to balls at a point $\zeta$ is defined analogously to RTS to balls. In [23], it was shown that the limit for the HTS defined in (1.15) exists if and only if the limit for the analogous RTS defined in (1.14) exists. Moreover, they show that the HTS distribution exists and is exponential (i.e., $G(t) = \mathrm{e}^{-t}$) if and only if the RTS distribution exists and is exponential.

For many mixing systems it is known that the HTS are exponential around almost every point. For example, this was shown for Axiom A diffeomorphisms in [25], transitive Markov chains in [33] and uniformly expanding maps of the interval in [13]. Note that in these papers the authors were also interested in the (Poisson) statistics of subsequent returns to some shrinking sets. For various results on some systems with some strong hyperbolicity properties see also e.g. [1,2,10].

Note that for systems with good mixing properties, but at some special (periodic) points, [25] obtained similar distributions for the HTS/RTS with an exponential parameter $0 < \theta < 1$ (i.e., $G(t) = \mathrm{e}^{-\theta t}$); while for the sequence of successive returns to neighbourhoods of these points a 'compound Poisson distribution' was proved in [24].

For non-uniformly hyperbolic systems less is known. A major breakthrough in the study of HTS/RTS for non-uniformly hyperbolic maps was made in [26], where they gave a set of conditions which, when satisfied, imply exponential RTS to cylinders and/or balls. Their principal application was to maps of the interval with an indifferent fixed point. They also provided similar conditions to imply (Poisson) laws for the subsequent visits of points to shrinking sets. (See Sect. 5).

Another important paper in this direction was [7], in which they showed that the RTS for a map are the same as the RTS for the first return map. (The first return map to a set $U \subset \mathcal{X}$ is the map $F = f^{r_U}$.) Since it is often the case that the first return maps for non-uniformly hyperbolic dynamical systems are much better behaved (possibly hyperbolic) than the original system, this provided an extremely useful tool in this theory. For example, they proved that if $f : I \to I$ is a unimodal map for which the critical point is nowhere dense, and for which an acip $\mu$ exists, then the relevant first return systems $(U, F, \mu_U)$ have a 'Rychlik' property. They then showed that such systems, studied in [34], must have exponential RTS, and hence the original system $(I, f, \mu)$ also has exponential RTS (to balls around $\mu$-a.e. point).

The presence of a recurrent critical point means that the first return map itself will not satisfy this Rychlik property. To overcome this problem in [9] special induced maps, $(U, F)$, were used, where for $x \in U$ we have $F(x) = f^{\mathrm{ind}(x)}(x)$ for some inducing time $\mathrm{ind}(x) \in \mathbb{N}$ that is not necessarily the first return time of $x$ to $U$. The fact that these particular maps can be seen as first return maps in the canonical Markov extension, the 'Hofbauer tower', meant that they were still able to exploit the main result of [7] to get exponential RTS around $\mu$-a.e. point for unimodal maps $f : I \to I$ with an acip $\mu$ as long as $f$ satisfies a polynomial growth condition along the critical orbit. In [8] this result was improved to include any multimodal map with an acip, irrespective of the growth along the critical orbits, and of the speed of mixing.

We would like to remark that in the case of partially hyperbolic dynamical systems, [17] proved exponential RTS, using techniques similar to [33]. In fact the theory there also covers the (Poisson) statistics of subsequent returns to shrinking sets of balls. These statistics were also considered for toral automorphisms, using a different method, in [15].

We note that for dynamical systems $(\mathcal{X}, \mathcal{B}, f, \mu)$ where $\mu$ is an equilibrium state, the RTS/HTS to the dynamically defined cylinders are often well understood, see for example [2]. However, for non-uniformly hyperbolic dynamical systems it is not always possible to go from these strong results to the corresponding results for balls. We would like to emphasise that in this paper we focus on the HTS to balls, rather than cylinders.

### 1.3 Main results

Our first main result, which obtains EVLs from HTS, is the following.

**Theorem 1** *Let $(\mathcal{X}, \mathcal{B}, \mu, f)$ be a dynamical system where $\mu$ is an acip, and consider $\zeta \in \mathcal{X}$ for which Lebesgue's Differentiation Theorem holds.*

- *If we have HTS to balls centred on* $\zeta \in \mathcal{X}$, *then we have an EVL for* $M_n$ *which applies to the observables* (1.10) *achieving a maximum at* $\zeta$.
- *If we have exponential HTS* ($G(t) = e^{-t}$) *to balls at* $\zeta \in \mathcal{X}$, *then we have an EVL for* $M_n$ *which coincides with that of* $\hat{M}_n$ (*meaning that* (1.9) *holds*). *In particular, this EVL must be one of the 3 classical types. Moreover, if* $g$ *is of type* $g_i$, *for some* $i \in \{1, 2, 3\}$, *then we have an EVL for* $M_n$ *of type* $EV_i$.

We next define a class of multimodal interval maps $f : I \to I$. We denote the finite set of critical points by Crit. We say that $c \in$ Crit is *non-flat* if there exists a diffeomorphism $\psi_c : \mathbb{R} \to \mathbb{R}$ with $\psi_c(0) = 0$ and $1 < \ell_c < \infty$ such that for $x$ close to $c$, $f(x) = f(c) \pm |\psi_c(x - c)|^{\ell_c}$. The value of $\ell_c$ is known as the *critical order* of $c$. Let

$$NF^k := \left\{ f : I \to I : f \text{ is } C^k, \text{ each } c \in \text{Crit is non-flat and } \inf_{f^n(p)=p} |Df^n(p)| > 1 \right\}.$$

The following is a simple corollary of Theorem 1 and [8, Theorem 3]. It generalises the result of Collet in [14] from unimodal maps with exponential growth on the critical point to multimodal maps where we only need to know that there is an acip.

**Corollary 1** *Suppose that* $f \in NF^2$ *and* $f$ *has an acip* $\mu$. *Then* $(I, f, \mu)$ *has an EVL for* $M_n$ *which coincides with that of* $\hat{M}_n$, *and this holds for* $\mu$*-a.e.* $\zeta \in \mathcal{X}$ *fixed at the choice of the observable in* (1.10). *Moreover, the EVL is of type* $EV_i$ *when the observables are of type* $g_i$, *for each* $i \in \{1, 2, 3\}$.

Now, we state a result in the other direction, i.e., we show how to get HTS from EVLs.

**Theorem 2** *Let* $(\mathcal{X}, \mathcal{B}, \mu, f)$ *be a dynamical system where* $\mu$ *is an acip and consider* $\zeta \in \mathcal{X}$ *for which Lebesgue's Differentiation Theorem holds.*

- *If we have an EVL for* $M_n$ *which applies to the observables* (1.10) *achieving a maximum at* $\zeta \in \mathcal{X}$ *then we have HTS to balls at* $\zeta$.
- *If we have an EVL for* $M_n$ *which coincides with that of* $\hat{M}_n$, *then we have exponential HTS* ($G(t) = e^{-t}$) *to balls at* $\zeta$.

The following is immediate by the above and [21, Theorem 1] (see (1.9)).

**Corollary 2** *Let* $(\mathcal{X}, \mathcal{B}, \mu, f)$ *be a dynamical system where* $\mu$ *is an acip and consider* $\zeta \in \mathcal{X}$ *for which Lebesgue's Differentiation Theorem holds. If* $D_2(u_n)$ (*or* $D_1(u_n)$) *and* $D'(u_n)$ *hold for a stochastic process* $X_0, X_1, \ldots$ *defined by* (1.1) *and* (1.10), *where* $u_n$ *is a sequence of levels satisfying* (1.5), *then we have exponential HTS to balls at* $\zeta$.

The following is an immediate corollary of Theorem 2 and the main theorem of [20].

**Corollary 3** *For every Benedicks-Carleson quadratic map* $f_a$ (*with* $a \in \mathcal{BC}$) *we have exponential HTS to balls around the critical point or the critical value.*

The next result is a byproduct of Theorems 1, 2 and the fact that under $D_1(u_n)$ the only possible limit laws for partial maximums are the classical $EV_i$ for $i \in \{1, 2, 3\}$. Since this is not as immediate as the other corollaries, we include a short proof in Sect. 2.

**Corollary 4** *Let $(\mathcal{X}, \mathcal{B}, \mu, f)$ be a dynamical system $\mu$ is an acip and consider $\zeta \in \mathcal{X}$ for which Lebesgue's Differentiation Theorem holds. If $D_1(u_n)$ holds for a stochastic process $X_0, X_1, \ldots$ defined by* (1.1) *and* (1.10)*, where $u_n$ is a sequence of levels satisfying* (1.5)*, then the only possible HTS to balls around $\zeta$ are of exponential type, meaning that, there is $\theta > 0$ such that $G(t) = e^{-\theta t}$.*

Note that for this corollary to be non-trivial, we must assume that there exists a distribution for HTS. This may not always be the case. For example, in [11,12] it was shown that for certain circle diffeomorphisms there are sequences of intervals $\{U_n\}_{n\in\mathbb{N}}$, $\{V_n\}_{n\in\mathbb{N}}$ which both shrink to the same point $\zeta$, but yield different HTS laws. Note that in these cases $D_1(u_n)$ also fails.

*Remark 2* In this paper we are concerned with observables of the form $\varphi(x) = g(\mathrm{dist}(x, \zeta))$ where $\zeta$ is some typical point for our measure $\mu$ and $g$ is a particular kind of function which is adapted to give different types of EVL. In our proof of the link between HTS and EVL for such observables it is useful to be able to express $\mu(B_\epsilon(\zeta))$ in terms of $\epsilon^d$ where $d$ is the dimension of the space. For measures $\mu$ other than acips, this is highly problematic. In many cases, for a measure $\mu$, the limit $\lim_{\epsilon\to 0} \frac{\log \mu(B_\epsilon(\zeta))}{\log \epsilon}$ exists for $\mu$-a.e. $\zeta$ (in fact this constant can be referred to as the dimension of the measure [32]). However, removing the logs in this expression, we would typically expect the liminf and limsup to be 0 and infinity respectively even for well behaved systems. This occurs for example for the doubling map $f : x \mapsto 2x \bmod 1$ when $\mu$ is the $(\alpha, 1-\alpha)$-Bernoulli measure and $\alpha \neq 1/2$. This follows from the fact that the measure of sets can be computed via certain Birkhoff averages, using the Gibbs property, and the fact that these averages should typically fluctuate, which can be seen from the Law of the Iterated Logarithm in this case, see [16]. We therefore expect serious intrinsic problems in proving results for general measures for this kind of observable with measures which are not acips. On the other hand, if, given a system $(X, f, \mu)$, we are prepared to change our observable to one which reflects properties of $\mu$, i.e. replacing $g(\mathrm{dist}(x, \zeta))$ with $g(\psi(x, \zeta))$ for some $\psi$, then it is possible to prove analogous results to those presented here. This is the subject of forthcoming work.

As we have already mentioned, Corollary 1 generalises the result of Collet in [14], which was for $C^2$ non-uniformly hyperbolic maps of the interval (admitting a Young tower). However, a close look to Collet's arguments allows us to conclude that his result still prevails in higher dimensions. In fact, one can show that if we consider non-uniformly expanding maps (in any finite dimensional compact manifold), admitting a so-called Young tower with exponential return times to the base, then for any sequence of r.v. $X_0, X_1, \ldots$ defined as in (1.1) and for a sequence of levels $u_n$ such that $n\mu(X_0 > u_n) \to \tau > 0$, conditions $D_2(u_n)$ and $D'(u_n)$ hold. This means that by the above theorems, we can prove both EVLs and HTS for these maps. Due to numerous definitions required for that setting, we leave both the theorems and the proofs on this subject to Sect. 6.

Theorems 1 and 2 give us new tools to investigate the recurrence of dynamical systems, principally by allowing us to use the wealth of theory for HTS which has been developed in recent years to prove EVLs. We note that in Corollary 1, the dynamical systems involved need not have any fast rate of decay of correlations at all. Indeed, a priori the relevant system may only have summable decay of correlations. As in Sect. 6 where we consider higher dimensional maps admitting Young towers, there are situations where it is actually easier to check conditions like $D_2(u_n)$ and $D'(u_n)$ in order to get laws for HTS. In fact, to our knowledge, exponential HTS to balls have never been proved before for higher dimensional non-uniformly expanding systems: in such cases, inducing schemes with the nice properties of one-dimensional dynamics are much harder to find. Also the dynamical systems we present in this paper should provide models which can be used in investigating Extreme Value Theory both analytically and numerically. Namely, the simple fact that we get EVLs from deterministic models may be an extra advantage for numerical simulation since there is no need to generate random numbers. This means that this theory may reveal very useful for testing GEV (Generalised Extreme Value distribution) fitting for data corresponding to phenomena for which there is an underlying deterministic model.

The next question that arises is: what about subsequent visits to $U_n$ or subsequent exceedances of the level $u_n$? Namely, we are interested in the *point processes* associated to the instants of occurrence of returns to $U_n$ and exceedances of the level $u_n$. If we have either exponential HTS or a classical EVL then time between hits or exceedances is exponentially distributed. This means that we should expect a Poisson limit for the point processes. We show in Sect. 3 that the relation between HTS and EVL does indeed extend to the laws for the subsequent visits/exceedances (we postpone the precise definitions and results to Sects. 3–5). More precisely, we show that the point process of hitting times has a Poisson limit if and only if the point process of exceedances has a Poisson limit. We next discuss how to obtain a Poisson law in these two different contexts. In Sect. 4 we give conditions which guarantee a Poisson limit for the point process of exceedance times. This part of the paper can be seen as a generalisation of [21]. Moreover, we show that these conditions can be verified in the settings from [14,20], leading to Poisson statistics for both point processes for the systems considered. In Sect. 5 we show that in many cases for multimodal maps it can be shown that the HTS behave asymptotically as a Poisson distribution.

In a parallel work [27], EVLs have also been proved for dynamical systems with acips. For example they proved EVLs for flows, and also considered observables $\varphi$ with multiple maxima. We point out that in the setting of multimodal maps and observables $\varphi$ of the form $\varphi(x) = g(d(x, \zeta))$ where $\zeta$ is some typical point of an acip $\mu$, our Corollary 1 improves on their results since we do not require any knowledge of the decay of correlations.

Throughout this paper the notation $A_n \sim B_n$ means that $\lim_{n\to\infty} \frac{A_n}{B_n} = 1$. Also, if $\{\delta_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ has $\delta_n \to 0$ as $n \to \infty$, then for each $\zeta \in \mathcal{X}$, let $\kappa \in (0, \infty)$ be such that $|B_{\delta_n}(\zeta)| \sim \kappa \cdot \delta_n^d$. Let $x \in \mathbb{R}$. We denote the integer part of $x$ by $\lfloor x \rfloor$ and define $\lceil x \rceil := x$ if $x = \lfloor x \rfloor$, and $\lceil x \rceil := \lfloor x \rfloor + 1$ otherwise.

## 2 Proofs of our results on HTS and EVL

In this section we prove Theorems 1, 2 and Corollary 4.

*Proof of Theorem 1* Let $\rho(\zeta)=\frac{d\mu}{d\mathrm{Leb}}(\zeta)\in\mathbb{R}_0^+$ and set

$$
\begin{aligned}
&u_n = g_1\left((\kappa\rho(\zeta)n)^{-1/d}\right)+p\left(g_1\left((\kappa\rho(\zeta)n)^{-1/d}\right)\right)\tfrac{y}{d}, && \text{for } y\in\mathbb{R}, \text{ for type } g_1;\\
&u_n = g_2\left((\kappa\rho(\zeta)n)^{-1/d}\right)y, && \text{for } y>0, \text{ for type } g_2;\\
&u_n = D-\left(D-g_3\left((\kappa\rho(\zeta)n)^{-1/d}\right)\right)(-y), && \text{for } y<0, \text{ for type } g_3.
\end{aligned}
$$

We remark that the choices of the normalising sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$, which determine the sequence of levels $\{u_n\}_{n\in\mathbb{N}}$ as in (1.4), are made accordingly to the behaviour of the tail of d.f. $F$, given by $F(u)=\mu(X_0\le u)$. [31, Corollary 1.6.3] gives a canonical choice in terms of the tail of $F$. In fact, up to linear scaling the normalising sequences are unique by Khintchine's Theorem (see [31, Theorem 1.2.3]). Since, as we have explained in Remark 1, the shape of $g^{-1}$ determines the tail of the d.f. F, the definitions of the levels $u_n$ above are just a reflection of this fact.

For all $n$ we have

$$
\begin{aligned}
\{x : M_n(x)\le u_n\} &= \bigcap_{j=0}^{n-1}\{x : X_j(x)\le u_n\} = \bigcap_{j=0}^{n-1}\{x : g(\mathrm{dist}(f^j(x),\zeta))\le u_n\}\\
&= \bigcap_{j=0}^{n-1}\{x : \mathrm{dist}(f^j(x),\zeta)\ge g^{-1}(u_n)\} = \{x : r_{B_{g^{-1}(u_n)}(\zeta)}(x)\ge n\}
\end{aligned}
\tag{2.1}
$$

Now, observe that (1.11)–(1.13) imply

$$
\begin{aligned}
g_1^{-1}(u_n) &= g_1^{-1}\left[g_1\left((\kappa\rho(\zeta)n)^{-1/d}\right)+p\left(g_1\left((\kappa\rho(\zeta)n)^{-1/d}\right)\right)\frac{y}{d}\right]\\
&\sim g_1^{-1}\left[g_1\left((\kappa\rho(\zeta)n)^{-1/d}\right)\right]\mathrm{e}^{-y/d} = \left(\frac{\mathrm{e}^{-y}}{\kappa\rho(\zeta)n}\right)^{1/d};\\
g_2^{-1}(u_n) &= g_2^{-1}\left[g_2\left((\kappa\rho(\zeta)n)^{-1/d}\right)y\right]\sim g_2^{-1}\left[g_2\left((\kappa\rho(\zeta)n)^{-1/d}\right)\right]y^{-\beta}\\
&= \left(\frac{y^{-\beta d}}{\kappa\rho(\zeta)n}\right)^{1/d};\\
g_3^{-1}(u_n) &= g_3^{-1}\left[D-\left(D-g_3\left((\kappa\rho(\zeta)n)^{-1/d}\right)\right)(-y)\right]\\
&\sim g_3^{-1}\left[D-\left(D-g_3\left((\kappa\rho(\zeta)n)^{-1/d}\right)\right)\right](-y)^{\gamma} = \left(\frac{(-y)^{\gamma d}}{\kappa\rho(\zeta)n}\right)^{1/d}.
\end{aligned}
$$

Thus, we may write

$$
g^{-1}(u_n)\sim\left(\frac{\tau(y)}{\kappa\rho(\zeta)n}\right)^{1/d},
$$

meaning that

$$g_i^{-1}(u_n) \sim \left(\frac{\tau_i(y)}{\kappa\rho(\zeta)n}\right)^{1/d}, \quad \forall i \in \{1,2,3\}$$

where $\tau_1(y) = \mathrm{e}^{-y}$ for $y \in \mathbb{R}$, $\tau_2(y) = y^{-\beta d}$ for $y > 0$, and $\tau_3(y) = (-y)^{\gamma d}$ for $y < 0$.

Since Lebesgue's Differentiation Theorem holds for $\zeta \in \mathcal{X}$, we have $\frac{\mu(B_\delta(\zeta))}{|B_\delta(\zeta)|} \to \rho(\zeta)$ as $\delta \to 0$. Consequently, since it is obvious that $g^{-1}(u_n) \to 0$ as $n \to \infty$, then

$$\mu\left(B_{g^{-1}(u_n)}(\zeta)\right) \sim \rho(\zeta)|B_{g^{-1}(u_n)}(\zeta)| \sim \rho(\zeta)\kappa(g^{-1}(u_n))^d = \rho(\zeta)\kappa\frac{\tau(y)}{\kappa\rho(\zeta)n} = \frac{\tau(y)}{n}.$$

Thus, we have

$$n \sim \frac{\tau(y)}{\mu\left(B_{g^{-1}(u_n)}(\zeta)\right)}. \tag{2.2}$$

Now, we claim that using (2.1) and (2.2), we have

$$\lim_{n\to\infty}\mu(\{x : M_n(x) \le u_n\}) = \lim_{n\to\infty}\mu\left(\left\{x : r_{B_{g^{-1}(u_n)}(\zeta)}(x) \ge \frac{\tau(y)}{\mu\left(B_{g^{-1}(u_n)}(\zeta)\right)}\right\}\right) \tag{2.3}$$

$$= G(\tau(y)), \tag{2.4}$$

which gives the first part of the theorem.

To see that (2.3) holds, observe that by (2.1) and (2.2) we have

$$\begin{aligned}&\left|\mu(\{M_n \le u_n\}) - \mu\left(\left\{r_{B_{g^{-1}(u_n)}(\zeta)} \ge \frac{\tau(y)}{\mu\left(B_{g^{-1}(u_n)}(\zeta)\right)}\right\}\right)\right| \\ &\quad = \left|\mu\left(\left\{r_{B_{g^{-1}(u_n)}(\zeta)} \ge n\right\}\right) - \mu\left(\left\{r_{B_{g^{-1}(u_n)}(\zeta)} \ge (1+\varepsilon_n)n\right\}\right)\right|,\end{aligned}$$

where $\{\varepsilon_n\}_{n\in\mathbb{N}}$ is such that $\varepsilon_n \to 0$ as $n \to \infty$. Since we have

$$\begin{aligned}&\left\{r_{B_{g^{-1}(u_n)}(\zeta)} \ge m\right\} \Big\backslash \left\{r_{B_{g^{-1}(u_n)}(\zeta)} \ge m+k\right\} \\ &\quad \subset \bigcup_{j=m}^{m+k-1} f^{-j}\left(B_{g^{-1}(u_n)}(\zeta)\right), \quad \forall m,k \in \mathbb{N},\end{aligned} \tag{2.5}$$

it follows by stationarity that

$$\begin{aligned}&\left|\mu\left(\left\{r_{B_{g^{-1}(u_n)}(\zeta)} \ge n\right\}\right) - \mu\left(\left\{r_{B_{g^{-1}(u_n)}(\zeta)} \ge (1+\varepsilon_n)n\right\}\right)\right| \\ &\quad \le |\varepsilon_n| n\mu\left(B_{g^{-1}(u_n)}(\zeta)\right) \sim |\varepsilon_n|\tau \to 0,\end{aligned}$$

as $n \to \infty$, completing the proof of (2.3).

Next we will use the exponential HTS hypothesis, that is $G(t) = \mathrm{e}^{-t}$, to show the second part of the theorem.

Under the exponential HTS assumption, by (2.4) it follows immediately that $\lim_{n\to\infty} \mu(\{x : M_n(x) \le u_n\}) = \mathrm{e}^{-\tau(y)}$. Recall that in the corresponding i.i.d setting, i.e. when we are considering $\{x : \hat{M}_n(x) \le u_n\}$ rather than $\{x : M_n(x) \le u_n\}$, (1.5) is equivalent to (1.7). Therefore we also have $\lim_{n\to\infty} \mu(\{x : \hat{M}_n(x) \le u_n\}) = \mathrm{e}^{-\tau(y)}$, since $n\mu(X_0 > u_n) = n\mu\,(B_{g^{-1}(u_n)}(\zeta)) \to \tau(y)$, as $n \to \infty$. As explained in the introduction, this means that in the i.i.d. setting $G(\tau)$ must be of the three classical types. It remains to show that if the observable is of type $g_i$ then $\lim_{n\to\infty} \mu(\{x : M_n(x) \le u_n\}) = \mathrm{e}^{-\tau(y)}$ means that the EVL that applies to $M_n$ (rather than $\hat{M}_n$) is also of type $EV_i$, for each $i \in \{1, 2, 3\}$.

Type $g_1$: In this case we have $\mathrm{e}^{-\tau_1(y)} = \mathrm{e}^{-\mathrm{e}^{-y}}$, for all $y \in \mathbb{R}$, that corresponds to the Gumbel e.v.d. and so we have an EVL for $M_n$ of type $EV_1$.

Type $g_2$: We obtain $\mathrm{e}^{-\tau_2(y)} = \mathrm{e}^{-y^{-\beta d}}$ for $y > 0$. To conclude that in this case we have the Fréchet e.v.d. with parameter $\beta d$, we only have to check that for $y \le 0$, $\mu(\{x : M_n(x) \le u_n\}) = 0$. Since $g_2((\kappa\rho(\zeta)n)^{-1/d}) > 0$ (for all large $n$) and

$$\mu(\{x : M_n(x) \le u_n\}) = \mu\left(\left\{x : M_n(x) \le g_2\left((\kappa\rho(\zeta)n)^{-1/d}\right) y\right\}\right) \to \mathrm{e}^{-y^{-\beta d}}$$

as $n \to \infty$. Letting $y \downarrow 0$, it follows that $\mu(\{x : M_n(x) \le 0\}) \to 0$, and, for $y < 0$,

$$\begin{aligned}\mu(\{x : M_n(x) \le u_n\}) &= \mu\left(\left\{x : M_n(x) \le g_2\left((\kappa\rho(\zeta)n)^{-1/d}\right) y\right\}\right)\\ &\le \mu(\{x : M_n(x) \le 0\}) \to 0.\end{aligned}$$

So, we have, in this case, an EVL for $M_n$ of type $EV_2$.

Type $g_3$: For $y < 0$, we have $\mathrm{e}^{-\tau_3(y)} = \mathrm{e}^{-(-y)^{\gamma d}}$. To conclude that in this case we have the Weibull e.v.d. with parameter $\gamma d$, we only need to check that for $y \ge 0$, $\mu(\{x : M_n(x) \le u_n\}) = 1$. In fact, for $y \ge 0$, since $D - g_3((\kappa\rho(\zeta)n)^{-1/d}) > 0$, we have

$$\begin{aligned}\mu(\{x : M_n(x) \le u_n\}) &= \mu\left(\left\{x : M_n(x) \le \left(D - g_3\left((\kappa\rho(\zeta)n)^{-1/d}\right)\right) y + D\right\}\right)\\ &\ge \mu(\{x : M_n(x) \le D\}) = 1.\end{aligned}$$

So we have, in this case, an EVL for $M_n$ of type $EV_3$. □

For the proof of Theorem 2, we will require the following lemma. This is essentially contained in [31, Theorem 1.6.2], but we give a proof for completeness.

**Lemma 2.1** *Suppose that $(X, \mathcal{B}, f, \mu)$ is a dynamical system where $\mu$ is an acip. Furthermore, let $\varphi(x) = g(\mathrm{dist}(x, \zeta))$ for some $\zeta \in X$ where $g$ is one of the three types described above. Then, for each $y \in \mathbb{R}$, there exists a sequence $\{u_n(y)\}_{n\in\mathbb{N}}$ as in (1.4) such that*

$$n\mu(\{x : \varphi(x) > u_n(y)\}) \xrightarrow[n\to\infty]{} \tau(y) \ge 0.$$

*Moreover, for every $t > 0$ there exists $y \in \mathbb{R}$ such that $\tau(y) = t$.*

*Proof* We will prove the lemma in the case when $g$ is of type $g_2$, which means that we will only need to consider $y > 0$. For the other two types of $g$, the argument is the same, but with minor adjustments, see [31, Theorem 1.6.2].

First we show that we can always find a sequence $\{\gamma_n\}_{n\in\mathbb{N}}$ such that

$$n\mu(X_0 > \gamma_n) \xrightarrow[n\to\infty]{} 1.$$

Take $\gamma_n := \inf\{y : \mu(X_0 \le y) \ge 1-1/n\}$, and let us show that it has the desired property. Note that $n\mu(X_0 > \gamma_n) \le 1$, which means that $\limsup_{n\to\infty} n\mu(X_0 > \gamma_n) \le 1$. Using (1.12), for any $z < 1$, we have

$$\begin{aligned}\liminf_{n\to\infty} \frac{\mu(X_0 > \gamma_n)}{\mu(X_0 > \gamma_n z)} &= \liminf_{n\to\infty} \frac{\rho(\zeta)\kappa\left(g_2^{-1}(\gamma_n)\right)^d}{\rho(\zeta)\kappa\left(g_2^{-1}(z\gamma_n)\right)^d} \\ &= \liminf_{n\to\infty} \frac{\left(g_2^{-1}(\gamma_n)\right)^d}{\left(z^{-\beta}g_2^{-1}(\gamma_n)\right)^d} = z^{\beta d}.\end{aligned}$$

Since, by definition of $\gamma_n$, for any $z < 1$, $n\mu(X_0 > \gamma_n z) \ge 1$, letting $z \to 1$, it follows immediately that $\liminf_{n\to\infty} n\mu(X_0 > \gamma_n) \ge 1$.

Now let $u_n(y) = \gamma_n y$, which means that, for all $n \in \mathbb{N}$, we are taking $a_n = \gamma_n^{-1}$ and $b_n = 0$ in (1.4). Then, using (1.12), it follows that for all $y > 0$

$$\begin{aligned}n\mu(X_0 > \gamma_n y) &= n\mu\left(\{x : \mathrm{dist}(x,\zeta) < g_2^{-1}(\gamma_n y)\}\right) \sim n\rho(\zeta)\kappa\left(g_2^{-1}(\gamma_n y)\right)^d \\ &\sim n\rho(\zeta)\kappa(y^{-\beta}g_2^{-1}(\gamma_n))^d \sim y^{-\beta d}n\rho(\zeta)\kappa\left(g_2^{-1}(\gamma_n)\right)^d \\ &\sim y^{-\beta d}n\mu(X_0 > \gamma_n) \xrightarrow[n\to\infty]{} y^{-\beta d}.\end{aligned}$$

So taking $y = t^{-1/(\beta d)} > 0$ would suit our purposes. □

*Proof of Theorem 2* We first assume that for every $y \in \mathbb{R}$ and some sequence $u_n = u_n(y)$ as in (1.4) such that $n\mu(\{x : \varphi(x) > u_n(y)\}) \xrightarrow[n\to\infty]{} \tau(y)$, we have

$$\lim_{n\to\infty} \mu\left(\{x : M_n(x) \le u_n(y)\}\right) = H(\tau(y)).$$

Observe that, by Khintchine's Theorem (see [31, Theorem 1.2.3]), up to linear scaling the normalising sequences are unique, which means that we may assume that they are the ones given by Lemma 2.1. Hence given $t > 0$, Lemma 2.1 implies that there exists $y \in \mathbb{R}$ such that

$$n\mu\left(\{x : \varphi(x) > u_n(y)\}\right) \xrightarrow[n\to\infty]{} t.$$

Given $\{\delta_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ with $\delta_n \xrightarrow[n\to\infty]{} 0$, we define

$$\ell_n := \lfloor t/(\kappa\rho(\zeta)\delta_n^d)\rfloor.$$

We will prove

$$g^{-1}\left(u_{\ell_n}\right) \sim \delta_n. \tag{2.6}$$

If $n$ is sufficiently large, then

$$\begin{aligned}\{x : \varphi(x) > u_n\} &= \{x : g(\mathrm{dist}(x,\zeta)) > u_n\} \\ &= \{x : \mathrm{dist}(x,\zeta) < g^{-1}(u_n)\} = B_{g^{-1}(u_n)}(\zeta).\end{aligned}$$

Hence, by assumption on the sequence $u_n$, we have $n\mu(B_{g^{-1}(u_n)}(\zeta)) \xrightarrow[n\to\infty]{} \tau(y) = t$. As Lebesgue's Differentiation Theorem holds for $\zeta \in \mathcal{X}$, we have $\frac{\mu(B_\delta(\zeta))}{|B_\delta(\zeta)|} \to \rho(\zeta)$ as $\delta \to 0$. Consequently, since it is obvious that $g^{-1}(u_n) \to 0$ as $n \to \infty$, then $n|B_{g^{-1}(u_n)}(\zeta)| \xrightarrow[n\to\infty]{} t/\rho(\zeta)$. Thus, we may write $g^{-1}(u_n) \sim (\frac{t}{\kappa n\rho(\zeta)})^{1/d}$ and substituting $n$ by $\ell_n$ we are immediately led to (2.6) by definition of $\ell_n$.

Next, using Lebesgue's Differentiation Theorem, again, we get $\mu(B_{\delta_n}(\zeta)) \sim \rho(\zeta)\kappa\delta_n^d$ which easily implies that by definition of $\ell_n$,

$$\frac{t}{\mu\left(B_{\delta_n}(\zeta)\right)} \sim \ell_n. \tag{2.7}$$

Now we note that, as in (2.1)

$$\begin{aligned}\{x : M_{\ell_n}(x) \le u_{\ell_n}\} &= \bigcap_{j=0}^{\ell_n-1}\{x : X_j(x) \le u_{\ell_n}\} = \bigcap_{j=0}^{\ell_n-1}\{x : g(\mathrm{dist}(f^j(x),\zeta)) \le u_{\ell_n}\} \\ &= \bigcap_{j=0}^{\ell_n-1}\{x : \mathrm{dist}(f^j(x),\zeta) \ge g^{-1}(u_{\ell_n})\} = \{x : r_{B_{g^{-1}(u_{\ell_n})}(\zeta)}(x) \ge \ell_n\}.\end{aligned} \tag{2.8}$$

At this point, we claim that

$$\lim_{n\to\infty}\mu\left(\left\{x : r_{B_{\delta_n}(\zeta)}(x) \ge \frac{t}{\mu(B_{\delta_n}(\zeta))}\right\}\right) = \lim_{n\to\infty}\mu\left(\{x : M_{\ell_n}(x) \le u_{\ell_n}\}\right). \tag{2.9}$$

Then, the first part of the theorem follows, once we observe that, by hypothesis, we have

$$\mu\left(\{x : M_{\ell_n}(x) \le u_{\ell_n}\}\right) \xrightarrow[n\to\infty]{} H(\tau(y)) = H(t).$$

For the second part of the theorem, first notice that for the i.i.d. setting, i.e. when we are considering $\{x : \hat{M}_n(x) \le u_n\}$ rather than $\{x : M_n(x) \le u_n\}$, (1.5) is equivalent to (1.7). Therefore, $\mu(\{x : \hat{M}_n(x) \leq u_n\}) \to \mathrm{e}^{-\tau(y)}$ as $n \to \infty$. Hence if the EVL of $M_n$ coincides with that of $\hat{M}_n$, then we also have $H(\tau(y)) = \mathrm{e}^{-\tau(y)}$.

It remains to show that (2.9) holds. First, observe that

$$\begin{aligned}
&\mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \frac{t}{\mu(B_{\delta_n}(\zeta))}\right\}\right)\\
&\quad = \mu\left(\{M_{\ell_n} \leq u_{\ell_n}\}\right) + \left(\mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \ell_n\right\}\right) - \mu\left(\{M_{\ell_n} \leq u_{\ell_n}\}\right)\right)\\
&\qquad + \left(\mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \frac{t}{\mu(B_{\delta_n}(\zeta))}\right\}\right) - \mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \ell_n\right\}\right)\right).
\end{aligned}$$

For the third term on the right, note that by (2.7) we have

$$\begin{aligned}
&\left|\mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \ell_n\right\}\right) - \mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \frac{t}{\mu(B_{\delta_n}(\zeta))}\right\}\right)\right|\\
&\quad = \left|\mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \ell_n\right\}\right) - \mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq (1+\varepsilon_n)\ell_n\right\}\right)\right|,
\end{aligned}$$

for some sequence $\{\varepsilon_n\}_n \in \mathbb{N}$ such that $\varepsilon_n \to 0$, as $n \to \infty$. By (2.5), (2.7) and stationarity it follows that

$$\left|\mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \ell_n\right\}\right) - \mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq (1+\varepsilon_n)\ell_n\right\}\right)\right| \leq |\varepsilon_n|\ell_n\mu\left(B_{\delta_n}(\zeta)\right) \sim |\varepsilon_n|t \to 0,$$

as $n \to \infty$.

For the remaining term, using (2.6)–(2.8), we have

$$\begin{aligned}
&\left|\mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \ell_n\right\}\right) - \mu\left(\{M_{\ell_n} \leq u_{\ell_n}\}\right)\right|\\
&\quad = \left|\mu\left(\left\{r_{B_{\delta_n}(\zeta)} \geq \ell_n\right\}\right) - \mu\left(\{r_{B_{g^{-1}(u_{\ell_n})}(\zeta)} \geq \ell_n\}\right)\right|\\
&\quad \leq \sum_{i=1}^{\ell_n} \mu\left(f^{-i}\left(B_{\delta_n}(\zeta) \triangle B_{g^{-1}(u_{\ell_n})}(\zeta)\right)\right)\\
&\quad = \ell_n\mu\left(B_{\delta_n}(\zeta) \triangle B_{g^{-1}(u_{\ell_n})}(\zeta)\right)\\
&\quad \sim \frac{t}{\mu\left(B_{\delta_n}(\zeta)\right)}\left|\mu\left(B_{\delta_n}(\zeta)\right) - \mu\left(B_{g^{-1}(u_{\ell_n})}(\zeta)\right)\right|\\
&\quad = t\left|1 - \frac{\mu\left(B_{g^{-1}(u_{\ell_n})}(\zeta)\right)}{\mu\left(B_{\delta_n}(\zeta)\right)}\right| \to 0
\end{aligned}$$

as $n \to \infty$, which ends the proof of (2.9). □

*Proof of Corollary 4* Let us assume the existence of HTS to balls around $\zeta$ (not necessarily exponential). Then the first part of Theorem 1 assures the existence of an EVL as in (1.3) for $M_n$ defined in (1.2). This fact and the hypothesis that $D_1(u_n)$ holds allows us to use [31, Theorem 3.7.1] to conclude that there is $\theta > 0$ such that

$\lim_{n\to\infty} \mu(M_n \le u_n) = \mathrm{e}^{-\theta\tau}$. Finally, we use the first part of Theorem 2 to conclude that we have HTS to balls centred on $\zeta$ of exponential type. □

## 3 Relation between hitting times and exceedance point processes

We have already seen how to relate HTS and EVL. We next show that if we enrich the process and the statistics by considering either multiple returns or multiple exceedances we can take the parallelism even further.

Given a sequence $\{\delta_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ such that $\delta_n \xrightarrow[n\to\infty]{} 0$, for each $j \in \mathbb{N}$, we define the $j$th *waiting (or inter-hitting) time* as

$$w^j_{B_{\delta_n}(\zeta)}(x) = r_{B_{\delta_n}(\zeta)}\left(f^{w^1_{B_{\delta_n}(\zeta)}(x)+\cdots+w^{j-1}_{B_{\delta_n}(\zeta)}(x)}(x)\right), \tag{3.1}$$

and the $j$th *hitting time* as

$$r^j_{B_{\delta_n}(\zeta)}(x) = \sum_{i=1}^{j} w^i_{B_{\delta_n}(\zeta)}(x).$$

We define the *Hitting Times Point Process* (HTPP) by counting the number of hitting times during the time interval $[0,t)$. However, since $\mu(B_{\delta_n}(\zeta)) \to 0$, as $n \to \infty$, then by Kac's Theorem, the expected waiting time between hits is diverging to $\infty$ as $n$ increases. This fact suggests a time re-scaling using the factor $v^*_n := 1/\mu(B_{\delta_n}(\zeta))$, which is precisely the expected inter-hitting time. Hence, for any $x \in \mathcal{X}$ and every $t \ge 0$ define

$$N^*_n(t) = N^*_n([0,t),x) := \sup\left\{j : r^j_{B_{\delta_n}(\zeta)}(x) \le v^*_n t\right\} = \sum_{j=0}^{\lfloor v^*_n t\rfloor} \mathbf{1}_{B_{\delta_n}(\zeta)} \circ f^j \tag{3.2}$$

When $x \in B_{\delta_n}(\zeta)$ and we consider the conditional measure $\mu_{B_{\delta_n}(\zeta)}$ instead of $\mu$, then we refer to $N^*_n(t)$ as the *Return Times Point Process* (RTPP).

If we have exponential HTS, ($G(t) = \mathrm{e}^{-t}$ in (1.14)), then the distribution of the waiting time before hitting $B_{\delta_n}(\zeta)$ is asymptotically exponential. Also, if we assume that our systems are mixing, because in that case we can think that the process gets renewed when we come back to $B_{\delta_n}(\zeta)$, then one may look at the hitting times as the sum of almost independent r.v.s that are almost exponentially distributed. Hence, one would expect that the hitting times, when properly re-scaled, should form a point process with a Poisson type behaviour at the limit.

As discussed in Sect. 1.2, for hyperbolic systems, it is indeed the case that we do get a Poisson Process as the limit of HTPP. The theory in [7,8,26] implies that if $f \in NF^2$ has an acip then we have a Poisson limit for the HTPP. We postpone a sketch of this fact to Sect. 5, in order to keep our focus on the relation between HTS and EVL here. However, we would like to remark that a key difference between proofs for the first hitting time and for showing that we have a Poisson Point Process, if we are using the theory started in [26], is that a further mixing condition is required.

Now, we turn to an EVL point of view. In this context, one is concerned with the occurrence of exceedances of the level $u_n$ for the stationary stochastic process $X_0, X_1, \ldots$ In particular, we are interested in counting the number of exceedances, among a random sample $X_0, \ldots, X_{n-1}$ of size $n$. As in the previous sections, we consider the stationary stochastic process defined by (1.1) and a sequence of levels $\{u_n\}_{n\in\mathbb{N}}$ such that $n\mu(X_0 > u_n) \to \tau > 0$, as $n \to \infty$. We define the *exceedance point process* (EPP) by counting the number of exceedances during the time interval $[0, t)$. We re-scale time using the factor $v_n := 1/\mu(X > u_n)$ given by Kac's Theorem, again. Then for any $x \in \mathcal{X}$ and every $t \geq 0$, set

$$N_n(t) = N_n([0, t), x) := \sum_{j=0}^{\lfloor v_n t\rfloor} \mathbf{1}_{X_j > u_n}. \tag{3.3}$$

The limit laws for these point processes can be used to assess the impact and damage caused by rare events since they describe their time occurrences, their individual impacts and accumulated effects. Assuming that the process is mixing, we almost have a situation of many Bernoulli trials where the expected number of successes is almost constant ($n\mu(X > u_n) \to \tau > 0$). Thus, we expect a Poisson law as a limit. In fact, one should expect that the exceedance instants, when properly normalised, should form a point process with a Poisson Process as a limit, also. This is the content of [31, Theorem 5.2.1] which states that under $D_1(u_n)$ and $D'(u_n)$, the EPP $N_n$, when properly normalised, converges in distribution to a Poisson Process. (See [31, Chapter 5], [28] and references therein for more information on the subject).

Similarly to Theorems 1 and 2, we show that if there exists a limiting continuous time stochastic process for the HTPP, when properly normalised, then the same holds for the EPP and vice-versa. In the sequel $\xrightarrow{d}$ denotes convergence in distribution.

**Theorem 3** *Let $(\mathcal{X}, \mathcal{B}, \mu, f)$ be a dynamical system where $\mu$ is an acip and consider $\zeta \in \mathcal{X}$ for which Lebesgue's Differentiation Theorem holds. Suppose that for any sequence $\delta_n \xrightarrow[n\to\infty]{} 0$ we have that the HTPP defined in (3.2) is such that $N_n^* \xrightarrow[n\to\infty]{d} N$, where $N$ is a continuous time stochastic process. Then, for the EPP defined in (3.3) we also have $N_n \xrightarrow[n\to\infty]{d} N$.*

*Proof* The result follows immediately once we set $\delta_n = g^{-1}(u_n)$ and observe that for every $j, n \in \mathbb{N}$ and $x \in \mathcal{X}$ we have $\{x : X_j > u_n\} = \{x : f^j(x) \in B_{g^{-1}(u_n)}(\zeta)\}$, which implies that $N_n(t) = N_n^*(t)$, for all $t \geq 0$. □

**Corollary 5** *Suppose that $f \in NF^2$ and $f$ has an acip $\mu$. Then, denoting by $N_n$ the associated EPP as in (3.3), we have $N_n \xrightarrow{d} N$, as $n \to \infty$, where $N$ denotes a Poisson Process with intensity 1.*

The fact that the maps in this corollary satisfy the conditions of Theorem 3 follows from the sketch in Sect. 5. So the result is otherwise immediate.

**Theorem 4** *Let $(\mathcal{X}, \mathcal{B}, \mu, f)$ be a dynamical system where $\mu$ is an acip and consider $\zeta \in \mathcal{X}$ for which Lebesgue's Differentiation Theorem holds. Suppose that for a sequence of levels $\{u_n\}_{n\in\mathbb{N}}$ such that $n\mu(X_0 > u_n) \to \tau > 0$, as $n \to \infty$, the EPP defined in (3.3) is such that $N_n \xrightarrow[n\to\infty]{d} N$, where $N$ is a continuous time stochastic process. Then, for the HTPP defined in (3.3) we also have $N_n^* \xrightarrow[n\to\infty]{d} N$.*

*Proof* Given a sequence $\{\delta_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ with $\delta_n \xrightarrow[n\to\infty]{} 0$ we define, as in the proof of Theorem 2, the sequence $\ell_n$ such that $\delta_n \sim g^{-1}(u_{\ell_n})$. Set $k_n := \max\{v_n^*, v_{\ell_n}\}$ and observe that $|N_n^*(t) - N_{\ell_n}(t)| \leq \sum_{j=0}^{k_n} \mathbf{1}_{B_{\delta_n}(\zeta)\triangle B_{g^{-1}(u_{\ell_n})}(\zeta)} \circ f^j$. Using stationarity we get

$$\begin{aligned}\mu\left(|N_n^*(t) - N_{\ell_n}(t)| > 0\right) &\leq k_n\mu\left(B_{\delta_n}(\zeta)\triangle B_{g^{-1}(u_{\ell_n})}(\zeta)\right)\\ &= k_n\left|\mu\left(B_{\delta_n}(\zeta)\right) - \mu\left(B_{g^{-1}(u_{\ell_n})}(\zeta)\right)\right| \xrightarrow[n\to\infty]{} 0,\end{aligned}$$

by definition of $\ell_n$. The result now follows immediately by Slutsky's Theorem (see [18, Theorem 6.3.15]). $\square$

## 4 Poisson statistics via EVL

As we have already mentioned, [31, Theorem 5.2.1] states that for a stationary stochastic process satisfying $D_1(u_n)$ and $D'(u_n)$, the EPP $N_n$ defined in (3.3) converges in distribution to a Poisson Process.

The main result in [21] states that in order to prove an EVL for stationary stochastic processes arising from a dynamical system, it suffices to show conditions $D_2(u_n)$ and $D'(u_n)$. This proved to be an advantage over [31, Theorem 3.5.2] since the mixing information of systems is usually known through decay of correlations that can be easily used to prove $D_2(u_n)$, as opposed to condition $D_1(u_n)$ appearing in [31, Theorem 3.5.2].

Our goal here is to prove that we still get the Poisson limit if we relax $D_1(u_n)$ so that it suffices to have sufficiently fast decay of correlations of the dynamical systems that generate the stochastic processes. However, for that purpose, one needs to strengthen $D_2(u_n)$ in order to cope with multiple events. (Something similar was necessary in the corresponding theory in [26]). For that reason we introduce condition $D_3(u_n)$ below, that still follows from sufficiently fast decay of correlations, as $D_2(u_n)$ did, and together with $D'(u_n)$ allows us to obtain the Poisson limit for the EPP.

Let $\mathcal{S}$ denote the semi-ring of subsets of $\mathbb{R}_0^+$ whose elements are intervals of the type $[a, b)$, for $a, b \in \mathbb{R}_0^+$. Let $\mathcal{R}$ denote the ring generated by $\mathcal{S}$. Recall that for every $A \in \mathcal{R}$ there are $k \in \mathbb{N}$ and $k$ intervals $I_1, \ldots, I_k \in \mathcal{S}$ such that $A = \cup_{i=1}^k I_j$. In order to fix notation, let $a_j, b_j \in \mathbb{R}_0^+$ be such that $I_j = [a_j, b_j) \in \mathcal{S}$. For $I = [a, b) \in \mathcal{S}$ and $\alpha \in \mathbb{R}$, we denote $\alpha I := [\alpha a, \alpha b)$ and $I + \alpha := [a + \alpha, b + \alpha)$. Similarly, for $A \in \mathcal{R}$ define $\alpha A := \alpha I_1 \cup \cdots \cup \alpha I_k$ and $A + \alpha := (I_1 + \alpha) \cup \cdots \cup (I_k + \alpha)$.

For every $A \in \mathcal{R}$ we define

$$M(A) := \max\{X_i : i \in A \cap \mathbb{Z}\}.$$

In the particular case where $A = [0, n)$ we simply write, as before, $M_n = M[0, n)$.
At this point, we propose:

**Condition** ($D_3(u_n)$) Let $A \in \mathcal{R}$ and $t \in \mathbb{N}$.
We say that $D_3(u_n)$ holds for the sequence $X_0, X_1, \ldots$ if

$$\mu\left(\{X_0 > u_n\} \cap \{M(A+t) \leq u_n\}\right) - \mu(\{X_0 > u_n\})\mu(\{M(A) \leq u_n\}) \leq \gamma(n, t),$$

where $\gamma(n, t)$ is nonincreasing in $t$ for each $n$ and $n\gamma(n, t_n) \to 0$ as $n \to \infty$ for some sequence $t_n = o(n)$, which means that $t_n/n \to 0$ as $n \to \infty$.

Recalling the definition of the EPP $N_n(t) = N_n[0, t)$ given in (3.3), we set

$$N_n[a, b) := N(b) - N(a) = \sum_{j=\lceil v_n a\rceil}^{\lfloor v_n b\rfloor} \mathbf{1}_{\{X_j > u_n\}}.$$

We now state the main result of this section that gives the Poisson statistics for the EPP under $D_3(u_n)$ and $D'(u_n)$.

**Theorem 5** *Let $X_1, X_2, \ldots$ be a stationary stochastic process for which conditions $D_3(u_n)$ and $D'(u_n)$ hold for a sequence of levels $u_n$ such that $n\mu(X_0 > u_n) \to \tau > 0$, as $n \to \infty$. Then the EPP $N_n$ defined in (3.3) is such that $N_n \xrightarrow{d} N$, as $n \to \infty$, where $N$ denotes a Poisson Process with intensity* 1.

As a consequence of this theorem, Theorem 4 and the results in [20] we get:

**Corollary 6** *For any Benedicks-Carleson quadratic map $f_a$ (with $a \in \mathcal{BC}$), consider a stochastic process $X_0, X_1, \ldots$ defined by (1.1) and (1.10), with $\zeta$ being either the critical point or the critical value. Then, denoting by $N_n$ the associated EPP as in (3.3), we have $N_n \xrightarrow{d} N$, as $n \to \infty$, where $N$ denotes a Poisson Process with intensity* 1. *Moreover, if we consider $N_n^*$, the HTPP as in (3.2), for balls around either the critical point or the critical value, then the same limit also applies to $N_n^*$.*

With minor adjustments to [14], we can use Theorem 5 to show that, similarly to Corollary 5, interval maps with exponential decay of correlations have Poisson statistics for the EPP. However, we will not state this result here, since we prove a more general result (which works in higher dimensions) in Sect. 6.

### 4.1 Proofs of the results

In this section we prove Theorem 5 and Corollary 6. The key is Proposition 1 whose proof we prepare with the following two Lemmas. These are very similar to ones in

[14, Section 3] and [21, Section 3], so we omit the details of the proofs. However, note that in contrast to the lemmas in the papers mentioned above, we need them to take care of events that depend on nonconsecutive random variables.

**Lemma 4.1** *For any $\ell \in \mathbb{N}$ and $u \in \mathbb{R}$ we have*

$$\sum_{j=0}^{\ell-1} \mu(X_j > u) \geq \mu(M_\ell > u)$$
$$\geq \sum_{j=0}^{\ell-1} \mu(X_j > u) - \sum_{j=0}^{\ell-1} \sum_{i=0, i \neq j}^{\ell-1} \mu(\{X_j > u\} \cap \{X_i > u\})$$

*Proof* This is a straightforward consequence of the formula for the probability of a multiple union of events. See for example the first Theorem of Chapter 4 in [19].

**Lemma 4.2** *Assume that $r, s, \ell, t$ are nonnegative integers. Suppose that $A, B \in \mathcal{R}$ are such that $A \subset B$. Set $\ell := \#\{j \in \mathbb{N} : j \in B \setminus A\}$. Assume that* $\min\{x : x \in A\} \geq r + t$ *and let $A_0 = [0, r+t)$. Then, we have*

$$0 \leq \mu(M(A) \leq u) - \mu(M(B) \leq u) \leq \ell \cdot \mu(X > u) \tag{4.1}$$

*and*

$$\left| \mu(M(A_0 \cup A) \leq u) - \mu(M(A) \leq u) + \sum_{i=0}^{r-1} \mu\left(\{X > u\} \cap \{M(A - i) \leq u\}\right) \right|$$
$$\leq 2r \sum_{i=1}^{r-1} \mu(\{X > u\} \cap \{X_i > u\}) + t\mu(X > u). \tag{4.2}$$

The proof of this lemma can easily be done by following the proof of [21, Lemma 3.2] with minor adjustments.

**Proposition 1** *Let $A \in \mathcal{R}$ be such that that $A = \bigcup_{j=1}^{p} I_j$ where $I_j = [a_j, b_j) \in \mathcal{S}$, $j = 1, \ldots, p$ and $a_1 < b_1 < a_2 < \cdots < b_{p-1} < a_p < b_p$. Let $\{u_n\}_{n \in \mathbb{N}}$ be such that $n\mu(X_0 > u_n) \to \tau > 0$, as $n \to \infty$, for some $\tau \geq 0$. Assume that conditions $D_3(u_n)$ and $D'(u_n)$ hold. Then,*

$$\mu\left(M(nA) \leq u_n\right) \xrightarrow[n \to +\infty]{} \prod_{j=1}^{p} \mu(M(nI_j) \leq u_n) = \prod_{j=1}^{p} e^{-\tau(b_j - a_j)}.$$

*Proof* Let $h := \inf_{j \in \{1,\ldots,p\}}\{b_j - a_j\}$ and $H := \lceil \sup\{x : x \in A\} \rceil$. Take $k > 2/h$ and $n$ sufficiently large. Note this guarantees that if we partition $n[0, H] \cap \mathbb{Z}$ into blocks of length $r_n := \lfloor n/k \rfloor$, $J_1 = [Hn - r_n, Hn)$, $J_2 = [Hn - 2r_n, Hn - r_n)$,..., $J_{Hk} = [Hn - Hkr_n, n - (Hk-1)r_n)$, $J_{Hk+1} = [0, Hn - Hkr_n)$, then there is more

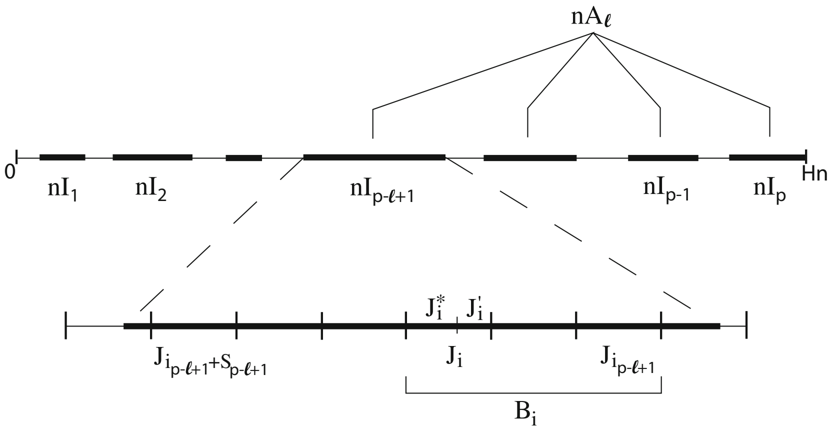


**Fig. 1** Notation

than one of these blocks contained in $nI_i$. Let $S_\ell = S_\ell(k)$ be the number of blocks $J_j$ contained in $nI_\ell$, that is,

$$S_\ell := \#\{j \in \{1, \ldots, Hk\} : J_j \subset nI_\ell\}.$$

As we have already observed $S_\ell > 1 \; \forall \ell \in \{1, \ldots, p\}$. For each $\ell \in \{1, \ldots, p\}$, we define

$$A_\ell := \bigcup_{i=1}^{\ell} I_{p-i+1}.$$

Set $i_\ell := \min\{j \in \{1, \ldots, k\} : J_j \subset nI_\ell\}$. Then $J_{i_\ell}, J_{i_\ell+1}, \ldots, J_{i_\ell+s_\ell} \subset nI_\ell$. Now, fix $\ell$ and for each $i \in \{i_{p-\ell+1}, \ldots, i_{p-\ell+1} + S_{p-\ell+1}\}$ let

$$B_i := \bigcup_{j=i_{p-\ell+1}}^{i} J_j, \;\; J_i^* := [Hn - ir_n, Hn - (i-1)r_n - t_n) \;\text{ and }\; J_i' := J_i - J_i^*.$$

Note that $|J_i^*| = r_n - t_n$ and $|J_i'| = t_n$. See Fig. 1 for more of an idea of the notation here.

We have,

$$\begin{aligned}
&\left|\mu(M(B_i \cup nA_{\ell-1}) \le u_n) - (1 - r_n\mu(X > u_n))\mu(M(B_{i-1} \cup nA_{\ell-1}) \le u_n)\right| \\
&\qquad = \Big|\mu(M(B_i \cup nA_{\ell-1}) \le u_n) - \mu(M(B_{i-1} \cup nA_{\ell-1}) \le u_n) \\
&\qquad\quad + r_n\mu(X > u_n)\mu(M(B_{i-1} \cup nA_{\ell-1}) \le u_n)\Big| \\
&\qquad \le \Big|\mu(M(B_i \cup nA_{\ell-1}) \le u_n) - \mu(M(B_{i-1} \cup nA_{\ell-1}) \le u_n)
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{j=0}^{r_n-t_n-1} \mu(\{X_{j+Hn-ir_n} > u_n\} \cap \{M(B_i \cup nA_{\ell-1}) \le u_n)\}\Big| \\
&+ \Big|(r_n - t_n)\mu(X > u_n)\mu(M(B_{i-1} \cup nA_{\ell-1}) \le u_n) \\
&- \sum_{j=0}^{r_n-t_n-1} \mu(\{X_{j+Hn-ir_n} > u_n\} \cap \{M(B_i \cup nA_{\ell-1}) \le u_n)\}\Big| \\
&+ t_n\mu(X > u_n).
\end{aligned}
$$

By Lemma 4.2, we obtain

$$
\begin{aligned}
&\Big|\mu(M(B_i \cup nA_{\ell-1}) \le u_n) - (1 - r_n\mu(X > u_n))\mu\left(M(B_{i-1} \cup nA_{\ell-1}) \le u_n\right)\Big| \\
&\le 2(r_n - t_n) \sum_{j=1}^{r_n-t_n-1} \mu(\{X > u_n\} \cap \{X_j > u_n\}) + t_n\mu(X > u_n) \\
&\quad + \sum_{j=0}^{r_n-t_n-1} \Big|\mu(X > u_n)\mu(M(B_{i-1} \cup nA_{\ell-1}) \le u_n) \\
&\quad - \mu(\{X > u_n\} \cap \{M((B_{i-1} \cup nA_{\ell-1}) - d_j) \le u_n\})\Big| + t_n\mu(X > u_n),
\end{aligned}
$$

where $d_j = (j + Hn - ir_n)$. Now using condition $D_3(u_n)$, we obtain

$$
\begin{aligned}
&\Big|\mu(M(B_i \cup nA_{\ell-1}) \le u_n) - (1 - r_n\mu(X > u_n))\mu(M(B_{i-1} \cup nA_{\ell-1}) \le u_n)\Big| \\
&\le 2(r_n - t_n) \sum_{j=1}^{r_n-t_n-1} \mu(\{X>u_n\}\cap\{X_j>u_n\}) + 2t_n\mu(X > u_n) + (r_n - t_n)\gamma(n, t_n).
\end{aligned}
$$

Set

$$
\begin{aligned}
\Upsilon_{k,n} := {}& 2(r_n - t_n) \sum_{j=1}^{r_n-t_n-1} \mu(\{X > u_n\} \cap \{X_j > u_n\}) + 2t_n\mu(X > u_n) \\
&+ (r_n - t_n)\gamma(n, t_n).
\end{aligned}
$$

Recalling (1.5), we may assume that $n$ and $k$ are sufficiently large so that $\frac{n}{k}\mu(X > u_n) < 2$ and $|1 - r_n\mu(X > u_n)| < 1$ which implies

$$
\begin{aligned}
&\big|\mu(M(B_{S_{p-\ell+1}} \cup nA_{\ell-1}) \le u_n) - (1 - r_n\mu(X > u_n)) \\
&\quad \mu(M(B_{S_{p-\ell+1}-1} \cup nA_{\ell-1}) \le u_n)\big| \le \Upsilon_{k,n},
\end{aligned}
$$

and

$$\begin{aligned}&\Big|\mu(M(B_{S_{p-\ell+1}}\cup nA_{\ell-1})\le u_n)-(1-r_n\mu(X>u_n))^2\mu(M(B_{S_{p-\ell+1}-2}\cup nA_{\ell-1})\le u_n)\Big|\\&\quad\le\big|\mu(M(B_{S_{p-\ell+1}}\cup nA_{\ell-1})\le u_n)-(1-r_n\mu(X>u_n))\mu(M(B_{S_{p-\ell+1}-1}\\&\qquad\cup nA_{\ell-1})\le u_n)\big|+|1-r_n\mu(X>u_n)|\,\big|\mu(M(B_{S_{p-\ell+1}-1}\cup nA_{\ell-1})\le u_n)\\&\qquad-(1-r_n\mu(X>u_n))\mu(M(B_{S_{p-\ell+1}-2}\cup nA_{\ell-1})\big|\\&\quad\le 2\Upsilon_{k,n}.\end{aligned}$$

Inductively, we obtain

$$\begin{aligned}&\Big|\mu(M(B_{S_{p-l+1}}\cup nA_{\ell-1})\le u_n)-(1-r_n\mu(X>u_n))^{S_{p-\ell+1}}\mu(M(nA_{\ell-1})\le u_n)\Big|\\&\quad\le S_{p-\ell+1}\Upsilon_{k,n}.\end{aligned}$$

Using Lemma 4.2,

$$\begin{aligned}&\Big|\mu(M(nA_\ell)\le u_n)-(1-r_n\mu(X>u_n))^{S_{p-\ell+1}}\mu(M(nA_{\ell-1})\le u_n)\Big|\\&\quad\le\big|\mu(M(nA_\ell)\le u_n)-\mu(M(B_{S_{p-l+1}}\cup nA_{\ell-1})\le u_n)\big|\\&\qquad+\Big|\mu(M(B_{S_{p-l+1}}\cup nA_{\ell-1})\le u_n)-(1-r_n\mu(X>u_n))^{S_{p-\ell+1}}\mu(M(nA_{\ell-1})\le u_n)\Big|\\&\quad\le\Big|\mu(M(nI_{p-\ell+1}\cup nA_{\ell-1})\le u_n)-\mu(M(\cup_{i=i_\ell}^{S_{p-\ell+1}}J_i\cup nA_{\ell-1})\le u_n)\Big|\\&\qquad+S_{p-l+1}\Upsilon_{k,n}\\&\quad\le 2r_n\mu(X>u_n)+S_{p-l+1}\Upsilon_{k,n}.\end{aligned}$$

In the next step we have

$$\begin{aligned}&\Big|\mu(M(nA_\ell)\le u_n)-(1-r_n\mu(X>u_n))^{S_{p-\ell+1}+S_{p-\ell+2}}\mu(M(nA_{\ell-2})\le u_n)\Big|\\&\quad\le\Big|\mu(M(nA_\ell)\le u_n)-(1-r_n\mu(X>u_n))^{S_{p-\ell+1}}\mu(M(nA_{\ell-1})\le u_n)\Big|\\&\qquad+|1-r_n\mu(X>u_n)|^{S_{p-\ell+1}}\Big|\mu(M(nA_{\ell-1})\le u_n)\\&\qquad-(1-r_n\mu(X>u_n))^{S_{p-\ell+2}}\mu(M(nA_{\ell-2})\le u_n)\Big|\\&\quad\le 4r_n\mu(X>u_n)+(S_{p-\ell+1}+S_{p-\ell+2})\Upsilon_{k,n}.\end{aligned}$$

Therefore, by induction, we obtain

$$\begin{aligned}&\Big|\mu(M(nA_p)\le u_n)-(1-r_n\mu(X>u_n))^{\sum_{j=1}^p S_j}\Big|\\&\quad\le 2pr_n\mu(X>u_n)+\sum_{j=1}^p S_j\Upsilon_{k,n}.\end{aligned}$$

Now, it is easy to see that $S_j \sim k|I_j|$, for each $j \in \{1, \dots, p\}$. Consequently,

$$\lim_{k\to+\infty}\lim_{n\to+\infty}(1-r_n\mu(X>u_n))^{\sum_{j=1}^p S_j}$$
$$=\lim_{k\to+\infty}\lim_{n\to+\infty}\left(1-\left\lfloor\frac{n}{k}\right\rfloor\mu(X>u_n)\right)^{\sum_{j=1}^p S_j}$$
$$=\lim_{k\to+\infty}\left(1-\frac{\tau}{k}\right)^{\sum_{j=1}^p S_j}=\lim_{k\to+\infty}\left[\left(1-\frac{\tau}{k}\right)^{k\sum_{j=1}^p |I_j|}\right]^{\frac{\sum_{j=1}^p S_j}{k\sum_{j=1}^p |I_j|}}=\mathrm{e}^{-\tau\sum_{j=1}^p |I_j|}$$
$$=\prod_{j=1}^p \mathrm{e}^{-\tau(b_j-a_j)}.$$

To conclude the proof it suffices to show that

$$\lim_{k\to+\infty}\lim_{n\to+\infty}(2pr_n\mu(X>u_n)+kH\Upsilon_{k,n})=0.$$

We start by noting that, since $n\mu(X>u_n)\to\tau\geq 0$,

$$\lim_{k\to+\infty}\lim_{n\to+\infty}2pr_n\mu(X>u_n)=\lim_{k\to+\infty}\frac{2p\tau}{k}=0.$$

Next we need to check that

$$\lim_{k\to+\infty}\lim_{n\to+\infty}k\Upsilon_{k,n}=0,$$

which means,

$$\lim_{k\to+\infty}\lim_{n\to+\infty}2k(r_n-t_n)\sum_{j=1}^{r_n-t_n-1}\mu(\{X>u_n\}\cap\{X_j>u_n\})+2kt_n\mu(X>u_n)$$
$$+k(r_n-t_n)\gamma(n,t_n)=0.$$

Assume that $t=t_n$ where $t_n=o(n)$ is given by Condition $D_3(u_n)$. Now, observe that, by (1.5), for every $k\in\mathbb{N}$, we have $\lim_{n\to\infty}kt_n\mu(X>u_n)=0$. Finally, use $D_3(u_n)$ and $D'(u_n)$ to prove that the two remaining terms also go to 0. □

*Proof of Theorem 5* Since the Poisson Process has no fixed atoms, that is, points $t$ such that $\mu(N(\{t\})>0)>0$, the convergence is equivalent to convergence of finite dimensional distributions. But, because $N$ is a simple point process, without multiple events, we may use a criterion proposed by Kallenberg [29, Theorem 4.7] to show the stated convergence. Namely we need to verify that

(1) $\mathbb{E}(N_n(I))\xrightarrow[n\to\infty]{}\mathbb{E}(N(I))$, for all $I\in\mathcal{S}$;
(2) $\mu(N_n(A)=0)\xrightarrow[n\to\infty]{}\mu(N(A)=0)$, for all $A\in\mathcal{R}$,

where $\mathbb{E}(\cdot)$ denotes the expectation with respect to $\mu$.

First we show that condition (1) holds. Let $a, b \in \mathbb{R}^+$ be such that $I = [a, b)$, then, recalling that $v_n = 1/\mu(X_0 > u_n)$, we have

$$\begin{aligned}\mathbb{E}(N_n(I)) &= \mathbb{E}\left(\sum_{j=\lfloor v_n a\rfloor+1}^{\lfloor v_n b\rfloor} \mathbf{1}_{\{X_j>u_n\}}\right) = \sum_{j=\lfloor v_n a\rfloor+1}^{\lfloor v_n b\rfloor} E\left(\mathbf{1}_{\{X_j>u_n\}}\right)\\ &= \left(\lfloor v_n b\rfloor - (\lfloor v_n a\rfloor + 1)\right)\mu(X_0 > u_n)\\ &\sim (b-a)v_n\mu(X_0 > u_n) \xrightarrow[n\to\infty]{} (b-a) = \mathbb{E}(N(I)).\end{aligned}$$

To prove condition (2), let $s \in \mathbb{N}$ and $A = \cup_{i=1}^{s} I_i$ where $I_1, \ldots, I_s \in \mathcal{S}$ are disjoint. Also let $a_j, b_j \in \mathbb{R}^+$ be such that $I_j = [a_j, b_j)$. By Proposition 1, we have

$$\begin{aligned}\mu(N_n(A) = 0) &= \mu\left(\cap_{i=1}^{s}\{M(v_n I_j) \le u_n\}\right)\\ &\sim \mu\left(\cap_{i=1}^{s}\{M((n/\tau)I_j) \le u_n\}\right) \xrightarrow[n\to\infty]{} \prod_{j=1}^{s} \mathrm{e}^{-(b_j-a_j)}.\end{aligned}$$

The result follows at once since $\mu(N(A) = 0) = \prod_{i=1}^{s} \mu(N(I_j) = 0) = \prod_{j=1}^{s} \mathrm{e}^{-(b_j-a_j)}$. □

*Proof of Corollary 6* In [20,21], conditions $D_2(u_n)$ and $D'(u_n)$ were proved for stochastic processes $X_0, X_1, \ldots$ as in (1.1) and (1.10) with $\zeta$ being either the critical point or the critical value and observables of type $g_3$ ($g_3(x) = x$ for $\zeta = 1$ and $g_3 = 1 - ax^2$ for $\zeta = 0$).

Observe that independently of the type of $g$, the sequence $u_n$ is computed so that an exceedance of the level $u_n$ corresponds to a visit to the ball $B_{\delta_n}(\zeta)$, where $\delta_n$ is such that $\mu(B_{\delta_n}) \sim \tau/n$. This means that condition $D'(u_n)$ can be written in terms of returns to $B_{\delta_n}(\zeta)$ which implies that it holds for every sequence $X_0, X_1, \ldots$ independently of the shape of $g$.

Condition $D_3(u_n)$ follows from decay of correlations. In fact, from [30,35] one has that for all $\phi, \psi : M \to \mathbb{R}$ with bounded variation, there is $C, \alpha > 0$ independent of $\phi, \psi$ and $n$ such that

$$\left|\int \phi \cdot (\psi \circ f^t) d\mu - \int \phi d\mu \int \psi d\mu\right| \le C \mathrm{Var}(\phi) \|\psi\|_\infty \mathrm{e}^{-\alpha t}, \quad \forall t \in \mathbb{N}_0, \tag{4.3}$$

where $\mathrm{Var}(\phi)$ denotes the total variation of $\phi$. For each $A \in \mathcal{R}$, take $\phi = \mathbf{1}_{\{X_0>u_n\}}$ and $\psi = \mathbf{1}_{\{M(A)\le u_n\}}$, then (4.3) implies that Condition $D_3(u_n)$ holds with $\gamma(n, t) = \gamma(t) := 2C\mathrm{e}^{-\alpha t}$ and for the sequence $t_n = \sqrt{n}$, for example. □

## 5 Poisson statistics for first return times

The purpose of this section is to discuss what is known about the Poisson statistics of first return times to balls. The main focus is on showing that a map $f \in NF^2$ with an acip must have the RTPP asymptotically converging to a Poisson Process. However, for more generality we will introduce the ideas assuming that our phase space $\mathcal{X}$ is a Riemannian manifold. We note that a similar result to the main theorem [23] implies that the limit laws for the HTPP and RTPP are the same. So since the results we will cite below are usually given in terms of RTS, we will use this.

Similarly to the proof of Theorem 5, in order to show that the RTPP has a Poisson limit, it suffices to prove that for $k \in \mathbb{N}$ and a rectangle $R_k \subset \mathbb{R}^k$,

$$\left|\mu_{U_n}\left((w_{U_n}, w_{U_n}^2, \ldots, w_{U_n}^k) \in \frac{1}{\mu(U_n)} R_k\right) - \int\limits_{R_k} \Pi_{i=1}^k \mathrm{e}^{t_i}\, dt^k\right| \to 0 \text{ as } n \to \infty.$$

The main result of [7] is that the RTS for an inducing scheme is the same for the inducing scheme as for the original system. However, they remark in that paper that their methods extend to give the same Poisson statistics for the inducing scheme and the original system. In [8], the theory in [7] was extended to show that for multimodal maps of the interval the RTS of suitable inducing schemes converge to the RTS of the original system. The corresponding result for Poisson statistics follows similarly. For multimodal maps $f : I \to I$, with an acip $\mu$, those inducing schemes are Rychlik maps. Therefore to prove that the original $(I, f, \mu)$ has the RTPP converging to a Poisson process, we must show that the induced, Rychlik, maps also have this property. As we sketch below, this can be proved using [26, Theorem 2.6].

For a system $(X, F, \mu)$, we say that a partition $\mathcal{Q}$ is *uniform mixing* if there exists $\gamma_{\mathcal{Q}}(n) \to 0$ as $n \to \infty$, such that

$$\gamma_{\mathcal{Q}}(n) := \sup_{k,l} \sup_{\substack{A \in \sigma\mathcal{Q}_k \\ B \in F^{-(n+k)}\sigma\mathcal{Q}_l}} |\mu(A \cap B) - \mu(A)\mu(B)|\,.$$

Here $\mathcal{Q}_k := \bigvee_{j=0}^{k-1} F^{-j}\mathcal{Q}$ and $\sigma\mathcal{Q}_k$ is the sigma algebra generated by $\mathcal{Q}_k$. For our purposes $\mathcal{Q}$ will be $\{U, U^c\}$ where $U$ is a ball around $\zeta$.

By [26, Theorem 2.6], if we assume the system is uniform mixing for $\{U, U^c\}$, then for a rectangle $R_k \subset \mathbb{R}^k$,

$$\left|\mu_U\left((w_U, w_U^2, \ldots, w_U^k) \in \frac{1}{\mu(U)} R_k\right) - \int\limits_{R_k} \Pi_{i=1}^k \mathrm{e}^{t_i}\, dt^k\right| \le Err(k, U). \qquad (5.1)$$

Moreover, the term $Err(k, U)$ goes to 0 as $U$ shrinks to a point $\zeta$. In fact, we have $Err(k, U) = k(3d(U) + R(k, U))$ where $R(k, U) \to 0$ as $\mu(U) \to 0$ and the rate that $R(k, U)$ goes to zero depends on how $\gamma_Q$ shrinks with $U$. As was shown in [7], for Rychlik maps the quantity $d(U)$ tends 0 as $U \to \{\zeta\}$. Therefore it only remains to show that the Rychlik maps defined in [8] are uniform mixing for $\{U, U^c\}$.

Since we assumed that $(X, F, \mu)$ is Rychlik [34, Theorem 5] implies that the natural partition $\mathcal{P}_1$, consisting of maximal intervals on which $f$ is a homeomorphism, is Bernoulli, with exponential speed. Since $(X, F, \mu)$ is uniformly expanding, this implies that $\{U, U^c\}$ is also Bernoulli, with exponential speed. As noted in [26, Remark 2.5], this implies that $\{U, U^c\}$ is uniform mixing, as required.

The proof that the successive returns form a point process converging to a Poisson Process follows from (5.1) and the Kallenberg argument used in the proof of Theorem 5.

## 6 EVL and HTS in higher dimensions

In this section, we extend Collet's theory of maps with exponential decay of correlations from one dimension to higher dimensions. We conclude with an example.

Let $\mathcal{X}$ be as usual a $d$-dimensional compact Riemannian manifold and $f : \mathcal{X} \to \mathcal{X}$ a $C^2$ endomorphism. We say that $f$ *admits a Young tower* if there exists a ball $\Delta \subset \mathcal{X}$, a countable partition $\mathcal{P}$ (mod 0) of $\Delta$ into topological balls $\Delta_i$ with smooth boundaries, and a return time function $R : \Delta \to \mathbb{N}$ piecewise constant on elements of $\mathcal{P}$ satisfying the following properties:

**(Y$_1$)** *Markov*: for each $\Delta_i \in \mathcal{P}$ and $R = R(\Delta_i)$, $f^R : \Delta_i \to \Delta$ is a $C^2$ diffeomorphism (and in particular a bijection). Thus the induced map

$$F : \Delta \to \Delta \;\text{ given by }\; F(x) = f^{R(x)}(x)$$

is defined almost everywhere and satisfies the classical Markov property. We consider also the *separation time* $s(x, y)$ given by the maximum integer such that $F^i(x)$ and $F^i(y)$ belong to the same element of the partition $\mathcal{P}$ for all $i \leq s(x, y)$, which we assume to be defined and finite for almost every pair of points $x, y \in \Delta$.

**(Y$_2$)** *Uniform backward contraction*: There exist $C > 0$ and $0 < \beta < 1$ such that for $x, y \in \Delta$ and any $0 \leq n \leq s(x, y)$ we have

$$\mathrm{dist}(f^n(x), f^n(y)) \leq C\beta^{s(x,y)-n}.$$

**(Y$_3$)** *Bounded distortion*: For any $x, y \in \Delta$ and any $0 \leq k \leq n < s(x, y)$ we have

$$\log \prod_{i=k}^{n} \frac{\det Df(f^i(x))}{\det Df(f^i(x))} \leq C\beta^{s(x,y)-n}$$

**(Y$_4$)** *Integrable return times*:

$$\int R \, d\mathrm{Leb} < \infty$$

In this section we only consider maps admitting a Young tower *with exponential return time tail* which means that we will replace condition **(Y$_4$)** by the following stronger one

**($\mathbf{Y}_4'$)** *Exponential tail decay*: There are $C, \alpha > 0$ such that

$$\mathrm{Leb}(\{R > n\}) = C\mathrm{e}^{-\alpha n}.$$

These systems have been studied, in a more general context, by Young [36,37], where several examples can also be found. Among the properties proved by L.S. Young we mention the existence of an $F$-invariant measure $\mu_0$ that is equivalent to Lebesgue measure on $\Delta$ (meaning that its density is bounded above and below by a constant). After saturating one gets an absolutely continuous (w.r.t. Lebesgue), $f$-invariant probability given by

$$\mu(A) = \bar{R}^{-1} \sum_{\ell=0}^{\infty} \mu_0 \left( f^{-\ell}(A) \cap \{R > \ell\} \right), \tag{6.1}$$

where $\bar{R} = \int_\Delta R d\mu_0$. One of the main achievements in [36,37] is the fact that the decay of the tail of return times determines the speed of decay of correlations for Hölder continuous (or Lipschitz) observables. Namely, if $\phi : \mathcal{X} \to \mathbb{R}$ is Hölder continuous of exponent $0 < \iota \leq 1$, with *Hölder constant*

$$K_\iota(\phi) := \sup_{x \neq y} \frac{|\phi(x) - \phi(y)|}{(\mathrm{dist}(x, y))^\iota},$$

$\psi : \mathcal{X} \to \mathbb{R}$ is in $L^\infty(\mathrm{Leb})$ and the tower has exponential tail, then there are $C > 0$ and $\alpha' > 0$ such that

$$\left| \int \phi \cdot (\psi \circ f^t) d\mu - \int \phi d\mu \int \psi d\mu \right| \leq C K_\iota(\phi) \|\psi\|_\infty \mathrm{e}^{-\alpha t}, \quad \forall t \in \mathbb{N}_0. \tag{6.2}$$

**Theorem 6** *Let $\mathcal{X}$ be a $d$-dimensional compact Riemannian manifold and assume that $f : \mathcal{X} \to \mathcal{X}$ is a $C^2$ endomorphism admitting a Young tower with exponential tail. Consider a stochastic process $X_0, X_1, \ldots$ defined by* (1.1) *and* (1.10)*, for some choice of $\zeta \in \mathcal{X}$. Then, for* Leb*-almost every $\zeta \in \mathcal{X}$ chosen, conditions $D_3(u_n)$ (or $D_2(u_n)$) and $D'(u_n)$ hold, where $u_n$ is a sequence of levels satisfying* (1.5).

Together with the results in Sects. 1.3, 3 and 4 we get the following corollary.

**Corollary 7** *Let $\mathcal{X}$ be a $d$-dimensional compact Riemannian manifold and assume that $f : \mathcal{X} \to \mathcal{X}$ is a $C^2$ endomorphism admitting a Young tower with exponential tail. Consider a stochastic process $X_0, X_1, \ldots$ defined by* (1.1) *and* (1.10) *for some $\zeta \in \mathcal{X}$. Then, for* Leb*-almost every choice of $\zeta \in \mathcal{X}$, the following assertions hold*:

(1) *We have an EVL for $M_n$, defined in* (1.2)*, which coincides with that one of $\hat{M}_n$ defined in* (1.6)*. In particular, it must be of one of the three classical types. Moreover, for every $i \in \{1, 2, 3\}$, if $g$ is of type $g_i$ then we have an EVL for $M_n$ of type $EV_i$.*
(2) *We have exponential HTS to balls at $\zeta \in \mathcal{X}$.*

(3) *The EPP $N_n$ defined in* (3.3) *is such that* $N_n \xrightarrow{d} N$, *as* $n \to \infty$, *where N denotes a Poisson Process with intensity* 1.
(4) *The same applies to the HTPP* $N_n^*$ *defined in* (3.2).

## 6.1 Proof of Theorem 6

To show this result, one needs only to realise that Collet's proof of [14, Theorem 1] may be mimicked in our multi-dimensional setting with minor adjustments. Thus, instead of repeating all the arguments, we will prove that $D_3(u_n)$ and $D'(u_n)$ hold just by redoing the parts that need to be adapted to this more general higher dimensional setting.

Let $\{E_v\}_{v\in\mathbb{N}}$ be the sequence of sets defined by

$$E_v = \left\{y:\ \exists j \in \{1,\ldots,(\log v)^5\},\ |y - f^j(y)| \leq v^{-1}\right\}\ .$$

This is the set of points which recur 'too quickly'. We not only need to control the set of points which recur too quickly, but also the set of points for which a neighbour recurs too quickly. For positive numbers $\omega$ and $\rho$ to be fixed below, we define a sequence of measurable sets $\{F_v\}_{v\in\mathbb{N}}$ by

$$F_v = \left\{x:\ \mu\big(B_{v^{-\omega}}(x) \cap E_{v^\omega}\big) \geq \kappa\, v^{-(d+\rho)\omega}\right\}\ .$$

The following proposition gathers most of the information we need to prove Theorem 6. The proof of its statements can be done by adapting slightly the proofs of Lemma 2.2, Proposition 2.3, Corollary 2.4 and Lemma 2.5 from [14]. Note that strictly speaking, items (b) and (c) are not used explicitly for the proof of Theorem 6, but we include them here for comparison with [14], as well as [27].

**Proposition 2** *Under* ($\mathbf{Y_4'}$),

(a) *there are two positive constants C and* $\theta$ *such that for any Lebesgue measurable set A, we have*

$$\mu(A) \leq C\mathrm{Leb}(A)^\theta.$$

(b) *there exist positive constants* $C$, $\alpha'$ *and* $\eta < 1$ *such that for any integer* $v$ *and any* $\epsilon > 0$ *we have*

$$\mu(\mathcal{E}_v(\epsilon)) \leq C\left(v^2\epsilon^\eta + e^{-\alpha' v}\right).$$

(c) *there exist positive constants* $C'$ *and* $\beta' < 1$ *such that for any integer* $v$

$$\mu(E_v) \leq C'\, v^{-\beta'}\ .$$

(d) *there exist positive numbers* $\rho$ *and* $\omega$ *such that* $\mathrm{Leb}(\cap_{i\geq 1}\cup_{v\geq i} F_v) = 0$.

As we have seen in the proof of Corollary 6, it is very easy to show that $D_3(u_n)$ holds when we have decay of correlations for observables of bounded variation. However, in this setting, decay of correlations is only available for Hölder continuous functions against $L^\infty$ ones, instead (see (6.2)). This means that we cannot use the test function $\phi = \mathbf{1}_{\{X_0>u_n\}}$, as we did before. However, proceeding as in [14, Lemma 3.3], if we use a suitable Hölder approximation one can easily get:

**Lemma 6.1** *Assume that there exists a rate function* $\Theta : \mathbb{N} \to \mathbb{R}$, *such that for every Hölder continuous (or Lipschitz) observable* $\phi$ *and all* $L^\infty$ *observable* $\psi$ *we have*:

$$\left|\int \phi \cdot (\psi \circ f^t) d\mu - \int \phi d\mu \int \psi d\mu\right| \leq K_\iota(\phi)\|\psi\|_\infty \Theta(t), \quad \forall t \in \mathbb{N}_0.$$

*Then, for every* $\zeta \in \mathcal{X}$, $0 < s < 1$, $\eta > 0$ *and all measurable set* $W$ *we have*

$$\left|\mu(B_s(\zeta) \cap f^{-t}(W)) - \mu(B_s(\zeta))\mu(W)\right| \leq s^{-(1+\eta)}\Theta(t) + O(s^{\theta(d+\eta)}),$$

*where* $\theta$ *is the number given in Proposition* 2 (a).

*Proof of Theorem 6* First let us show that $D_3(u_n)$ holds. Since in this setting we have exponential decay of correlations for Hölder continuous functions (see (6.2)) and $\{X_0 > u_n\} = B_{g^{-1}(u_n)}(\zeta)$ then by Lemma 6.1 we may take

$$\gamma(n,t) = O\left((g^{-1}(u_n))^{-1-\eta} \mathrm{e}^{-\alpha t} + (g^{-1}(u_n))^{\theta(d+\eta)}\right).$$

Hence, recalling that $g^{-1}(u_n) \sim (\frac{\tau}{\kappa\rho(\zeta)n})^{1/d}$, if we consider $t_n = \sqrt{n}$, for example, and choose $\eta$ from Lemma 6.1 so that $\theta(d+\eta)/d > 2$ (where $\theta$ is given by Proposition 2 (a)), then we easily get that $n\gamma(n,t_n) \xrightarrow[n\to\infty]{} 0$ which gives $D_3(u_n)$.

Now, it only remains to show that $D'(u_n)$ also holds. Recall the stochastic process $X_0, X_1, \ldots$ given by (1.1) for observables defined by (1.10), achieving a global maximum at $\zeta \in \mathcal{X}$. At this point, we describe the full Lebesgue measure set of points $\zeta \in \mathcal{X}$ for which Theorem 6 holds. We take $\zeta$ for which Lebesgue's differentiation theorem holds (with respect to the measure $\mu$) and $\zeta \in \cup_{i\geq 1} \cap_{j \geq i} \mathcal{X} \setminus F_j$, which by Proposition 2 (d) is also a full Lebesgue measure set. For each such $\zeta$, let $v_0(\zeta) \in \mathbb{N}$ be such that $\zeta \notin F_j$ for all $j \geq v_0(\zeta)$.

We consider a turning instant $t = t(n) = \lfloor (\log n)^2 \rfloor$, and split the sum in $D'(u_n)$ into the period before $t$ and after $t$.

For the later we use exponential decay of correlations (6.2) and Lemma 6.1 to get, for some $C > 0$,

$$\begin{aligned} S_2(t,n,k) &:= n \sum_{j=t}^{\lfloor n/k \rfloor} \mu(\{X_0 > u_n\} \cap \{X_j > u_n\}) \\ &\leq n \left\lfloor \frac{n}{k} \right\rfloor \mu(X_0 > u_n)^2 + n \left\lfloor \frac{n}{k} \right\rfloor (g^{-1}(u_n))^{\theta(d+\eta)} \\ &\quad + n \left\lfloor \frac{n}{k} \right\rfloor (g^{-1}(u_n))^{-1-\eta} C \mathrm{e}^{-\alpha t}. \end{aligned}$$

Recalling that $\mu(X_0 > u_n) \sim \tau n^{-1}$ and $g^{-1}(u_n) \sim (\frac{\tau}{\kappa\rho(\zeta)n})^{1/d}$, we have

$$S_2(t,n,k) = O\left(\frac{1}{k} + \frac{n^2}{k} n^{-\theta(d+\eta)/d} + \frac{n^2}{k} n^{(1+\eta)/d} \mathrm{e}^{-\alpha' \log^2(n)}\right).$$

So, if we chose $\eta$ so that $\theta(d+\eta)/d > 2$ then $\lim_{k\to\infty} \limsup_{n\to\infty} S_2(t,n,k) = 0$.

We are left with the first period from 1 to $t$ and the respective sum

$$S_1(t,n) := n \sum_{j=1}^{t} \mu(\{X_0 > u_n\} \cap \{X_j > u_n\}).$$

We set $\upsilon = \upsilon(n) = \lfloor (3g^{-1}(u_n))^{-1/\omega} \rfloor$, where $\omega$ is given in Proposition 2 (d). Observe that

$$\{X_0 > u_n\} = B_{g^{-1}(u_n)}(\zeta) \subset B_{\upsilon^{-\omega}}(\zeta)$$

and, if $y \in \{X_0 > u_n\} \cap \{X_j > u_n\}$, then

$$\mathrm{dist}(f^j(y), y) \leq \mathrm{dist}(f^j(y), \zeta) + \mathrm{dist}(\zeta, y) \leq 2g^{-1}(u_n) < \upsilon^{-\omega},$$

which implies that

$$\{X_0 > u_n\} \cap \{X_j > u_n\} \subset B_{\upsilon^{-\omega}}(\zeta) \cap E_{\upsilon^{\omega}}. \tag{6.3}$$

We take $n$ so large that $\upsilon = \upsilon(n) \geq \upsilon_0(\zeta)$. Hence $\zeta \notin F_\upsilon$. Using (6.3), the definition of $F_\upsilon$ and the fact $g^{-1}(u_n) \sim (\frac{\tau}{\kappa\rho(\zeta)n})^{1/d}$, we have

$$\mu(\{X_0 > u_n\} \cap \{X_j > u_n\}) = O(\upsilon^{-\omega(d+\rho)}) = O(n^{-(d+\rho)/d}).$$

Hence, $\limsup_{n\to\infty} S_1(t,n) \leq \limsup_{n\to\infty} O(n \log^2(n) n^{-(d+\rho)/d}) = 0$. □

### 6.2 An example

Here we present a $C^1$ open class of local diffeomorphisms with no critical points that are non-uniformly expanding in the sense of [3,4]. Namely, let $f : M \to M$ be a $C^1$ local diffeomorphism, we say that $f$ is non-uniformly expanding if there exists a constant $\lambda > 0$ such that for Lebesgue almost all points $x \in M$ the following *non-uniform expansivity* condition is satisfied:

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \log \|Df^{-1}_{f^i(x)}\|^{-1} \geq \lambda > 0. \tag{6.4}$$

Condition (6.4) implies that the *expansion time* function

$$\mathcal{E}(x) = \min\left\{N : \frac{1}{n}\sum_{i=0}^{n-1}\log\|Df^{-1}_{f^i(x)}\|^{-1} \geq \lambda/2 \ \ \forall n \geq N\right\}$$

is defined and finite almost everywhere in $M$. We think of this as the waiting time before the exponential derivative growth kicks in. We are now able to define the *Hyperbolic tail set*, at time $n \in \mathbb{N}$,

$$\Gamma_n = \{x \in I : \ \mathcal{E}(x) > n\}, \tag{6.5}$$

which can be seen as the set of points that at time $n$ have not reached a satisfactory exponential growth of the derivative. Applying [5] and [22] together shows that these maps admit a Young tower whose return time tail is related to the volume decay rate of the hyperbolic tail set.

The class we consider here is obtained by deformation of a uniformly expanding map by isotopy inside some small region. In general, these maps are not expanding: deformation can be made in such way that the new map has periodic saddles. We follow the construction in [3,4].

Let $M$ be any compact Riemannian $d$-dimensional manifold supporting some uniformly expanding map $f_0$: there exists $\sigma_0 > 1$ such that $\|Df_0(x)v\| > \sigma_0\|v\|$ for every $x \in M$ and every $v \in T_xM$. Let $V \subset M$ be small compact domain, so that $f_0|_V$ is one-to-one. Let $f_1$ be a $C^1$ map coinciding with $f_0$ in $M\backslash V$ for which the following holds:

(1) $f_1$ is *volume expanding everywhere*: there is $\sigma_1 > 1$ such that

$$|\det Df_1(x)| > \sigma_1, \quad \text{for every } x \in M;$$

(2) $f_1$ is *not too contracting on* $V$: there is small $\delta > 0$ such that

$$\|Df_1(x)^{-1}\| < 1 + \delta, \quad \text{for every } x \in V.$$

We consider the class of maps $f$ in a small $C^1$-neighbourhood $\mathcal{F}$ of $f_1$.

In [3, Section 6] it was shown that these maps satisfy condition (6.4) and there exist $C, \gamma > 0$ such that $\mathrm{Leb}(\Gamma_n) \leq C\mathrm{e}^{-\gamma n}$ for all $n \in \mathbb{N}$. Now, using the results in [22] this implies that every map $f \in \mathcal{F}$ admits a Young tower for which conditions **($\mathbf{Y_1}$)**–**($\mathbf{Y_4'}$)** are satisfied. This means that we can apply Theorem 6 and obtain that all assertions of Corollary 7 hold for this class of maps $\mathcal{F}$.

To make this example a bit more concrete, we give the following construction in the 2-torus. Consider the doubling map in $S^1$, or in other words, let $g : S^1 \to S^1$ be such that $g(x) = 2x$ if $x \in [0, 1/2]$ and $g(x) = 2x - 1$ if $x \in (1/2, 1]$. Take $f_0 : S^1 \times S^1 \to S^1 \times S^1$ to be the uniformly expanding map given by $f_0(x, y) = (g(x), g(y))$.

Let $k : [0.05, 0.1] \to \mathbb{R}$ be such that $k(0.05) = 0.045$, $k'(0.05) = 0.9$, $k(0.1) = 0.2$, $k'(0.1) = 2$, $k'(x) \geq 0.9$, for all $x \in [0.05, 0.1]$ and $k'(x^*) > 1$, where $x^*$ is the

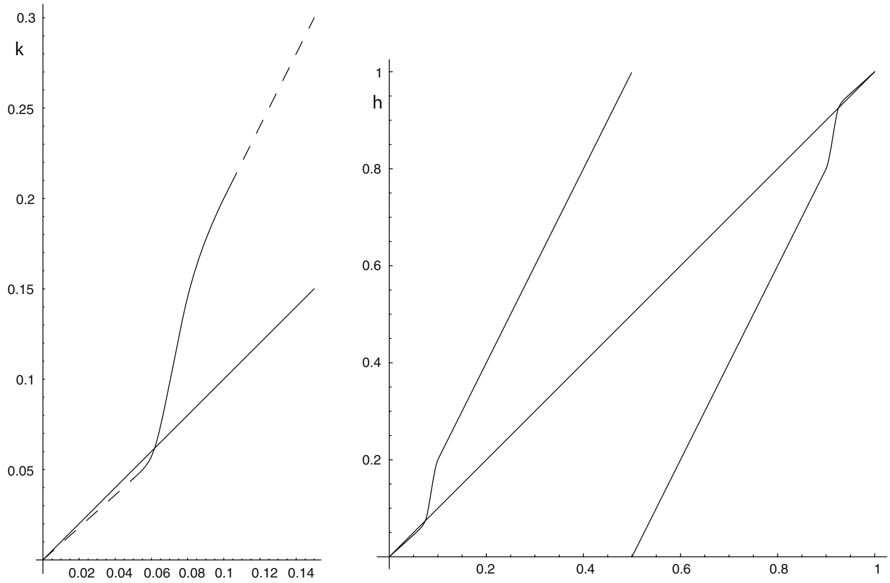


**Fig. 2** Plots of functions $k$ and $h$

only solution of $k(x) = x$. See the plot on the left of Fig. 2. Now, let $\ell : [0, 1/2] \to [0, 1]$ be given by

$$\ell(x) = \begin{cases} 0.9x & \text{if } x \in [0, 0.05) \\ k(x) & \text{if } x \in [0.05, 0.1) \\ 2x & \text{if } x \in [0.1, 1/2], \end{cases}$$

and define $h : S^1 \to S^1$ by $h(x) = \ell(x)$ if $x \in [0, 1/2]$ and $h(x) = 1 - \ell(1 - x)$ if $x \in (1/2, 1]$. See plot on the right of Fig. 2.

Consider a $C^1$ family of maps $\psi : S^1 \times [0, 1] \to S^1$ such that $\psi_0(x) = \psi(x, 0) = g(x)$, $\psi_1(x) = \psi(x, 1) = h(x)$, for all $x \in S^1$ and $\frac{\partial \psi(x,t)}{\partial x}|_{(x,t)} \geq 0.9$, for all $x \in S^1$, $t \in [0, 1]$. Finally, consider $f_1 : S^1 \times S^1 \to S^1 \times S^1$ given by

$$f_1(x, y) = \left(\psi\left(x, (1 - 10y)\mathbf{1}_{[0,0.1]} + (1 - 10(1 - y))\mathbf{1}_{[0.9,1]}\right), g(y)\right).$$

Taking $V = [-0.1, 0.1] \times [-0.1, 0.1] \subset S^1 \times S^1$, we have $f_0$ is one-to-one in $V$, $f_1 = f_0$ outside $V$, $f_1$ is not uniformly expanding since 0 is a saddle and conditions (1), (2) are also easily checked.

**Acknowledgments** We would like to thank J.F. Alves for useful suggestions regarding the example of a non-uniformly expanding system given in Sect. 6.2. We thank the referee for useful comments.

## References


1. Abadi, M.: Sharp error terms and necessary conditions for exponential hitting times in mixing processes. Ann. Probab. **32**, 243–264 (2004)
2. Abadi, M., Galves, A.: Inequalities for the occurrence times of rare events in mixing processes. The state of the art. Markov Process. Relat. Fields **7**, 97–112 (2001)
3. Alves, J.F.: Strong statistical stability of non-uniformly expanding maps. Nonlinearity **17**, 1193–1215 (2004)
4. Alves, J.F., Bonatti, C., Viana, M.: SRB measures for partially hyperbolic systems whose central direction is mostly expanding. Invent. Math. **140**, 351–398 (2000)
5. Alves, J.F., Luzzatto, S., Pinheiro, V.: Markov structures and decay of correlations for non-uniformly expanding dynamical systems. Ann. Inst. H. Poincaré Anal. Non Linéaire **22**, 817–839 (2005)
6. Benedicks, M., Carleson, L.: On iterations of $1 - ax^2$ on $(-1, 1)$. Ann. Math. **122**, 1–25 (1985)
7. Bruin, H., Saussol, B., Troubetzkoy, S., Vaienti, S.: Return time statistics via inducing. Ergod. Theory Dyn. Syst. **23**, 991–1013 (2003)
8. Bruin, H., Todd, M.: Return time statistics for invariant measures for interval maps with positive Lyapunov exponent. Stoch. Dyn. **9**, 81–100 (2009)
9. Bruin, H., Vaienti, S.: Return time statistics for unimodal maps. Fund. Math. **176**, 77–94 (2003)
10. Coelho, Z.: Asymptotic laws for symbolic dynamical systems. Topics in Symbolic Dynamics and Applications (Temuco 1997). LMS Lecture Notes Series, vol. 279, pp. 123–165. Cambridge University Press, London (2000)
11. Coelho, Z.: The loss of tightness of time distributions for homeomorphisms of the circle. Trans. Am. Math. Soc. **356**, 4427–4445 (2004)
12. Coelho, Z., de Faria, E.: Limit laws of entrance times for homeomorphisms of the circle. Israel J. Math. **93**, 93–112 (1996)
13. Collet, P.: Some ergodic properties of maps of the interval. Dynamical Systems (Temuco, 1991/1992), (Travaux en cours, 52), pp. 55–91. Herman, Paris (1996)
14. Collet, P.: Statistics of closest return for some non-uniformly hyperbolic systems. Ergod. Theory Dyn. Syst. **21**, 401–420 (2001)
15. Denker, M., Gordin, M., Sharova, A.A.: Poisson limit theorem for toral automorphisms. Ill. J. Math. **48**, 1–20 (2004)
16. Denker, M., Philipp, W.: Approximation by Brownian motion for Gibbs measures and flows under a function. Ergod. Theory Dyn. Syst. **4**, 541–552 (1984)
17. Dolgopyat, D.: Limit theorems for partially hyperbolic systems. Trans. Am. Math. Soc. **356**, 1637–1689 (2004)
18. Dudewicz, E.J., Mishra, S.N.: Modern Mathematical Statistics. Wiley Series in Probability and Mathematical Statistics. Wiley, London (1988)
19. Feller, W.: An introduction to Probability Theory and its Applications, vol. I. Wiley Publications in Statistics, New York (1952)
20. Freitas, A.C.M., Freitas, J.M.: Extreme values for Benedicks Carleson maps. Ergod. Theory Dyn. Syst. **28**(4), 1117–1133 (2008)
21. Freitas, A.C.M., Freitas, J.M.: On the link between dependence and independence in extreme value theory for dynamical systems. Stat. Probab. Lett. **78**, 1088–1093 (2008)
22. Gouëzel, S.: Decay of correlations for nonuniformly expanding systems. Bull. Soc. Math. France **134**, 1–31 (2006)
23. Haydn, N., Lacroix, Y., Vaienti, S.: Hitting and return times in ergodic dynamical systems. Ann. Probab. **33**, 2043–2050 (2005)
24. Haydn, N., Vaienti, S.: The compound Poisson distribution and return times in dynamical systems. Probab. Theory Relat. Fields **144**, 517–542 (2009)
25. Hirata, M.: Poisson law for axiom a diffeomorphisms. Ergod.Theory Dyn. Syst. **13**, 533–556 (1993)
26. Hirata, M., Saussol, B., Vaienti, S.: Statistics of return times: a general framework and new applications. Comm. Math. Phys. **206**, 33–55 (1999)
27. Holland, M., Nicol, M., Torok, A.: Extreme value distributions for non-uniformly hyperbolic dynamical systems (preprint) http://www.math.uh.edu/~nicol/papers.html
28. Hsing, T., Hüsler, J., Leadbetter, M.R.: On the exceedance point process for a stationary sequence. Probab. Theory Relat. Fields **78**, 97–112 (1988)
29. Kallenberg, O.: Random Measures. Academic Press Inc., New York (1986)

30. Keller, G., Nowicki, T.: Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps. Comm. Math. Phys. **149**, 31–69 (1992)
31. Lindgren, G., Leadbetter, M.R., Rootzén, H.: Extremes and related properties of random sequences and processes. Springer Series in Statistics. Springer, New York (1983)
32. Pesin, Y.: Dimension Theory in Dynamical Systems. University of Chicago Press, Chicago (1997)
33. Pitskel, B.: Poisson limit law for Markov chains. Ergod. Theory Dynam. Syst. **11**, 501–513 (1991)
34. Rychlik, M.: Bounded variation and invariant measures. Stud. Math. **76**, 69–80 (1983)
35. Young, L.S.: Decay of correlations for certain quadratic maps. Comm. Math. Phys. **146**, 123–138 (1992)
36. Young, L.S.: Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. **147**(2), 585–650 (1998)
37. Young, L.S.: Recurrence times and rates of mixing. Isr. J. Math. **110**, 153–188 (1999)