EXTREME VALUE THEORY FOR LORENZ MAPS AND FLOWS, HENÓN-LIKE DIFFEOMORPHISMS AND A CLASS OF HYPERBOLIC SYSTEMS WITH SINGULARITIES

CHINMAYA GUPTA, MARK HOLLAND, AND MATTHEW NICOL

Abstract. We establish that Hölder observations on certain non-uniformly hyperbolic diffeomorphisms and on certain hyperbolic systems with singularities exhibit the same extreme value statistics as iid processes with the same distribution function. Dynamical systems to which our results apply include Lozi-like maps, hyperbolic billiards, Lorenz maps and Henón diffeomorphisms. Suspension flows of such systems, including the Lorenz flow, exhibit the same extreme laws.

1. Introduction

Suppose $\{X_n\}$ is a stationary stochastic process and define $\{M_n\}$, the sequence of successive maxima by $M_n := \max\{X_1, \ldots, X_n\}$. Analogously, if $\{X_t\}$ is a continuous time stochastic process define $M_T := \sup_{0 \le t \le T}\{X_t\}$. In order to simplify the discussion we focus in this introduction on the discrete-time case; the ideas presented for discrete time have straightforward counterparts for continuous time.

There is a well developed theory [12, 7, 18] assuming $\{X_n\}$ are independent for the limiting distribution of $\{M_n\}$ under linear scaling $a_n(M_n - b_n)$ defined by constants $a_n > 0, b_n \in \mathbb{R}$. It is known that there are only three non-degenerate distributions $G(x)$ such that $\lim_{n\to\infty} P(a_n(M_n - b_n) \le x) = G(x)$ (up to location $G(x) \to G(x + b)$ and scale $G(x) \to G(ax)$, $a > 0$, changes). These distributions are called extreme type distributions, Type I, II or III [12]. We say a stationary process $\{X_n\}$ satisfies the law of types if, when under linear scaling the successive maxima $\{M_n\}$ converge to a non-degenerate distribution then the distribution is a Type I, II or III distribution. We recall the form of these extremal distributions:

Type I

$$G(x) = e^{-e^{-x}}, \qquad -\infty < x < \infty.$$

Type II

$$G(x) = \begin{cases} 0 & \text{if } x \le 0; \\ e^{-x^{-\alpha}} & \text{for some } \alpha > 0 \text{ if } x > 0. \end{cases}$$

Type III

$$G(x) = \begin{cases} e^{-(-x)^{\alpha}} & \text{for some } \alpha > 0 \text{ if } x \le 0; \\ 1 & \text{if } x > 0. \end{cases}$$

Date: May 22, 2009.

2000 *Mathematics Subject Classification.* Primary 37A50; Secondary 60G70.

Key words and phrases. $D(u_n)$, $D'(u_n)$, non-uniformly hyperbolic systems, Lozi maps, hyperbolic billiards, Lorenz equations, Henón diffeomorphisms.

The research of CG and MN was supported in part by the National Science Foundation grants DMS-0607345 and DMS-0600927. MH thanks the University of Houston for hospitality.

If $\{X_n\}$ is a stationary sequence we let $\{\hat{X}_n\}$ denote the associated stationary, independent sequence, that is $\{\hat{X}_n\}$ is independent and $\hat{X}_1$ has the same distribution as X_1. For $x \in \mathbb{R}$ and sequences a_n, b_n we define $u_n(x) = \frac{x}{a_n} + b_n$ so that $P(M_n \leq u_n(x)) = P(a_n(M_n - b_n) \leq x)$. For fixed x we will often drop the dependence upon x and write simply the sequence u_n. Leadbetter [12] gives two conditions called $D(u_n)$ and $D'(u_n)$ for suitable sequences u_n which imply that $P(a_n(M_n - b_n) \leq x) \to G(x)$ is equivalent to $P(a_n(\hat{M}_n - b_n) \leq x) \to G(x)$. If $\{X_n\}$ satisfies condition $D(u_n)$ then the stochastic process satisfies the law of types. Moreover it is known that for $u_n = x/a_n + b_n$, $nP(X_0 \geq u_n) \to \tau$ is equivalent to $P(\hat{M}_n \leq u_n) \to e^{-\tau}$ when $D(u_n)$ and $D'(u_n)$ hold, and hence we have a strategy for determining the extreme type distribution for dependent sequences.

There are, however, no general techniques for proving conditions $D(u_n)$ and $D'(u_n)$ and the latter is usually hard. Collet [2] in an elegant paper used the rate of decay of correlation of Hölder observations to establish $D(u_n)$ for certain one-dimensional non-uniformly hyperbolic maps. Freitas et al [4], based on Collet's work, in turn gave a condition $D_2(u_n)$ which has the full force of $D(u_n)$ in that together with $D'(u_n)$ it ensures the equivalence of $P(a_n(M_n - b_n) \leq x) \to G(x)$ and $P(a_n(\hat{M}_n - b_n) \leq u_n) \to G(x)$. Condition $D_2(u_n)$ is easier to establish in the dynamical setting by estimating the rate of decay of correlations of Hölder continuous observables or those of bounded variation. We establish condition $D_2(u_n)$ in this paper for the time-series of certain observations on non-uniformly hyperbolic diffeomorphisms of a D dimensional manifold modeled by a Young tower [23]. A main contribution is that we extend Collet's approach to handle dynamical systems with stable foliations. We also establish condition $D'(u_n)$ for certain systems (hyperbolic billiards, Lozi maps, Hénon diffeomorphisms, Lorenz maps and flows) and hence show that from the point of view of extreme value theory such systems behave as iid processes.

We now state conditions $D_2(u_n)$ and $D'(u_n)$. If $\{X_n\}$ is a stochastic process define

$$M_{j,l} := \max\{X_j, X_{j+1}, \dots, X_{j+l}\}.$$

We will often write $M_{0,n}$ as M_n.

Condition $D_2(u_n)$ [4] We say condition $D_2(u_n)$ holds for the sequence $X_0, X_1, \dots,$ if for any integers l,t and n

$$|\mu(X_0 > u_n, M_{t,l} \leq u_n) - \mu(X_0 > u_n)\mu(M_l \leq u_n)| \leq \gamma(n,t)$$

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

Condition $D'(u_n)$ [12] We say condition $D'(u_n)$ holds for the sequence $X_0, X_1, \dots,$ if

$$\lim_{k\to\infty} \limsup_n n \sum_{j=1}^{[n/k]} \mu(X_0 > u_n, X_j > u_n) = 0.$$

Collet [2] demonstrated a technique involving maximal functions for establishing $D'(u_n)$ for one dimensional non-uniformly expanding maps modeled by a Young tower. His argument relies heavily on the absence of a stable direction and the boundedness of the derivative and these are obstacles to generalizing his argument. The one-dimensional feature can be generalized to expanding maps in higher dimension [6].

In this paper we assume T to be a non-uniformly hyperbolic diffeomorphism of a D dimensional manifold modeled by a Young tower [23] with SRB measure μ and establish

$D_2(u_n)$ for the process $X_n(x) = -\log(d(x_0, T^n x))$. For these systems, if $D'(u_n)$ can be verified, then the process has the same extreme value statistics as its associated iid process, even for more general observations. The following observation is straightforward, see [5] [10, Lemma 1.3]:

Lemma 1.1. *Assume $g(x)$ is a function with a unique minimum value of zero at x_0 (we have in mind the function $g(x) = d(x, x_0)$).*

The following are equivalent, where $\alpha > 0$:

(1) *A Type I law for $x \mapsto -\log g(x)$ with $a_n = 1$ and $b_n = \log n$;*
(2) *A Type II law for $x \mapsto g(x)^{-\alpha}$ with $a_n = n^{-\alpha}$ and $b_n = 0$;*
(3) *A Type III law for $x \mapsto \mathcal{C} - g(x)^{\alpha}$ with $a_n = n^{\alpha}$ and $b_n = \mathcal{C}$;*

We do not have a general method to establish $D'(u_n)$ for all systems modeled by a Young Tower, but we identify a class of two-dimensional maps for which the condition will hold. Our method of proof for $D'(u_n)$ in these cases is an extension of the argument in Collet [2] and in the case of Hénon diffeomorphisms uses a result of Melbourne et al [17] on large deviations. While the condition $D_2(u_n)$ requires only mild assumptions on the tails of the tower, our method of proof of $D'(u_n)$ in this paper requires exponential tails for the Young Tower. We note that Poisson-limit laws for return-time statistics in the Axiom-A setting have been established by Hirata [9] and in the uniformly partially hyperbolic setting by Dolgopyat [3]. For recent related work on extreme value theory for deterministic dynamical systems see [5, 4, 6, 10]. Our results also have implication for the hitting time statistics for such systems which we describe in Section 1.2.

1.1. **Statement of results.** Let M be a Riemannian manifold of dimension 2, with Lebesgue measure m and let $T : M \to M$ be a (local) diffeomorphism modeled by a Young Tower. We will assume in addition to the usual structure of a Young Tower that the density h of the T invariant SRB measure μ is in $L^{1+\sigma}(m)$. The Young Tower assumption implies that there exists a subset $\Lambda \subset M$ such that Λ has a hyperbolic product structure and that (P1)-(P4) of [23] hold.

By taking T to be a local diffeomorphism we allow the map T to have discontinuities or its derivative to have singularities. We let S denote the singular set for T (which could be empty) and let $\mathcal{S} = \bigcup_k T^{-k} S$. The set S will either be a finite set of points or a finite collection of smooth curves. The following definition specifies the degree of hyperbolicity we require in the presence of singularities.

Definition 1.1. We say that a local diffeomorphism T is *hyperbolic with singularities* if the following hold. Let DT_u be the Jacobian restricted to the unstable direction and let $\mathcal{S}$ denote the singularity set (together with its preimages) for T. Then

(1) there exists $C > 0$ and $\lambda > 1$ such that for all $x \in M \setminus \mathcal{S}$ and $n > 0$, $|DT_u^n(x)| \geq C\lambda^n$.
(2) the derivative $DT \in L^1(\mu)$.
(3) tangent vectors to local unstable manifolds lie in a strict cone with slope bounded away from 0 and ∞.

Theorem 1.1. *Let $T : M \to M$ be a two-dimensional (local) hyperbolic diffeomorphism modeled by a Young tower with exponential decay of correlations. Suppose that T is hyperbolic with singularities in the sense of Definition 1.1. Then for μ a.e. x_0 the stochastic process defined by $X_n(x) = -\log(d(x_0, T^n x))$ satisfies a Type I extreme value law.*

A central ingredient in the proof of Theorem 1.1 is Theorem 3.1. Applications of Theorem 3.1 will be considered in Sections 4.2, 4.3 and 4.4 to Lozi maps, hyperbolic billiards, and Lorenz maps respectively. We will show that the assumptions of Theorem 1.1 hold for these systems but defer the description of the maps and proof to the relevant sections. We obtain as a corollary:

Corollary 1.1. *If T is a Lozi map, a hyperbolic billiard map, or a Lorenz map then for μ a.e. x_0 the stochastic process defined by $X_n(x) = -\log(d(x_0, T^n x))$ satisfies a Type I extreme value law.*

We also obtain extreme value laws for the Lorenz equations. We again defer the formal description to Section 4.4 but as a corollary:

Corollary 1.2. *Let $T_t : \mathbb{R}^3 \to \mathbb{R}^3$ be the Lorenz flow. Then a Type I extreme value law holds for the process defined by $X_t(x) = -\log(d(x_0, T_t x))$ for μ a.e. x_0.*

We also obtain extreme value laws for other non-uniformly hyperbolic systems. In particular we consider the Hénon family of maps $T(x, y) = (1 - ax^2 + y, bx)$ for $a \simeq 2$ and $|b| \ll 1$. For a positive Lebesgue measure set of parameters $(a, b) \in \mathbb{R}^2$ the Hénon family of maps admit a Young tower with exponential tails (and hence have an SRB measure μ with exponential decay of correlations), see [23, 1]. In particular it is shown that there exists a subset $\Lambda \subset \mathbb{R}^2$ with a hyperbolic product structure and properties (P1)-(P4) of [23] hold. In contrast to Definition 1.1 Hénon maps do not admit invariant cone fields and do not have uniform derivative growth estimates. However we are able to show:

Theorem 1.2. *Let $T : M \to M$ be a Hénon diffeomorphism modeled by a Young tower with exponential decay of correlations. Then, for μ a.e. x_0, the stochastic process defined by $X_n(x) = -\log(d(x_0, T^n x))$ satisfies a Type I extreme value law.*

The proof of Theorems 1.1 and 1.2 will come in two parts. In the first part we verify condition $D_2(u_n)$, but keep the setting fairly general (we do not restrict to manifolds of dimension 2). In the second part we verify condition $D'(u_n)$. This uses ideas from [2] but the arguments need to be modified in the context of proving Theorem 1.1 due to the unbounded derivative, see Section 3. To check $D'(u_n)$ for Hénon maps we need to establish geometrical control of the unstable manifold (whose closure is the Hénon attractor). This will be discussed in Section 5. Furthermore a large deviations estimate is needed to establish control on the derivative growth along typical orbits.

1.2. **Relationship to return time statistics.** For a diffeomorphism $T : M \to M$ preserving a probability measure μ, we may consider hitting and return time statistics as follows. For a set $A \subset M$, let $R_A(x)$ denote the first time $j \geq 1$ such that $T^j(x) \in A$. Given a sequence of sets $\{U_n\}_{n \in \mathbb{N}}$, with $\mu(U_n) \to 0$ then we say that the system has hitting time statistics (HTS) $G(t)$ for $\{U_n\}$ if for all $t \geq 0$

$$\lim_{n \to \infty} \mu\left(R_{U_n} \geq \frac{t}{\mu(U_n)}\right) = G(t). \tag{1}$$

We say that the system has HTS $G(t)$ to balls at x_0 is for any sequence $\delta_n \subset \mathbb{R}^+$, with $\delta_n \to 0$ as $n \to \infty$ we have HTS $G(t)$ for $U_n = B_{\delta_n}(x_0)$.

Analogously we say that *return time statistics* (RTS) $G(t)$ holds for $\{U_n\}$ if we can replace the measure μ by the conditional measure μ_{U_n} in equation (1), where $\mu_A = \frac{\mu|A}{\mu(A)}$.

RTS to balls is also defined analogously to HTS to balls.

In [6] an equivalence between extreme value laws and hitting time statistics was obtained for a dynamical system (M, T, μ) admitting an absolutely continuous invariant probability measure μ. Hence from [6, Corollary 2] we have the following:

Corollary 1.3. *If T is a Lozi map, a hyperbolic billiard map, a Lorenz map or a Hénon map then for μ a.e. $x_0 \in M$ we have HTS to balls at x_0 with limit $G(t) = e^{-t}$.*

Remark 1.1. The techniques we use in this paper assume exponential decay of correlations and an invariant density in $L^{1+\sigma}$ for $\sigma > 0$. We would like to extend if possible these results to dynamical systems modeled by a general Young tower, with merely summable return time function.

2. Condition $D_2(u_n)$.

In this section we let M be a Riemannian manifold of dimension D, with Lebesgue measure m and let $T : M \to M$ be a diffeomorphism modeled by a Young Tower. As in Section 1.1 we assume that the invariant density h is in $L^{1+\sigma}(m)$ and there is a set Λ with a hyperbolic product structure. Define $\Delta_0 := \Lambda$. Let $\Lambda_{0,i}$ be a countable partition of Δ_0. Let $R : \Delta_0 \to \mathbb{N}$ be an $L^1(m)$ roof function with the property that

$$R|_{\Lambda_{0,i}} := R_i.$$

Define the Young Tower by

$$\Delta = \cup_{i,l \le R_i - 1} \{(x, l) : x \in \Lambda_{0,i}\}$$

and the tower map $F : \Delta \to \Delta$ by

$$F(x,l) = \begin{cases} (x, l+1) & \text{if } x \in \Lambda_{0,i}, l < R_i - 1 \\ (T^{R_i} x, 0) & \text{if } x \in \Lambda_{0,i}, l = R_i - 1 \end{cases}.$$

For convenience, we will refer to $\Delta_0 := \cup_i (\Lambda_{0,i}, 0)$ as the base of the tower Δ and denote $\Lambda_i := \Lambda_{0,i}$. We define $\Delta_l = \{(x, l) : l < R(x)\}$, the lth level of the tower. Define the map $f = F^R : \Delta_0 \to \Delta_0$. We may form a quotiented tower (see [23] for details) by introducing an equivalence relation for points on the same stable manifold. Much of the analysis for the statistical properties of the tower in [23] (but not in this paper) is performed on the quotiented tower, we will merely list the features that we will use.

There exists an invariant measure m_0 for $f : \Delta_0 \to \Delta_0$ which has absolutely continuous conditional measures on local unstable manifolds in Δ_0, with density bounded uniformly from above and below.

The tower structure allows us to construct an invariant measure ν for F on Δ by defining for a measurable set $B \subset \Lambda_l$, $\nu(B) = \frac{m_0(F^{-l}B)}{\int \Lambda_0 R dm_0}$ and extending the definition to disjoint unions of such sets in the obvious way. We define a projection $\pi : \Delta \to M$ by $\pi(x, l) = T^l(x)$. We note that $\pi \circ F = T \circ \pi$. The invariant measure μ, which is an SRB measure for $T : M \to M$, is given by $\mu = \pi_* \nu$. We have assumed (this is not a standard Tower assumption) that $h = \frac{\mathrm{d}\mu}{\mathrm{d}m} \in L^{1+\sigma}(m)$. $W^s_\eta(x)$ will denote the local stable manifold through x and $B_r(x)$ will denote the ball of radius r centered at the point x. We lift a function $\phi : M \to \mathbb{R}$ to Δ by defining, with abuse of notation, $\phi(x, l) = \phi(T^l x)$.

As a consequence of [23, (P2)], there exists an $\alpha \in (0, 1)$ and a $C > 0$ such that $d(T^n x, T^n y) \le C\alpha^n$ for all $y \in W^s_\eta(x)$. In particular, $|T^k W^s_\eta(x)| \le C\alpha^k$ where $|\dots|$

denotes, with abuse of notation, length of a curve in the Riemannian metric. If we define the set

$$B_{r,k}(x_0) = \left\{ x : T^k(W^s_\eta(x)) \cap \partial B_r(x_0) \neq \emptyset \right\}$$

we see that $T^k B_{r,n}(x_0)$ must lie completely within an annulus of width $2C\alpha^n$ of the boundary of $B_r(x_0)$. Hence $m(B_{r,n}(x_0)) \leq r^D C_1 \alpha^n$. It then follows from Hölder's inequality that

$$\mu(B_{r,k}(x_0)) \leq C_2 \|h\|_{1+\sigma} r^{\frac{D\sigma}{1+\sigma}} \alpha^{\frac{k\sigma}{1+\sigma}}$$

with the right hand side of the inequality independent of the reference point x_0. We will use this observation repeatedly in subsequent proofs.

Henceforth, we will fix a reference point x_0 in the support of μ and define a stochastic process X_n given by $X_n(x) = -\log d(T^n x, x_0)$. We are interested in the distribution of the maximum of X_n, denoted by

$$M_n = \max\{X_0, X_1, \ldots, X_n\}.$$

We will prove the condition $D_2(u_n)$ [4] for the sequence $u_n = v + \frac{1}{D}\log n$.

We define $\kappa(n)$ to be the rate of decay of correlations of Hölder functions with respect to the SRB measure μ on the manifold: so that

$$\left| \int_M \phi\psi \circ T^n d\mu - \int_M \phi d\mu \int_M \psi d\mu \right| \leq \kappa(n) \|\phi\|_{Lip} \|\psi\|_{Lip}$$

for all Lipschitz ϕ, $\psi : M \to \mathbb{R}$. In fact on a quotiented Young Tower (see Lemma 2.1) we may use the L^∞ norm of ψ in the estimate above if ψ is defined on the quotiented tower and in general a faster decay rate than $\kappa(n)$.

We now state the theorem proved in this section:

Theorem 2.1. *Suppose there exists an $\epsilon > \frac{D}{\sigma}$ such that $n^{\frac{1+D+\epsilon}{D}}\kappa(\sqrt{n}) \to 0$ as $n \to \infty$. Then the stochastic process $(X_n)_{n\in\mathbb{N}}$ satisfies the condition $D_2(u_n)$, namely, for any integers j, l and n,*

$$|\mu(\{X_0 > u_n\} \cap \{M_{j,l} \leq u_n\}) - \mu(\{X_0 > u_n\})\,\mu(\{M_{0,l} \leq u_n\})| \leq \gamma(n,j) \tag{2}$$

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

Once condition $D_2(u_n)$ is proved, from [2, 4] we obtain as a corollary:

Corollary 2.1. *Let $T : M \to M$ be a diffeomorphism modeled by a Young tower satisfying the hypotheses of Theorem 2.1, let $X_n = -\log d(x_0, T^n x)$ and $u_n = v + \frac{1}{D}\log n$. Then, if condition $D'(u_n)$ holds for the sequence X_n*

$$\lim_{n\to\infty} P(M_n \leq u_n) = e^{-h(x_0)e^{-Dv}}.$$

Proof of Corollary 2.1. The following facts are evident from [2]. We use the notation of [4]. Firstly, for any $l \in \mathbb{N}$ and $u \in \mathbb{R}$

$$\sum_{j=0}^{l-1} P(X_j > u) \geq P(M_l > u) \geq \sum_{j=0}^{l-1} P(X_j > u) - \sum_{j=0}^{l-1} \sum_{i\neq j, i=0}^{l-1} P(\{X_j > u\} \cap \{X_i > u\}).$$

Secondly, if $t, r, m, s \in \mathbb{N}$, then

$$0 \leq P(M_r \leq u) - P(M_{r+l} \leq u) \leq lP(X > u)$$

and

$$\left|P(M_{s+t+m} \leq u) - P(M_m \leq u) + \sum_{j=0}^{s-1} P(\{X > u\} \cap \{M_{s+t-j,m} \leq u\})\right|$$
$$\leq 2s \sum_{j=1}^{s-1} P(\{X > u\} \cap \{X_j > u\}) + tP(X > u).$$

It follows from a measure-theoretic manipulation (see [2]) that

$$|P(M_n \leq u_n) - (1 - pP(X > u_n))^q| \leq qtP(X > u_n) + q\Gamma_n$$

where p, q, t are suitable increasing, unbounded functions of n, $pq = n$, $t = o(n) < p$ and

$$\Gamma_n = p\gamma(n,t) + tP(X > u_n) + 2p \sum_{j=1}^{p-1} P(\{X > u_n\} \cap \{X_j > u_n\}).$$

It is clear from Lebesgue's differentiation theorem that

$$\lim_{n\to\infty} nP(X > u_n) = h(x_0)e^{-Dv}$$

$\mu - a.e. x_0$ so

$$\lim_{n\to\infty} P(M_n \leq u_n) = e^{-h(x_0)e^{-Dv}}$$

provided $q\Gamma_n \to 0$. This follows from $D_2(u_n)$ and $D'(u_n)$.

□

An immediate corollary to Corollary (2.1), using Lemma (1.1) is:

Corollary 2.2. *Let $T : M \to M$ be a diffeomorphism modeled by a Young tower satisfying the hypotheses of Theorem 2.1, let $X_n = -\log d(x_0, T^n x)$ and $u_n = v + \frac{1}{D}\log n$. Suppose $D'(u_n)$ holds for X_n. Then if $\beta > 0$,*

1. *A Type I law holds for $-\log d(T^j x, x_0)$ with $a_n = 1$ and $b_n = \frac{1}{D}\log n$, i.e.,*
$$\lim_{n\to\infty} \mu\left(\max_{0\leq j\leq n} -\log d(T^j x, x_0) \leq v + \frac{1}{D}\log n\right) = e^{-h(x_0)e^{-Dv}}$$
2. *A Type II law holds for $d(x_0, T^j x)^{-D\beta}$ with $a_n = n^{-\beta}$ and $b_n = 0$, i.e.,*
$$\lim_{n\to\infty} \mu\left(\max_{0\leq j\leq n} d(T^j x, x_0)^{-D\beta} \leq \frac{\eta}{n^{-\beta}}\right) = e^{-h(x_0)\eta^{-\frac{1}{\beta}}}$$
3. *A Type III law holds for $C - d(x_0, T^j x)^{D\beta}$ with $a_n = n^{\beta}$ and $b_n = C$, i.e.,*
$$\lim_{n\to\infty} \mu\left(\max_{0\leq j\leq n} C - d(x_0, T^j x)^{D\beta} \leq C + \frac{\eta}{n^{\beta}}\right) = e^{-h(x_0)(-\eta)^{\frac{1}{\beta}}}.$$

We start with a lemma that clarifies the role of the rate of decay of correlations $\kappa(n)$.

Lemma 2.1. *Suppose $\Phi : M \to \mathbb{R}$ is Hölder continuous and $\Psi_{a,b}$ is the indicator function $1_{\{X_a\leq u_n, X_{a+1}\leq u_n, \ldots, X_{a+b}\leq u_n\}}$.*

Then there exist constants $A_1(\Phi)$ and $A_2(\Phi)$, such that

$$\left|\int \Phi\Psi_{0,l}\circ T^j\, d\mu - \int \Phi\, d\mu \int \Psi_{0,l}\, d\mu\right| \leq A_1 \left(\frac{e^{-v}}{n^{1/D}}\right)^{\frac{D\sigma}{1+\sigma}} \alpha^{\frac{j\sigma}{2+2\sigma}} + A_2\kappa(j) \tag{3}$$

Proof. Define the function $\tilde{\Phi} : \Delta \to \mathbb{R}$ by $\tilde{\Phi}(x,l) = \Phi(T^l x)$ and the function $\hat{\Psi}_{[j/2],l}(x,r) = \Psi_{[j/2],l}(T^r x)$. For simplicity of notation we will write $[j/2]$ as $j/2$. We choose a reference local unstable manifold $W^u_\eta(\tilde{x})$ and using the hyperbolic product structure, for each point x on a local stable manifold (lsm) $W^s_\eta(x)$ of the base of the tower Δ_0, we choose a reference point $\hat{x} \in W^u_\eta(\tilde{x})$ so that $x \in W^s_\eta(\hat{x})$.

We define the function $\tilde{\Psi}_{j/2,l}(x,r) := \hat{\Psi}_{j/2,l}(\hat{x},r)$. We note that $\tilde{\Psi}_{j/2,l}$ is constant along stable manifolds and the set of points where $\tilde{\Psi}_{j/2,l} \neq \hat{\Psi}_{j/2,l}$ along stable manifolds is by definition the set for which there exists an x and y belonging to the same lsm with $T^{j/2+i}(x) \in \{X_{j/2} \leq u_n, \ldots, X_{j/2+l} \leq u_n\}$ and $T^{j/2+i}(y) \notin \{X_{j/2} \leq u_n, \ldots, X_{j/2+l} \leq u_n\}$ which is completely contained inside $\cup_{k=j/2}^{j/2+l} B_{\frac{e^{-v}}{n^{1/D}}\alpha^k}$ which, in turn, is contained inside the annulus of width $2C\alpha^{j/2}$ around the boundary. Hence $\nu\{\hat{\Psi}_{j/2,l} \neq \tilde{\Psi}_{j/2,l}\} \leq C_3 \left(\frac{e^{-v}}{n^{1/D}}\right)^{\frac{D\sigma}{1+\sigma}} \alpha^{\frac{j\sigma}{2+2\sigma}}$.

By the decay of correlations as proved in [2, 23] (recall $\tilde{\Psi}$ is defined on the Tower quotiented in the stable direction), we have

$$\left|\int \tilde{\Phi}\tilde{\Psi}_{j/2,l} \circ F^{j/2} \mathrm{d}\nu - \int \tilde{\Phi} \mathrm{d}\nu \int \tilde{\Psi}_{j/2,l} \mathrm{d}\nu\right| \leq C_4 \|\Phi\|_{\mathrm{Lip}} \|\Psi\|_\infty \kappa(j)$$

where $\kappa(j) \to 0$ as $j \to \infty$ exponentially or polynomially fast depending on the tails of the Young tower being exponential or polynomial, hence

$$\begin{aligned}
&\left|\int \Phi\Psi_{j/2,l} \circ T^{j/2} \mathrm{d}\mu - \int \Phi \mathrm{d}\mu \int \Psi_{j/2,l} \mathrm{d}\mu\right| \\
&= \left|\int \tilde{\Phi}\hat{\Psi}_{j/2,l} \circ F^{j/2} \mathrm{d}\nu - \int \tilde{\Phi} \mathrm{d}\nu \int \hat{\Psi}_{j/2,l} \mathrm{d}\nu\right| \\
&\leq \left|\int \tilde{\Phi}(\hat{\Psi}_{j/2,l} - \tilde{\Psi}_{j/2,l}) \circ F^{j/2} \mathrm{d}\nu\right| + C_4 \|\Phi\|_{\mathrm{Lip}} \kappa(j) \\
&\quad + \left|\int \tilde{\Phi} \mathrm{d}\nu \int (\tilde{\Psi}_{j/2,l} - \hat{\Psi}_{j/2,l}) \mathrm{d}\nu\right| \\
&\leq C_3 (\sup|\Phi|) \left(\frac{e^{-v}}{n^{1/D}}\right)^{\frac{D\sigma}{1+\sigma}} \alpha^{\frac{j\sigma}{2+2\sigma}} + C_4 \|\Phi\|_{\mathrm{Lip}} \kappa(j).
\end{aligned}$$

The proof is complete if we notice that $\int \Psi_{0,l} \mathrm{d}\mu = \int \Psi_{j/2,l} \mathrm{d}\mu$ by the μ invariance of T and that $\Psi_{j/2,l} \circ T^{j/2} = \Psi_{j,l} = \Psi_{0,l} \circ T^j$.

□

Remark 2.1. The constants A_1 and A_2 are $\mathcal{O}(\|\Phi\|_{\mathrm{Lip}})$. We will require, in the proof of the condition $D_2(u_n)$ that the rate of decay of correlations $\kappa(n)$ be sufficiently fast.

Proof of Theorem 2.1. We approximate the indicator function $1_{\{X_0 \geq u_n\}}$ by a Lipschitz continuous function Φ as follows. The set $\{X_0 \geq u_n\}$ corresponds to a ball of radius $\frac{e^{-v}}{n^{1/D}}$ about x_0. We define Φ to be 1 on a ball of radius $\frac{e^{-v}}{n^{1/D}} - \left(\frac{e^{-v}}{n^{1/D}}\right)^{1+\epsilon}$ about x_0 and decaying to 0 at a linear rate so that on the boundary it takes the value 0. The Lipschitz norm of Φ is $\frac{n^{\frac{1+\epsilon}{D}}}{e^{-v(1+\epsilon)}}$. Therefore,

$$\begin{aligned}&\left|\int 1_{\{X_0\geq u_n\}}\Psi_{j/2,l}\circ T^{j/2}\mathrm{d}\mu-\mu(X_0\geq u_n)\int\Psi_{j/2,l}\mathrm{d}\mu\right|\\ \leq\ &\left|\int(1_{\{X_0\geq u_n\}}-\Phi)\Psi_{j/2,l}\mathrm{d}\mu\right|+C_3\left(\frac{e^{-v}}{n^{1/D}}\right)^{\frac{D\sigma}{1+\sigma}}\alpha^{\frac{j\sigma}{2+2\sigma}}\\ &+C_4\|\Phi\|_{\mathrm{Lip}}\kappa(j)+\left|\int(1_{\{X_0\geq u_n\}}-\Phi)\mathrm{d}\mu\int\Psi_{j/2,l}\mathrm{d}\mu\right|.\end{aligned}$$

However, the set where $1_{\{X_0\geq u_n\}}$ differs from Φ is an annulus of inner radius $\frac{e^{-v}}{n^{1/D}}-\left(\frac{e^{-v}}{n^{1/D}}\right)^{1+\epsilon}$ and outer radius $\frac{e^{-v}}{n^{1/D}}$, so

$$\left|\int(1_{\{X_0\geq u_n\}}-\Phi)\mathrm{d}m\right|\leq C_5\left(\frac{e^{-v}}{n^{1/D}}\right)^{D+\epsilon}$$

(this follows from the fact that for $|r|<1, r^D-(r-r^{1+\epsilon})^D\leq D^2((D-1)!)r^{D+\epsilon}$). Therefore,

$$\left|\int(1_{\{X_0\geq u_n\}}-\Phi)\mathrm{d}\mu\right|\ \leq\ C_5'\left(\frac{e^{-v}}{n^{1/D}}\right)^{\frac{\sigma(D+\epsilon)}{1+\sigma}}$$

Combined with the observation that $\Psi_{j/2,l}\circ T^{j/2}=\Psi_{j,l}$ and $\int\Psi_{j/2,l}\mathrm{d}\mu=\int\Psi_{0,l}\circ T^{j/2}\mathrm{d}\mu=\int\Psi_{0,l}\mathrm{d}\mu$ we have

$$|P(\{X_0\geq u_n\}\cap\{M_{j,l}\leq u_n\})-P(\{X_0\geq u_n\})P(\{M_{0,l}\}\leq u_n)|\leq\gamma(n,j)\tag{4}$$

where

$$\gamma(n,j)=2C_5'\left(\frac{e^{-v}}{n^{1/D}}\right)^{\frac{\sigma(D+\epsilon)}{1+\sigma}}+C_3\left(\frac{e^{-v}}{n^{1/D}}\right)^{\frac{D\sigma}{1+\sigma}}\alpha^{\frac{j\sigma}{2+2\sigma}}+C_6n^{\frac{1+\epsilon}{D}}\kappa(j).$$

If we choose $j_n=\sqrt{n}$, then $n\gamma(n,j_n)\to 0$ if and only if $n^{\frac{1+D+\epsilon}{D}}\kappa(\sqrt{n})\to 0$.

□

3. Condition $D'(u_n)$ for hyperbolic systems with singularities

In this section we establish condition $D'(u_n)$ for a class of two-dimensional uniformly hyperbolic maps with stable foliation modeled by a Young Tower with exponential tails. However we allow for the derivative map of T to have discontinuities or singularities. We let S denote the singular set for T and $\mathcal{S}=\bigcup_k T^{-k}S$.

Theorem 3.1. *Let $T:M\to M$ be a two-dimensional (local) hyperbolic diffeomorphism modeled by a Young tower with exponential decay of correlations. Furthermore suppose that T is hyperbolic with singularities in the sense of Definition 1.1. Then for μ a.e. x_0 the stochastic process defined by $X_n(x)=-\log(d(x_0,T^nx))$ satisfies condition $D'(u_n)$.*

The strategy of proof is the following: We first obtain a bound on the conditional measure μ_C on an unstable manifold for the set of points that return to a ϵ neighborhood of themselves in k steps. We then extend this bound to the set of those points that return to a $1/k$ neighborhood of themselves before $(\log k)^5$ steps. A Borel-Cantelli argument together with the Maximal Function technique used by Collet is then used to complete the proof of condition $D'(u_n)$. It is sufficient to estimate the conditional measure on unstable manifolds, as μ is defined via the pushforward of the conditional measure on a reference local unstable manifold. It helps to read the proof below as if Λ_i are intervals, though in certain applications the partition sets will be Cantor subsets of local unstable manifolds.

Define $\mathcal{E}_i(\epsilon)$ by

$$\mathcal{E}_i(\epsilon) := \left\{x : d(T^i x, x) \leq \epsilon\right\}.$$

Proposition 3.1. *Under the hypothesis of Theorem 3.1*

$$\mu(\mathcal{E}_k(\epsilon)) \leq O(1)\left(\sqrt{\epsilon} + k\eta^{\sqrt{-\log \epsilon}}\right)$$

for some $0 < \eta < 1$.

Proof. We fix a reference local unstable manifold (lum) $W^u_\eta(\tilde{x}) \subset \Delta_0$. The conditional measures on lum $W^u_\eta(x) \subset \Delta_0$ are related by holonomy along local stable manifolds (lsm) and the holonomy preserves conditional measure [23, Section 3, Lemma 1].

The partition of Δ_0 into sets $\{\Lambda_i\}$ induces a partition of $W^u_\eta(\tilde{x})$, and we will with abuse of notation denote the elements of this partition also by $\{\Lambda_i\}$. If we apply $T^{R_l - j}$ to $T^j(\Lambda_l)$ we land in $\Delta_0 \cap W^u_\eta$ for some lum W^u_η . Because holonomy preserves conditional measure we will slide along lsm to identify the image with the original lum $W^u_\eta(\tilde{x})$. Equivalently we could work on a quotiented tower but this would introduce more cumbersome notation.

As in Collet, the idea of proof is to push forward Λ_l (i.e $\Lambda_l \cap W^u_\eta(\tilde{x})$) to $T^j\Lambda_l$ for some $j < R_l$, considering the intersection of $\mathcal{E}_k(\epsilon)$ with intervals of monotonicity of T^k on a finer partition induced on $T^j\Lambda_l$: we partition into points which have not been separated by the Tower partition under k forward iterates. We then sum over the pullback to Λ_l, sum over $j < R_l$ and finally sum over l. This gives us $\mu(\mathcal{E}_k(\epsilon))$ by definition of μ.

For a sequence of integers (s_i) let $\Lambda_{s_1,\ldots,s_r}$ denote a cylinder set of the partition of $W^u_\eta(\tilde{x}) \cap \Delta_0$ induced by $\bigvee_{j=0}^{r-1}(T^R)^{-j}P_0$ such that if $x \in \Lambda_{s_1,\ldots,s_r}$ then $R(x) = R_{s_1}$, $R(fx) = R_{s_2}$ and $R(f^i x) = R_{s_i}$ for $i \leq r$. Note that $T^{R_{s_1}+\ldots+R_{s_{r-1}}}\Lambda_{s_1,\ldots,s_r} = \Lambda_{s_r}$ and $T^{R_{s_1}+\ldots+R_{s_p}}\Lambda_{s_1,\ldots,s_r} = \Lambda_{s_{p+1}}$ for $p = 1,\ldots,r-1$ (under our holonomy identification). For intervals I of $T^j(\Lambda_l)$ not separated by the tower under k forward iterates we will write $k = R_l - j + R_{s_1} + \cdots + R_{s_r} - p$, where the expression $R_{s_1} + \cdots + R_{s_r} - p$ is a random variable on $T^j(\Lambda_l)$ but is constant on such an interval I. The idea is that under T^k a point in such an interval $I \subset T^j(\Lambda_l)$ lies in $T^p\Lambda_{s_r}$ with $p < R_{s_r}$ i.e. lies in the column above Λ_{s_r} but has not been broken by intervening trips up the tower.

Recall that forward images of $\Delta_0 \cap W^u_\eta$ have tangent vectors lying in an unstable cone. We will consider the following cases:

(1) $|T^k(I)| > \delta$,
(2) $|T^k(I)| \leq \delta$,
  (2.a) $p > \sqrt{-\log \delta}$,
    (2.a.i) $R_l < \sqrt{-\log \delta}$,
    (2.a.ii) $R_l \geq \sqrt{-\log \delta}$,
  (2.b) $p \leq \sqrt{-\log \delta}$.

We will consider first of all the collection of all such sets I with $|T^k(I)| \geq \delta$. (We recall that $|\cdot|$ denotes arc length measure here). Due to assumption 1 in definition 1.1, we obtain for some $k \geq k_0$, $|D(T^k)_u| > m$. Thus if I is an interval and $J = \mathcal{E}_k(\epsilon) \cap I$, then $|T^k(J)| < C(m)\epsilon$.

This uses the observation that at the linear level if F is a 1-d smooth map and $|F'(x)| > m$, then $|\{x : |F(x) - x| < \epsilon\}| \leq 1/(m-1)\epsilon$ for small ϵ which is the essential idea used by

Collet. In this setting tangent vectors to $T^k(I)$ lie in a cone so $|T^k(J)| < C(m)\epsilon$ where $C(m)$ depends on the angle of the cone and decreases in m.

If $|T^k(I)| \geq \delta$ then by bounded distortion $\frac{|J|}{|I|} \leq C\frac{\epsilon}{\delta}$. The rest of this argument needs no modification from [2] in the case $|T^k(I)| > \delta$. We have, by bounded distortion, $\mu_C(\Lambda_l \cap T^{-j}J) \leq C\frac{\epsilon}{\delta}\mu_C(\Lambda_l \cap T^{-j}I)$. If we sum over all such disjoint intervals I contained in $T^j(\Lambda_l)$ with large image i.e. $|T^k(I)| > \delta$ then we get a bound of $C\frac{\epsilon}{\delta}\mu_C(\Lambda_l)$. Summing over all $j < R_l$ and over the index l we obtain a bound of $C\frac{\epsilon}{\delta}$.

Next we must consider monotonicity intervals I with $|T^k(I)| \leq \delta$. Such intervals arise, for example, out of frequent chopping with S which in turn prevents them from growing. We show that this set is small by exponential decay of correlations. Due to the unbounded derivative we cannot apply [2] directly.

Instead we balance the sums according as to whether the return time R is large (or small) relative to $c(\delta) := \sqrt{-\log \delta}$. Recall $p = R_l - j + R_{s_1} + \cdots + R_{s_r} - k$. Note that if I has not returned to the base then $p = R_l - j - k$. We consider two cases: $p > c(\delta)$ and $p \leq c(\delta)$.

We note that $T^k(I) = T^{\hat{k}}\Lambda_{s_r}$ (where $\hat{k} = R_{s_r} - p$) or in the case of no returns $T^k(I) = T^{j+k}(\Lambda_l)$. Denote by $I_{l,j,s_1,\dots,s_r} \subset T^j\Lambda_l$ an interval such that points have the itinerary $\Lambda_{s_1}, \dots, \Lambda_{s_r}$ under successive returns to the base. We estimate if $p > c(\delta)$:

$$\mathcal{E}_k(\epsilon) \cap T^j(\Lambda_l) = \bigcup_{R_{s_1},\dots,R_{s_r}} (I_{l,j,s_1,\dots,s_r} \cap \mathcal{E}_k(\epsilon)) \subset T^j(\Lambda_l) \cap \left(\bigcup_{m=0}^{k} T^{-m}(\{R > c(\delta)\}) \right), \tag{5}$$

where in the first union we consider all relevant sequences $R_{s_1}, \dots, R_{s_r}$ with $R_l + \sum_{i=1}^{r} R_{s_i} > j + k$ and $R_l + \sum_{i=1}^{r-1} R_{s_i} < j + k$. In the union, we also account for the case $R_l > j + k$. We now apply T^{-j} and intersect with Λ_l to estimate

$$\sum_{j=0}^{R_l} \sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(\{R > c(\delta)\})). \tag{6}$$

Finally we sum over Λ_l

$$\sum_{R_l} \sum_{j=0}^{R_l} \sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(\{R > c(\delta)\})). \tag{7}$$

We consider two cases $R_l > c(\delta)$ and $R_l \leq c(\delta)$. In the following θ will denote the exponential rate of decay of tails, $k(n) = \theta^n$, while $\theta_i < 1$ and $C_i > 0$, $i = 1, 2$, will be generic constants that do not depend on δ or k.

Consider first the restriction of (7) to $R_l < c(\delta)$.

$$\begin{aligned}\sum_{R_l<c(\delta)} \sum_{j=0}^{R_l-1} \sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(\{R > c(\delta)\})) \\ \leq \sum_{R_l<c(\delta)} \sum_{j=0}^{R_l-1} \sum_{m=0}^{k+c(\delta)} \mu_C(\Lambda_l \cap T^{-m}(\{R > c(\delta)\})) \\ \leq c(\delta) \sum_{R_l<c(\delta)} \sum_{m=0}^{k+c(\delta)} \mu_C(\Lambda_l \cap T^{-m}(\{R > c(\delta)\})) \\ \leq c(\delta) \sum_{m=0}^{k+c(\delta)} \mu_C(T^{-m}(\{R > c(\delta)\})) \\ \leq O(1)c(\delta)(k + c(\delta))\mu_C(\{R > c(\delta)\}).\end{aligned}$$

For $R_l > c(\delta)$ we bound the expression in (7) by

$$\begin{aligned}\sum_{R_l>c(\delta)} \sum_{j=0}^{R_l-1} \sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(\{R > c(\delta)\})) \\ \leq \sum_{R_l>c(\delta)} kR_l\mu_C(\Lambda_l) \leq O(1)k \sum_{m>c(\delta)} m\theta^m \\ \leq C_1 k\theta_1^{c(\delta)}\end{aligned}$$

We obtain the bound:

$$\begin{aligned}\sum_{R_l} \sum_{j=0}^{R_l} \sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(\{R > p\})) \\ \leq O(1)c(\delta)(k + c(\delta))\theta\sqrt{-\log\delta} + k\tilde{\theta}^{\sqrt{-\log\delta}} \leq C_2\theta_2^{\sqrt{-\log\delta}}.\end{aligned} \tag{8}$$

For the case $p < c(\delta)$ recall that

$$|\Lambda| = T^p(T^k I)$$

as $p + k = R_l - j + R_{s_1} + \ldots + R_{s_r}$ and I makes a full crossing under T^{p+k}.
Therefore, by bounded distortion,

$$T^k(I) \subset \{x \in \Lambda : |DT(T^i x)| > C\delta^{\frac{-1}{\sqrt{-\log\delta}}} \text{ for some } i < \sqrt{-\log\delta}\}.$$

Thus

$$T^k(I) \subset A_p := \cup_{i=1}^{p} T^{-i}\{|DT(x)| > C\delta^{\frac{-1}{\sqrt{-\log\delta}}}\}. \tag{9}$$

The measure of the union of such intervals I is bounded by

$$\sum_{R_l} \sum_{j=0}^{R_l} \sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(A_p)). \tag{10}$$

Recall $c(\delta) = \sqrt{-\log\delta}$, We balance the two cases $R_l < c(\delta)$ and $R \geq c(\delta)$ as we did for equation (7). The quantity (10) is bounded by

$$O(1)c(\delta)kA_p) + C_1k\theta_1^{c(\delta)} \leq O(1)c(\delta)k\sqrt{-\log\delta}\,\delta^{\frac{1}{\sqrt{-\log\delta}}} + C_1k\theta_1^{c(\delta)}. \tag{11}$$

Note that we have estimated the measure of A_p by using Markov's inequality. The term in (11) is then bounded by

$$O(1)\left(\sqrt{-\log\delta}\,k\left(\sqrt{-\log\delta}\,\delta^{\frac{1}{\sqrt{-\log\delta}}}\right) + k\tilde{\theta}^{\sqrt{-\log\delta}}\right) = C_3k\theta_3^{c(\delta)} \tag{12}$$

Collecting all estimates we get from equations (8) and (12):

$$\mu_C(\mathcal{E}_k(\epsilon)) \leq C_4\frac{\epsilon}{\delta} + C_2k\theta_2^{c(\delta)} + C_3k\theta_3^{c(\delta)} \tag{13}$$

By choosing $\delta = \sqrt{\epsilon}$ we obtain the following bound for some $\eta < 1$:

$$\mu_C(\mathcal{E}_k(\epsilon)) \leq O(1)\left(\sqrt{\epsilon} + k\eta^{\sqrt{-\log\epsilon}}\right).$$

□

We define the set E_k:

$$E_k = \{d(T^jx, x) \leq 1/k \text{ for some } 1 \leq j \leq (\log k)^5\}.$$

We will use our estimate of $\mu(\mathcal{E}_k(\epsilon))$ to estimate $\mu(E_k)$.

Proposition 3.2. *Under the hypothesis of Theorem 3.1, and for k large enough, there exists $\tilde{\eta} < 1$ such that $\mu(E_k) \leq \tilde{\eta}^{\sqrt{\log k}}$*

Proof. Under the hypothesis of Theorem 3.1 we obtain $\mu(E_k) \leq \sum_{j=1}^{(\log k)^5} \mu(\mathcal{E}_j(1/k))$. Using Proposition 3.1 we compute this as:

$$\mu(E_k) \leq O(1)\sum_{j=1}^{(\log k)^5}\left(\sqrt{\frac{1}{k}} + j\eta^{\sqrt{\log k}}\right) \leq \tilde{\eta}^{\sqrt{\log k}}$$

for some $0 < \tilde{\eta} < 1$. □

Collecting these results we are now ready to prove Theorem 3.1. The reader will observe that some estimates we have used in the proof of Theorem 3.1 are worse than the estimates we obtain in Proposition 3.1 and 3.2. This is so because Theorem 5.1, the non-uniformly hyperbolic analogue of Theorem 3.1, requires the estimation of these same inequalities, and the estimates we have used hold for the non-uniformly hyperbolic case also.

Proof of Theorem 3.1. Choose, arbitrarily, a number $\beta \in (0,1)$ and $\rho \in (0, \beta/3)$ and define the sets

$$F_n := \left\{x : \mu(B_{e^{-n^\beta}}(x) \cap E_{e^{n^\beta}}) \geq e^{-2n^\beta}e^{-n^{\beta\rho}}\right\}.$$

Define the Hardy-Littlewood maximal function M_n as

$$M_n(x) = \sup_{a>0}\frac{1}{B_a(x)}\int \chi_{E_n}(y)d\mu(y).$$

A theorem of Hardy and Littlewood [16, Theorem 2.19] implies that

$$\mu(|M_n| > c) \leq \frac{\|\chi_{E_n}\|_1}{c}$$

for all $c > 0$.

Since

$$F_n \subset \left\{ M_{e^{n^\beta}} \geq e^{-n^{\beta\rho}} \right\}$$

we have

$$\mu(F_n) \leq \|\chi_{E_{e^{n^\beta}}}\| e^{n^{\beta\rho}} \leq \mathcal{O}(1)(n^{20\beta} e^{n^{\beta\rho}} e^{-\tilde{\alpha} n^{\beta/2}}) \leq e^{-n^\gamma}$$

for some $\gamma > 0$. By the Borel Cantelli Lemma, $\mu(\overline{\lim} F_n) = 0$. Hence for a.e. x there exists an N_x such that $x \notin F_n$ for all $n \geq N_x$. For any $k = \left(\frac{1}{2}\log n\right)^{\frac{1}{\beta}} \geq N_x, x \notin F_k$ and hence we see that since for $j \leq (\log n)^2$

$$\{X \geq u_n\} \cap \{X \geq u_n\} \circ T^j \subset B_{\frac{1}{\sqrt{n}}}(x) \cap E_{\sqrt{n}}$$

we obtain

$$\mu(\{X \geq u_n\} \cap \{X \geq u_n\} \circ T^j) \leq \frac{e^{-(\log n)^\rho / 2^\rho}}{n}.$$

On summing from $j = 1$ to $(\log n)^2$ and taking limits, we obtain the desired result. □

4. Applications of Theorem 3.1

In this section we consider a range of applications of Theorem 3.1. We first consider Lozi maps and hyperbolic billiards. These are hyperbolic systems that admit invariant cone fields, but the derivative map DT is discontinuous or singular. In some of these examples the corresponding Young Towers are built over partition sets Λ_i which are Cantor sets. Hence we must carefully explain how Theorem 3.1 applies as it is written in the interest of clarity for $\{\Lambda_i\}$ a collection of intervals. We also consider extremal properties of Lorenz attractors by considering the Lorenz flow as a suspension over a hyperbolic map with a Young Tower. In each of the following case studies we review the model and then verify the appropriate conditions of Theorem 3.1.

4.1. **Hyperbolic product structures.** We recall the definition of a hyperbolic product structure for a subset $\Lambda \subset M$, see [23].

Definition 4.1. Λ has a hyperbolic product structure if there are families $\Gamma^u = \bigcup W^u_\eta$ and $\Gamma^s = \bigcup W^s_\eta$ of local unstable (reps. stable) manifolds with $\Lambda = (\cup W^u_\eta) \cap (\cup W^s_\eta)$. Each $W^u_\eta \in \Gamma^u$ crosses transversally each $W^s_\eta \in \Gamma^s$ (and vice versa) and $\dim(W^u_\eta) + \dim(W^s_\eta) = \dim(M)$.

Given Λ, a s-sublattice $\Lambda' \subset \Lambda$ is a subset of the form $\Gamma^u \cap \tilde{\Gamma}^s$ with $\tilde{\Gamma}^s \subset \Gamma^s$. Similarly a u-sublattice of Λ is a subset of the form $\tilde{\Gamma}^u \cap \Gamma^s$ with $\tilde{\Gamma}^u \subset \Gamma^u$.

We assume properties (P1)-(P4) of [23] hold. Briefly, these properties imply the existence of a countable Markov partition $\mathcal{P}_\Lambda$ of Λ into s-sublattices $\{\Lambda_i\}$ with return time function $R : \mathcal{P}_\Lambda \to \mathbb{N}$ such that $T^{R_i}\Lambda_i \subset \Lambda$ (with bounded distortion) and $T^{R_i}\Lambda_i$ is a u-sublattice.

In the setting of Lozi maps and hyperbolic billiards the partition sets Λ_i are Cantor sets. The base Δ_0 has a $f := F^R$ invariant measure m_0. The invariant measure m_0 on Δ_0 induces conditional measure μ_C on $W^u_\eta \in \Gamma^u$ which is the restriction of an absolutely continuous measure with density bounded above and below. The holonomy along stable manifolds in Γ^s preserves the conditional measures. The measure ν is defined on a set $A \subset \Delta_l$ by $\nu(A) = m_0(F^{-l}A)/\int_{\Delta_0} R dm_0$ which defines an F invariant measure on Δ.

For our examples $m_0(R > n) \leq \theta^n$ for some $\theta < 1$. There is a projection $\pi : \Delta \to M$ given by $\pi(x, l) = T^l x$. The T invariant SRB measure μ is given by $\mu(B) = \nu(\pi^{-1}B)$ for measurable sets $B \in M$.

4.2. **Lozi-Like Maps.** Let $R = [0, 1] \times [0, 1]$ and $T : R \to R$ be a continuous injective map. Assume that some iterate of T maps R into its interior and that there is a finite set S such that T is a C^2 diffeomorphism on $\mathcal{R}$ where $\mathcal{R} = R \setminus (S \times [0, 1])$. We further assume that T expands horizontally more than it folds and that horizontal expansion dominates the action of DT on vertical vectors.

Specifically, we assume that if $T = (T_1, T_2)$ then

$$\inf \left\{ \left(\left| \frac{\partial T_1}{\partial x} \right| - \left| \frac{\partial T_1}{\partial y} \right| \right) - \left(\left| \frac{\partial T_2}{\partial x} \right| + \left| \frac{\partial T_2}{\partial y} \right| \right) \right\} \geq 0 \tag{14}$$

$$\inf \left\{ \left| \frac{\partial T_1}{\partial x} \right| - \left| \frac{\partial T_1}{\partial y} \right| \right\} = u > 1 \tag{15}$$

$$\sup \left\{ \frac{|\partial T_1/\partial y| + |\partial T_2/\partial y|}{(|\partial T_1/\partial x| - |\partial T_1/\partial y|)^2} \right\} < 1 \tag{16}$$

and

$$\exists N \in \mathbb{N} : u^N > 2 \text{ and } T^k S \cap S = \emptyset \forall 1 \leq k \leq N. \tag{17}$$

Lozi maps are given by the map $T(x, y) = (1 + by - a|x|, x)$. It is first shown in [22] that Lozi-like maps have an invariant SRB measure. The tower that Young constructs for the Lozi map [23] has exponential tails for the return time function R.

Verification of Hypotheses of Theorem 3.1. We will show that Lozi-like maps are hyperbolic with singularities in the sense of Definition 1.1. For these maps there is a strict cone, hence Condition (3) is satisfied, Condition (1) follows from uniform hyperbolicity and as DT_u is integrable Condition (2) holds.

4.3. **Hyperbolic Billiards.** Our treatment and notation follows Young [23]. Let $\Gamma = \{\Gamma_i, i = 1, \ldots, k\}$ be a family of pairwise disjoint, simply connected C^3 curves with strictly positive curvature. We consider billiards on the domain $\mathbb{T}^2 \setminus \Gamma$. We assume the finite horizon condition, namely, that time between successive collisions is bounded. This is equivalent to the existence of an upper bound on the number of successive tangential collisions of any billiards trajectory with the family of scatterers Γ. We take $M := \Gamma \times \left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$ and let $T : M \to M$ be the Poincaré map.

We note that the authors in [23] restrict their verification of assumptions (P1)-(P5) in a metric called the p metric. The proof of Theorem 3.1 is written with the Euclidean metric in mind, but it adapts to the p metric as we show in this subsection.

Verification of Hypotheses of Theorem 3.1. We illustrate that the hyperbolic billiards map is hyperbolic with singularities (Definition 1.1). Hypothesis 1 is immediate in the p-metric due to the uniformly hyperbolic nature of the billiards map. The billiard maps are known to admit an invariant cone field, see for example [23, 11, 13]. Further, $|DT_u|_p$ is integrable with respect to the invariant measure μ because $|DT_u(x)|_p \approx \frac{1}{d(x,S)}$ where S is the singularity set for T, $S = \cup_i \Gamma_i \times \{\frac{\pm\pi}{2}\}$. We let $|\cdot|$ denotes the Euclidean distance in the (r,θ) co-ordinates. Since the density of the invariant measure is $d\mu = c\cos(\theta)drd\theta$, we note that

$$\|DT\|_1^\mu \approx \int_r \int_{\theta\in[-\pi/2,\pi/2]} \left|\frac{\cos(\theta)}{\theta - \frac{\pi}{2}}\right| d\theta dr < \infty$$

and hence $|DT_u|_p$ is integrable. Hyperbolic billiards therefore satisfy all the hypotheses of our theorem. It remains to show that the proof does not change by working with the p-metric.

First note that in [23] $|v|_p \le |v|$ (the inequality is given as $|v|_p \le |v|$ but C may be taken to be 1). In Proposition 3.1 we note that uniform hyperbolicity gives us $|DT_u^k|_p > m$ for any m for large enough k independent of x. This gives us $|DT_u^k| > m$ for the same values of k for any $x \in I$ where I is an interval of monotonicity obtained by the partitioning of $T^j(\Lambda_l)$ by iterating up till time $R_l - j + R_{s_1} + \cdots + R_{s_r}$. Using the Collet 1-d argument, we have that $|\mathcal{E}_k(\epsilon)\cap I| < \tilde{C}(m)\epsilon$ and so $|\mathcal{E}_k(\epsilon)\cap I|_p < \tilde{C}(m)\epsilon$. Since [23] verifies bounded distortion for the p metric, the argument involving partition intervals with the property $|T^k(I)|_p > \delta$ goes through unchanged in terms of the p-metric.

Next note that the bounded distortion estimates are with regard to the p metric. The conditional measure μ_C in equation (7) is in terms of the p metric. However, our arguments remain unchanged up to equation (8). If I is such that $|T^k(I)|_p < \delta$, then $|T^k(I)| \le \sqrt{\delta}$. Since $|\Lambda| = |DT^P(DT^k(I))|$ for $p = R_l - j + R_{s_1} + \cdots + R_{s_r} - k$, we can continue to argue as above with the estimates in equation (9) being replaced by

$$T^I \subset A_p := \cup_{i=1}^P T^{-i}\left\{|DP| > \delta^{\frac{-1}{2\sqrt{-\log\delta}}}\right\}. \tag{18}$$

Since $A_p \subset \left\{|DT|_p > \delta^{\frac{-1}{2\sqrt{-\log\delta}}}/C\right\}$, the estimates in equation (11) and equation (12) go through with the obvious changes.

4.4. **Lorenz attractors.** In this section we consider Lorenz maps and geometric Lorenz flows. The Lorenz equations

$$\dot{x} = 10(y-x), \quad \dot{y} = 28x - y - xz, \quad \dot{z} = xy - \frac{8}{3}z, \tag{19}$$

were introduced in 1963 by Lorenz [14], as a simplified nonlinear model for the weather. The mathematical study of these equations began with the geometric Lorenz flows, see [8, 21] which were shown to possess a strange attractor with sensitive dependence on initial conditions. Statistical properties of the actual Lorenz equations (19) were established in [19, 20]. It was proved that the attractor supports an SRB measure with positive Lyapunov exponents.

To establish extreme limit laws for the Lorenz equations we first consider the geometrical model. Let 0 be an equilibrium for a smooth (at least $C^{1+\epsilon}$) flow T_t on $\mathbb{R}^3$. For the corresponding vector field $Z : \mathbb{R}^3 \to \mathbb{R}^3$ suppose that the eigenvalues of $(DZ)_0$ are real and satisfy

$$\lambda_{ss} < \lambda_s < 0 < \lambda_u \quad \text{and} \quad \lambda_u > |\lambda_s|. \tag{20}$$

We choose coordinates (x_1, x_2, x_3) so that $(DZ)_0 = \text{diag}\{\lambda_u, \lambda_{ss}, \lambda_s\}$ and suppose that the flow T_t is $C^{1+\epsilon}$-linearizable in a neighborhood of 0. After rescaling, we may suppose that the flow is linearized in a neighborhood of the unit cube. Define the cross-sections $M = \{(x_1, x_2, 1) : |x_1|, |x_2| \le 1\}$, $M' = \{(1, x_2, x_3) : |x_2|, |x_3| \le 1\}$. The Poincaré map $T : M \to M$ (where defined) decomposes into $T = T_2 \circ T_1$ where $T_1 : M \to M'$ and

$T_2 : M' \to M$. Write $T(x) = T_{h(x)}(x)$ where $h : M \to \mathbb{R}^+$ is the first return time to M. It can then be shown that (see [15]) for $\beta = |\lambda_s|/\lambda_u \in (0,1)$ and $\beta' = |\lambda_{ss}|/\lambda_u > \beta$ we have $T_1(x_1, x_2, 1) = (1, x_1^{\beta'} x_2, x_1^{\beta})$, and $h(x) = -\lambda_u^{-1} \log |x_1| + h_0(x)$ where $h_0 \in C^{\beta}(M)$.

$T : M \to M$ has the following properties:

(A1) There exists a compact set $N \subset M$ such that $N \setminus W^s(0)$ is forward invariant under T. There exists an attracting set $\Lambda = \bigcap_{n\geq 0} T^n(N)$ with $T \mid \Lambda$ topologically transitive.

(A2) On N there exists an invariant cone field $\mathcal{C}^u$ which is mapped strictly inside itself. That is for all $x \in N$, $DT(x)\mathcal{C}^u(x) \subset \mathcal{C}^u(T(x))$.

(A3) There exists $C > 0$ and $\lambda > 1$ such that for all $v \in \mathcal{C}^u(x)$, $x \in N$, we have $|DT^n(x)v| \geq C\lambda^n |v|$, $\forall\, n \geq 0$.

The map T has a stable foliation, see [19]. That is, there exists a T-invariant $C^{1+\epsilon}$ foliation into stable leaves (including the singular leaf $W^s(0) \equiv \{x_1 = 0\}$), and a constant $\lambda_0 \in (0,1)$ such that for all x, y in the same leaf and all $n \geq 1$, $|T^n x - T^n y| \leq C\lambda_0^n$. Taking the quotient along stable leaves, a $C^{1+\epsilon}$ one-dimensional expanding map is obtained with a singularity at 0. This process can be reversed to recover the Lorenz flow T_t (with stable foliation corresponding to that of the stable foliation for T). This gives rise to a geometric Lorenz flow which we define as a three-dimensional flow with an equilibrium satisfying the eigenvalue conditions (20), and a return map $T : M \to M$ satisfying (A1), (A2) and (A3).

Verification of Hypotheses of Theorem 3.1 for Lorenz maps. It suffices to check that $T : M \to M$ satisfies the conditions of Definition 1.1. The hyperbolicity assumptions (A1), (A2) and (A3) of T ensure that Conditions (1) and (3) of Definition 1.1 are satisfied. The map $T : M \to M$ can be modeled by a Young tower with exponential tails.

Condition (2) of Definition 1.1 follows from the fact that the density for the Lorenz map is in L^∞ and $DT_u(x) \approx |x|^\alpha + g_0(x)$ ($-1 < \alpha < 0$ and $g_0(x)$ continuous and bounded) and hence DT_u is integrable with respect to μ.

4.4.1. *Lorenz flows.* We now consider limit laws for Lorenz flows. For the map $T : M \to M$ described above we can first regard the geometric Lorenz flow as a suspension flow $T_t : M^h \to M^h$ with $T_t(x, u) = (x, u+t)/\sim$ where the equivalence is $(x, h(x)) \sim (T(x), 0)$ and $M^h = \{(x, u) \in M \times \mathbb{R},\, 0 \leq u < h(x)\}/\sim$. The flow T_t admits an ergodic measure $\mu^h = \mu \times \text{Leb}/\overline{h}$ with $\overline{h} = \int_X h\, d\mu < \infty$. To analyze extremes we consider an observation $\phi : M^h \to \mathbb{R}$ with a unique logarithmic singularity at $(x, u) = (x_0, u_0)$. We assume that ϕ is continuous on $\mathbb{R}^3$ except possibly at (x_0, u_0). For the base transformation $T : M \to M$ we consider observations $\Phi : M \to M$ by maximizing ϕ over each fiber, namely $\Phi(x) = \max_{0 \leq u < h(x)} \phi(x, u)$. Given ϕ we define $M_T^\phi : M^h \to \mathbb{R}$ by

$$\phi_T(x, u) := \max\{\phi(T_t(x, u)) \mid 0 \leq t < T\}.$$

For the suspension flow on M^h we let d_M denote the Riemannian metric on M and define the (local) metric d_{M^h} on M^h by

$$d_{M^h}((x, u), (y, v)) = \sqrt{d_X(x, y)^2 + |u - v|^2}. \tag{21}$$

By [10, Theorem 2.5], the following result holds.

Theorem 4.1. *Let $T_t : X^h \to X^h$ be the suspension of the geometric Lorenz flow. Consider a measurable observation $\phi : M^h \to \mathbb{R}$ with unique maximum at $(x, u) = (x_0, u_0)$.*

Suppose in a neighborhood of (x_0, u_0), $\phi(x,u) = -\log d_{M^h}((x,u),(x_0,u_0))$. *Then for* μ^h *a.e.* $(x_0,u_0) \in M^h$:

$$\mu^h\{(x,u) \in \mathbb{R}^3 : \phi_T(x,u) \leq v + \log\left(\lfloor T/\overline{h} \rfloor\right)\} \to \exp\{-\rho^h(x_0,u_0)e^{-v}\}, \quad (T \to \infty).$$

where ρ^h *is the density of* μ^h.

The geometric Lorenz flow $T_t : \mathbb{R}^3 \to \mathbb{R}^3$ is at least C^1 and the Poincaré section $M \subset \mathbb{R}^3$ is a smooth transverse cross-section of the flow. To relate the geometric flow on $\mathbb{R}^3$ to that on M^h there is a projection $\pi : M^h \to \mathbb{R}^3$, $(x,t) \mapsto T_t(x)$, which is a local C^1 diffeomorphism. The invariant measure μ^h on the suspension M^h then determines a T_t-invariant measure μ on $\mathbb{R}^3$ by $\mu(A) = \mu^h(\pi^{-1}A)$ for measurable sets A. We let ρ^μ denote the density of μ. As a corollary:

Corollary 4.1. *Let* $T_t : \mathbb{R}^3 \to \mathbb{R}^3$ *be the geometric Lorenz flow. Consider an observation* $\phi : \mathbb{R}^3 \to \mathbb{R}$, *continuous except at* x_0, *with unique maximum at* $x_0 \in \mathbb{R}^3$. *Suppose in a neighborhood of* x_0, $\phi(x) = -\log d(x,x_0)$. *Then for* μ *a.e.* $x_0 \in \mathbb{R}^3$:

$$\mu\{x \in \mathbb{R}^3 : \phi_T(x) \leq v + \log\left(\lfloor T/\overline{h} \rfloor\right)\} \to \exp\{-\tilde{\rho}(x_0)e^{-v}\}, \quad (T \to \infty).$$

5. Condition $D'(u_n)$ for non-uniformly hyperbolic systems

In this section we show that $D'(u_n)$ holds for a class of diffeomorphisms that are non-uniformly hyperbolic. The main application is to the Hénon family of maps $T(x,y) = (1 - ax^2 + y, bx)$ for $a \simeq 2$ and $|b| \ll 1$. For a positive Lebesgue measure set of parameters $(a,b) \in \mathbb{R}^2$ it is shown in [23, 1] that there exists a subset $\Lambda \subset \mathbb{R}^2$ with a hyperbolic product structure and properties (P1)-(P4) of [23] hold. Moreover the partition sets $\Lambda_i \subset \Lambda$ are s-sublattice Cantor sets which consist of points whose orbits satisfy a slow recurrence to a critical set. Geometrically, the Hénon maps admit no invariant foliations of stable and unstable leaves. However the partition sets Λ_i have a good geometry in the sense of having uniform derivative growth estimates and bounded curvature estimates up until time of separation on the tower.

To keep the setting fairly general, we will consider non-uniformly hyperbolic systems with a Young Tower that satisfy (P1)-(P4). However we need to control the geometry of stable and unstable manifolds. To state the next definition we consider a sequence of integers (s_i) and let $\Lambda_{s_1,\ldots,s_r}$ denote the s-sublattice cylinder set of $\bigvee_{j=0}^{r-1}(T^R)^{-j}\Lambda$ with $T^{R_1+\ldots+R_{s_i}}\Lambda_{s_1,\ldots,s_r} \subset \Lambda$, $i < r$, a u-sublattice. In particular $T^{R_{s_1}+\ldots+R_{s_{r-1}}}\Lambda_{s_1,\ldots,s_r} = \Lambda_{s_r}$ and $T^{R_{s_1}+\ldots+R_{s_p}}\Lambda_{s_1,\ldots,s_r} = \Lambda_{s_{p+1}}$ for $p = 1,\ldots,r-1$. For any sublattice $\Lambda' \subset \Lambda$ define its spanning rectangle Q to the minimal topological rectangle containing Λ' with boundary consisting of curves in Γ^u and Γ^s.

Definition 5.1. $T : M \to M$ has a *bounded geometry* on Λ if for any cylinder $\Lambda_{s_1,\ldots,s_r}$ and Q its spanning rectangle, then for all $n \leq R(Q) = R_1 + \ldots + R_s$ there exist C^1 unstable (resp. stable) foliations U_n, V_n defined on all of Q with the properties:

(1) There exists $\lambda > 1$ such that for all parametrized curves $\tau_u(t)$ in U_n, $|DT^k\dot{\tau}_u(t)| \geq C\lambda^k|\dot{\tau}_u(t)|$ $(\forall k \leq n)$ where $\dot{\tau}_u(t)$ is the tangent vector to $\tau_u(t)$
(2) For all $\tau_u \in U_n$, $\exists\, \tilde{\kappa} > 0$ such that $\forall k \leq n$, curvature$(T^k(\tau_u)) < \tilde{\kappa}$.
(3) There exists $\tilde{\lambda} > 1, C > 0$ such that for every curve $\tau_s(t) \in V_n$ and every $x, y \in \tau_s(t)$: $d(T^k(x), T^k(y)) \leq C\tilde{\lambda}^{-k}$ $(\forall k \leq n)$.

It is shown in [1] that Hénon maps have a bounded geometry. Properties (1)-(3) in Definition 5.1 will be sufficient for $D'(u_n)$ to be proved in the absence of an invariant cone field.

We say a function $v : M \to \mathbb{R}$ has the exponential large deviations property if there exists a strictly convex function $c(\epsilon) > 0$ for $\epsilon \neq 0$, $c(0) = 0$ such that

$$\lim_{N\to\infty} \frac{1}{N} \log \mu(|\frac{1}{N} v_N - \bar{v}| > \epsilon) = -c(\epsilon)$$

where $v_N := \sum_{j=0}^{N} v \circ T^j$ and $\bar{v} = \int v d\mu$.

As a result of [17] the function $-\log|DT_u(x)|$ satisfies exponential large deviations.

Theorem 5.1. *Let $T : M \to M$ be a Hénon diffeomorphism modeled by a Young tower with exponential decay of correlations. Then, for μ a.e. x_0, the stochastic process defined by $X_n(x) = -\log(d(x_0, T^n x))$ satisfies condition $D'(u_n)$.*

Remark 5.1. Theorem 5.1 holds for two-dimensional non-uniformly hyperbolic diffeomorphisms with the properties

(1) the log derivative $\log|DT_u(x)|$ satisfies exponential large deviations.
(2) the derivative $DT \in L^1(\mu)$.
(3) T has bounded geometry in the sense of Definition 5.1.

Proposition 5.1. *Under the hypothesis of Theorem 5.1 for large enough k,*

$$\mu_C(\mathcal{E}_k(\epsilon)) \leq \mathcal{O}(1)\left(\sqrt{\epsilon} + k(-\log\epsilon)\theta^{\sqrt{-\log\epsilon}} + k(-\log\epsilon)^{\frac{3}{2}}\epsilon^{\frac{\sqrt{2}}{\sqrt{-\log\epsilon}}} + e^{-\alpha k}\right).$$

for some $\alpha > 0, 0 < \theta < 1$.

Proof. Consider the partition elements $I_{j,l,s_1,\dots,s_r}$ defined in the proof of Proposition 3.1. Using Condition (1) in Remark 5.1, we can conclude that $|DT_u^k| > (\sqrt{\lambda})^k$ for large enough k, except on a set whose measure decays exponentially in k. Choose $C > 1$,$\tilde{C} > C$ such that if $S_k := \{x : |DT_u^k(x)| < C(1+\delta)\}$ then, by bounded distortion, if $I \cap S_k \neq \emptyset$ then $I \subset \tilde{S}_k := \{x : |DT_u^k(x)| < \tilde{C}(1+\delta)\}$ for any interval I on the tower not separated till time k. We will bound the conditional measure of $\tilde{S}_k$ by $C_1 e^{-\alpha k}$. We will consider the following cases (where p is a random variable defined in what follows):

(1) $I \subset \tilde{S}_k$,
(2) $I \not\subseteq \tilde{S}_k$,
    (a) $|T^k(I)| > \delta$,
    (b) $|T^k(I)| \leq \delta$,
        (i) $p > \sqrt{-\log\delta}$,
            (A) $R_l < \sqrt{-\log\delta}$,
            (B) $R_l \geq \sqrt{-\log\delta}$,
        (ii) $p \leq \sqrt{-\log\delta}$,
            (A) $R_l < \sqrt{-\log\delta}$,
            (B) $R_l \geq \sqrt{-\log\delta}$.

The collection of all I such that $I \subset \tilde{S}_k$ has measure smaller than that of $\tilde{S}_k$. We now consider the collection of all I such that $I \not\subseteq \tilde{S}_k$ with $|T^k(I)| > \delta$. By the construction of I, separation does not occur until after time k. This allows us to use the bounded distortion estimates as in Proposition 3.1 leading to a contribution of the form $C\frac{\epsilon}{\delta}$. The summary of the argument is that if $J = I \cap \mathcal{E}_k(\epsilon)$ then $|DT_u^k| > C(1+\delta)$ implies, by bounded distortion, $\frac{|J|}{|I|} < C\frac{\epsilon}{\delta}$, (for perhaps a different C). By bounded distortion again $\mu_C(\Lambda_l \cap T^{-j}J) \le C\frac{\epsilon}{\delta}\mu_C(\Lambda_l \cap T^{-j}I)$. Summing over all I with $|T^k(I)| \ge \delta$, over all $j < R_l$ and over all l gives us a bound of $C\frac{\epsilon}{\delta}$.

Let $p = R_l - j + R_{s_1} + \cdots + R_{s_r} - k$. We will consider two cases: $p > c(\delta) := \sqrt{-\log\delta}$ and $p \le c(\delta)$. We know that $T^k(I) = T^{\hat{k}}(\Lambda_{s_r})$ with $\hat{k} = R_{s_r} - p$. As in 3.1, we estimate:

$$\mathcal{E}_k(\epsilon) \cap T^j(\Lambda_l) = \bigcup_{R_{s_1},\dots,R_{s_r}} (I_{j,l,s_1,\dots,s_r} \cap \mathcal{E}_k(\epsilon)) \subset T^j(\Lambda_l) \cap \left(\bigcup_{m=0}^{k} T^{-m}(\{R > c(\delta)\})\right). \tag{22}$$

In the first union we consider all sequences for which $R_l + \sum_{i=1}^{r} R_{s_i} > j + k$ and $R_l + \sum_{i=1}^{r-1} R_{s_i} < j + k$. Applying T^{-j} and intersecting with Λ_l, and then summing over all l we bound by

$$\sum_{R_l}\sum_{j=0}^{R_l}\sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(\{R > c(\delta)\})). \tag{23}$$

We now consider two further subcases: $R_l > c(\delta)$ and $R_l \le c(\delta)$. If $R_l < c(\delta)$ we can rewrite the sum in equation (23) as

$$\begin{aligned}
&\sum_{R_l < c(\delta)}\sum_{j=0}^{R_l-1}\sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(\{R > c(\delta)\})) \\
&\le \sum_{R_l < c(\delta)}\sum_{j=0}^{R_l-1}\sum_{m=0}^{k+c(\delta)} \mu_C(\Lambda_l \cap T^{-m}(\{R > c(\delta)\})) \\
&\le c(\delta)\sum_{R_l < c(\delta)}\sum_{m=0}^{k+c(\delta)} \mu_C(\Lambda_l \cap T^{-m}(\{R > c(\delta)\})) \\
&\le c(\delta)\sum_{m=0}^{k+c(\delta)} \mu_C(T^{-m}(\{R > c(\delta)\})) \\
&\le \mathcal{O}(1)c(\delta)(k + c(\delta))\mu_C(R > c(\delta)).
\end{aligned} \tag{24}$$

If $R_l \ge c(\delta)$, we bound equation (23) by

$$\begin{aligned}
&\sum_{R_l \ge c(\delta)}\sum_{j=0}^{R_l-1}\sum_{m=0}^{k} \mu_C(\Lambda_l \cap T^{-j-m}(R > c(\delta))) \\
&\le \sum_{R_l \ge c(\delta)} kR_l\mu_C(\Lambda_l) \\
&\le \mathcal{O}(1)k\sum_{m > \sqrt{-\log\delta}} m\theta^m \\
&\le \mathcal{O}(1)k\theta_1^{\sqrt{-\log\delta}}.
\end{aligned} \tag{25}$$

Using the estimates in equations (24) and equations (25), we bound the estimate in equation (23) by

$$\begin{aligned}
&\sum_{R_l}\sum_{j=0}^{R_l}\sum_{m=0}^{k}\mu_C(\Lambda_l\cap T^{-j-m}(R>c(\delta)))\\
&\le \mathcal{O}(1)\left(k\sqrt{-\log\delta}\,\theta_2^{\sqrt{-\log\delta}}+(-\log\delta)\theta_3^{\sqrt{-\log\delta}}\right)\\
&\le \mathcal{O}(1)k(-\log\delta)\tilde{\theta}^{\sqrt{-\log\delta}}. \qquad (26)
\end{aligned}$$

We now come to our last case: $p\le c(\delta)$. Since

$$\Lambda=T^p(T^k(I)),$$

we must have, by bounded distortion,

$$T^k(I)\subset\bigcup_{i=1}^{p}T^{-i}\{|DT(x)|>C\delta^{\frac{-1}{\sqrt{-\log\delta}}}\}. \qquad (27)$$

Let $A_p:=\bigcup_{i=1}^{p}T^{-i}\left\{|DT(x)|>C\delta^{\frac{-1}{\sqrt{-\log\delta}}}\right\}$. The measure of the union of such intervals I is bounded by

$$\sum_{R_l}\sum_{j=0}^{R_l}\sum_{m=0}^{k}\mu_C(\Lambda_l\cap T^{-j-m}(A_p)). \qquad (28)$$

Again, we break this estimate up using $R_l>c(\delta)$ and $R_l\le c(\delta)$. As done for equation (23) we get the estimate

$$\mathcal{O}(1)\left(k\sqrt{-\log\delta}\,\mu_C(A_p)+(-\log\delta)\mu_C(A_p)+k\theta_1^{\sqrt{-\log\delta}}\right). \qquad (29)$$

By using Markov's inequality, we estimate the measure of A_p to be smaller than

$$\sqrt{-\log\delta}\,\delta^{\frac{1}{\sqrt{-\log\delta}}}.$$

Using all these estimates in equation (28) we get

$$\begin{aligned}
&\mathcal{O}(1)\left(k(-\log\delta)\delta^{\frac{1}{\sqrt{-\log\delta}}}+(-\log\delta)^{\frac{3}{2}}\delta^{\frac{1}{\sqrt{-\log\delta}}}+k\theta_1^{\sqrt{-\log\delta}}\right)\\
&\le \mathcal{O}(1)\left(k(-\log\delta)^{\frac{3}{2}}\delta^{\frac{1}{\sqrt{-\log\delta}}}+k\theta_1^{\sqrt{-\log\delta}}\right).
\end{aligned}$$

Collecting all our estimates, we get

$$\mu_C(\mathcal{E}_k(\epsilon))\le\mathcal{O}(1)\left(\frac{\epsilon}{\delta}+k(-\log\delta)\theta^{\sqrt{-\log\delta}}+k(-\log\delta)^{\frac{3}{2}}\delta^{\frac{1}{\sqrt{-\log\delta}}}+e^{-\alpha k}\right).$$

Letting $\delta=\sqrt{\epsilon}$ we get

$$\mu_C(\mathcal{E}_k(\epsilon))\le\mathcal{O}(1)\left(\sqrt{\epsilon}+k(-\log\epsilon)\theta^{\sqrt{-\log\epsilon}}+k(-\log\epsilon)^{\frac{3}{2}}\epsilon^{\frac{\sqrt{2}}{\sqrt{-\log\epsilon}}}+e^{-\alpha k}\right).$$

□

Proposition 5.2. *Under the hypothesis Theorem 5.1, for k large enough,*

$$\mu(E_k) \leq \mathcal{O}(1)(\log k)^{20} e^{-\tilde{\alpha}\sqrt{\log k}}$$

for some $\tilde{\alpha} > 0$.

Proof. Since $-\log DT_u$ is Hölder, it satisfies exponential large deviations with a rate function [17]. We use large deviations to control the growth of $-\log DT_u$. We begin by noticing that $\mu(E_k) \leq \sum_{j=1}^{(\log k)^5} \mu(\mathcal{E}_j(1/k))$. The estimate for the measure of $\mathcal{E}_j(\epsilon)$ is not available for small j, so we find a $j*$ such that the estimate holds for all $j \geq j*$, and for each $j < j*$ we choose r_j such that $jr_j > j*$. We wish to show that for small j, $\mathcal{E}_j(\epsilon) \subset \mathcal{E}_l(\delta)$ with $l \geq j*$ and $\delta = \delta(DT, l, \epsilon)$.

From the large deviations estimate we conclude that for large enough jr_j,

$$\frac{1}{jr_j} \log \mu \left(\left| \frac{1}{jr_j} \sum_{s=0}^{jr_j - 1} \log DT_u(T^s x) - \lambda \right| > \epsilon \right) = -c(\epsilon)$$

from where it follows that

$$\left| \frac{1}{jr_j} \sum_{s=0}^{jr_j - 1} \log DT_u(T^s x) - \lambda \right| \leq \epsilon$$

except on a set A with $\mu(A) \leq Ce^{-c(\epsilon)jr_j}$. On the set A^c we have

$$\left| \sum_{s=0}^{jr_j - 1} \log DT_u(T^s x) - jr_j\lambda \right| \leq \epsilon jr_j$$

which implies that

$$e^{jr_j(\lambda - \epsilon)} \leq \left| \prod_{s=0}^{jr_j - 1} DT_u(T^s(x)) \right| = |DT^{jr_j} x| \leq e^{jr_j(\lambda + \epsilon)}.$$

If $x \in \mathcal{E}_j(\epsilon) \cap A^c$, then

$$|x - T^{jr_j} x| \leq \epsilon \sum_{s=0}^{r_j - 1} |DT^{sj}| \leq \epsilon(r_j - 1)M^{jr_j}$$

where $M = e^{\lambda+\epsilon}$. It follows that $\mathcal{E}_j(\epsilon) \cap A^c \subset \mathcal{E}_{jr_j}(\epsilon(r_j - 1)M^{jr_j})$.
Hence,

$$\mu(\mathcal{E}_j(\epsilon) \cap A^c) \leq \mu(\mathcal{E}_{jr_j}(\epsilon(r_j - 1)M^{jr_j})).$$

We will choose $jr_j \approx \sqrt{\log k}$ and $\epsilon = \frac{1}{k}$. Then

$$-\log(\epsilon(r_j - 1)M^{jr_j}) = \mathcal{O}(\log k)$$

which implies that there exist $\eta > 0$ such that $\sqrt{\epsilon(r_j - 1)M^{jr_j}} \leq k^{-\eta}$. Further, since, $x\theta^{\sqrt{x}} \leq \theta^{\frac{1}{2}\sqrt{x}}$ for $x \gg 1$,

$$\sqrt{\log k}\,\theta^{\left(-\frac{1}{2}\log\left(\frac{1}{k}\log k M^{\sqrt{\log k}}\right)\right)^{1/2}} \leq \sqrt{\log k}\,\theta^{\mathcal{O}(1)(\log k)^{1/2}}.$$

The third term in the estimate from Proposition 5.1 can be rewritten as

$$(-\log z)^{\frac{3}{2}} e^{-\sqrt{-2\log z}}$$

where $z = \frac{1}{k}\log k M^{\sqrt{\log k}}$, and by the choice of out sequence jr_j and ϵ we get

$$(\log k)^{\frac{3}{2}}\exp(-\sqrt{2\log k}).$$

Therefore, we get

$$\mu(\mathcal{E}_j(1/k)\cap A^c) \le k^{-\eta} + \sqrt{\log k}\,\theta^{\mathcal{O}(1)(\log k)^{1/2}} + (\log k)^{\frac{3}{2}}e^{-\sqrt{2\log k}} + e^{-\alpha\sqrt{\log k}} \tag{30}$$

This estimate was required to account for $j \le j*$. For $j > j*$ we get the direct estimate

$$\mu(\mathcal{E}_j(1/k)) \le \mathcal{O}(1)\left(\frac{1}{\sqrt{k}} + j(\log k)\theta^{\sqrt{\log k}} + j(\sqrt{\log k})^{\frac{3}{2}}e^{-\sqrt{2\log k}} + e^{-\alpha\sqrt{\log k}}\right).$$

Hence

$$\mu(E_k) = \sum_{j=1}^{(\log k)^5}\mu(\mathcal{E}_i(\frac{1}{k})) \le \mathcal{O}(1)\left(\begin{array}{c}\dfrac{(\log k)^5}{k^\eta} + (\log k)^{\frac{11}{2}}\theta^{\mathcal{O}(1)(\log k)^{\frac{1}{2}}} + \\ (\log k)^{\frac{15}{2}}e^{-\sqrt{2\log k}} + (\log k)^5 e^{-\alpha\sqrt{\log k}}\end{array}\right) + \mu(A)$$

and so, for some constant $\tilde{\alpha} > 0$, we get

$$\mu(E_k) \le \mathcal{O}(1)(\log k)^{20}e^{-\tilde{\alpha}\sqrt{\log k}}.$$

□

The remainder of the proof of Theorem 5.1 follows that of Theorem 3.1 without further modification.

References

[1] M. Benedicks and L. S. Young. Markov extensions and decay of correlations for certain Hénon maps. Asterisque No. 261 (2000), xi, 13–56.

[2] P. Collet. Statistics of closest return for some non-uniformly hyperbolic systems. *Erg. Th. Dyn. Syst.* **21** (2001), 401-420.

[3] D. Dolgopyat. 'Limit theorems for partially hyperbolic systems' *Trans AMS* **356** (2004) 1637-1689.

[4] A. C. M. Freitas and J. M. Freitas. On the link between dependence and independence in extreme value theory for dynamical systems, *Stat. Probab. Lett.*, **78**, (2008), 1088-1093.

[5] J. Freitas and A. Freitas. 'Extreme values for Benedicks Carleson maps'. To appear *Erg. Thy. Dyn. Sys.*.

[6] J. Freitas, A. Freitas and M. Todd. Hitting Times and Extreme Values, To appear *Probability Theory and Related Fields*.

[7] J. Galambos. *The Asymptotic Theory of Extreme Order Statistics*, John Wiley and Sons, 1978.

[8] J. Guckenheimer and R. F. Williams. Structural stability of Lorenz attractors. *Inst. Hautes Études Sci. Publ. Math.* **50** (1979) 59–72.

[9] M. Hirata. Poisson Limit Law for Axiom-A diffeomorphisms *Erg. Thy. Dyn. Sys., 13* , **533–556**,(1993)

[10] M. P. Holland, and M. Nicol and A. Török. Extreme value distributions for non-uniformly hyperbolic dynamical systems. Preprint 2009.

[11] A. Katok and J. M. Streleyn. Invariant manifolds, entropy and billiards; smooth maps with singularities, Springer Lecture Notes in Math. **1222**, (1986)

[12] M. R. Leadbetter, G. Lindgren and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, 1980.

[13] C. Liverani and M. Wojtkowski. *Dynamics Reported.* **4** (1995) 130–202.

[14] E. D. Lorenz. Deterministic nonperiodic flow. *J. Atmosph. Sci.* **20** (1963) 130–141.

[15] S. Luzzatto, I. Melbourne and F. Paccaut. The Lorenz attractor is mixing. *Commun. Math. Phys.* **260** (2005) 393–401.

[16] P. Mattila. *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability.* Cambridge Studies in Advanced Mathematics, **44**, 1995.

[17] I. Melbourne and M. Nicol, 'Large deviations for nonuniformly hyperbolic systems', *Transactions of the AMS*, **360**, (2008) 6661-6676.
[18] S. I. Resnick. *Extreme values, regular variation, and point processes*, Applied Probability Trust, **4**, Springer-Verlag, 1987.
[19] W. Tucker. The Lorenz attractor exists. *C. R. Acad. Sci. Paris* **328** (1999) 1197–1202.
[20] W. Tucker. A rigorous ODE solver and Smale's 14th problem. *Found. Comput. Math.* **2** (2002) 53–117.
[21] R. F. Williams. The structure of Lorenz attractors. *Inst. Hautes Études Sci. Publ. Math.* **50** (1979) 73–99.
[22] L. S. Young. Bowen-Ruelle measures for certain piecewise hyperbolic maps, *Trans. Amer. Math. Soc.*, **287**, (1985), 41-48.
[23] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650.

(Chinmaya Gupta) University of Houston, Department of Mathematics, 4800 Calhoun Rd., Houston TX, 77204, USA
URL: http://math.uh.edu/∼ccgupta
E-mail address: ccgupta@math.uh.edu

(Mark Holland) University of Exeter, Mathematics Research Institute, North Park Road, Exeter, EX4 4QF, UK
E-mail address: m.p.holland@exeter.ac.uk

(Matthew Nicol) University of Houston, Department of Mathematics, 4800 Calhoun Rd., Houston TX, 77204, USA
URL: http://math.uh.edu/∼nicol
E-mail address: nicol@math.uh.edu