

# EXTREMES AND EXTREMAL INDICES FOR LEVEL SET OBSERVABLES ON HYPERBOLIC SYSTEMS.

MEAGAN CARNEY, MARK HOLLAND, AND MATTHEW NICOL

**ABSTRACT.** Consider an ergodic measure preserving dynamical system  $(T, X, \mu)$ , and an observable  $\phi : X \rightarrow \mathbb{R}$ . For the time series  $X_n(x) = \phi(T^n(x))$ , we establish limit laws for the maximum process  $M_n = \max_{k \leq n} X_k$  in the case where  $\phi$  is an observable maximized on a line segment, and  $(T, X, \mu)$  is a hyperbolic dynamical system.

Such observables arise naturally in weather and climate applications. We consider the extreme value laws and extremal indices for these observables on hyperbolic toral automorphisms, Sinai dispersing billiards and coupled expanding maps. In particular we obtain clustering and nontrivial extremal indices due to self intersection of submanifolds under iteration by the dynamics, not arising from any periodicity.

## 1. INTRODUCTION

Suppose we have a time-series  $(X_n)$  of real-valued random variables defined on a probability space  $(X, \mu)$  and let  $M_n := \max\{X_1, \dots, X_n\}$  be the sequence of successive maxima of  $(X_i)$ . There is a well-developed theory for these maximum values in the setting of  $(X_n)$  i.i.d [12, 25]. If we consider a dynamical system  $(T, X, \mu)$  such that  $T : X \rightarrow X$  and an observable  $\phi : X \rightarrow \mathbb{R}$ , we can define a stochastic process by

$$X_n = \phi \circ T^n(x)$$

for  $x \in X$ . In the case of modeling deterministic physical phenomenon,  $T$  is usually taken as an ergodic, measure-preserving transformation,  $\mu$  a probability measure and  $\phi$  is a function with some regularity, for example (locally) Hölder [39]. In extreme value literature, it is typically assumed that  $\phi$  is a function of the distance  $d(x, p)$  to a distinguished point  $p$  for some metric  $d$  so that  $\phi(x) = f(d(x, p))$  for  $x \in X$ , and  $f$  is a monotone decreasing function  $f : (0, \infty) \rightarrow \mathbb{R}$ . In this instance  $\sup_{x \in X} \phi(x) = \lim_{x \rightarrow p} \phi(x)$ , and hence the set  $\{\phi(x) \geq u\}$  corresponds to a neighborhood about  $p$ . We shall refer to the set of all points  $x \in X$  for which  $\phi(x)$  achieves its maximum (with  $\sup_{x \in X} \phi(x) = \infty$  allowed) as the extremal set  $\mathcal{S}$ . For convenience (and almost by convention) the observation  $\phi(x) = -\log d(x, p)$

---

*Date:* January 23, 2021.

*2000 Mathematics Subject Classification.* 37A50, 60G70, 37B20, 60G10, 37C25.

*Key words and phrases.* Extreme Value Theory, Return Time Statistics, Stationary Stochastic Processes, Metastability.

MC and MN were supported in part by NSF Grants DMS 1600780 and DMS 2009923. MH acknowledges support of EPSRC grant: EP/P034489/. MC thanks the Max Planck Institute-PKS, Dresden, for their hospitality while part of this work was completed.

is often used, but scaling relations translate extreme value results for one functional form to another quite easily provided the extremal set of  $\phi$  is unchanged. If the observable  $\phi(x) = -\log d(x, p)$  is changed to another function of  $d(x, p)$ , then  $\mathcal{S}$  remains equal to  $\{p\}$ . However, if the underlying extremal set  $\mathcal{S}$  is changed, e.g. going from a point to a curve, then the proofs of extreme value results and the results themselves do not translate and new approaches are required. Indeed, even if the extremal set changes from one point to another, then the extreme value laws may change (e.g.  $p$  periodic versus  $p$  non-periodic give different distributional extreme value laws) [10, 16, 20, 36].

Since the value of the function  $\phi \circ T^n(x)$  is larger the closer  $T^n(x)$  is to the extremal set  $\mathcal{S}$ , there is a close relation between extreme value statistics for the time series  $X_n = \phi \circ T^n(x)$  and return-time statistics to nested sets about  $\mathcal{S}$  [5, 8, 11, 19, 20, 26, 31]. We focus on extreme value theory in this paper but it would be possible, though computationally very difficult, to derive return time distributions which are simple Poisson (in the cases in which the extremal index is  $\theta = 1$ ) and compound Poisson (in the cases in which the extremal index  $\theta < 1$ ). The parameters in the compound Poisson distribution would in particular be difficult to compute but this would constitute an interesting investigation. We expect this work could be carried out using basically the same toolkit from extreme value theory. These parameters are calculated for functions maximized at periodic orbits in the setting of a hyperbolic toral automorphism [10] and in [4] for functions maximized at periodic orbits in Sinai dispersing billiard systems. We discuss the concept of extremal index below, it is a number  $0 \leq \theta \leq 1$  which roughly quantifies the clustering of exceedances. We will say  $\theta = 1$  is a trivial extremal index and  $\theta < 1$  a nontrivial extremal index. For results along these lines see [10, 20, 24].

Recent literature has focused on the case where the extremal set  $\mathcal{S}$  is a single point  $\{p\}$ . In this paper we address some scenarios of interest where the observable is maximized on sets other than unique points in phase space, and in turn describe how the extreme value law depends on the geometry of  $\mathcal{S}$ . We also describe a dynamical mechanism giving rise to a nontrivial extremal index which is not due to periodicity. The recent preprint [30] provides a different and axiomatic approach to determining the limit laws (especially simple and compound Poisson distributions) for entry times into neighborhoods of sets of measure zero in dynamical systems. They present similar results to this paper on coupled map lattices and consider other dynamical and statistical examples, including some systems with polynomial decay of correlations. We address here cases that are not easily captured by axiomatic approaches. This happens for example, if the extremal set  $\mathcal{S}$  fails certain transversality assumptions relative to the local (or global) stable and unstable manifolds of the system. We discuss these situations further in Sections 2.1 and 4.

**1.1. Background on extremes for dynamical systems.** Suppose  $(X_n)$  is a stationary process with probability distribution function  $F_X(u) := \mu(X \leq u)$ . We define an extreme value law (EVL) in the following way. Given  $\tau \in \mathbb{R}$ , let  $u_n(\tau)$  be a sequence satisfying

$$n\mu(X_0 > u_n(\tau)) \rightarrow \tau, \quad (1.1)$$

as  $n \rightarrow \infty$ . We say that  $(X_n)$  satisfies an extreme value law if

$$\mu(M_n \leq u_n(\tau)) \rightarrow e^{-\theta\tau} \quad (1.2)$$

for some  $\theta \in (0, 1]$ . Here,  $\theta$  is called the extremal index and  $\frac{1}{\theta}$  roughly measures the average number of exceedances in a time window given that one exceedance has occurred. When  $(X_n)$  is i.i.d. and has a regularly varying tail it can be shown that this limit exists and  $\theta = 1$ .

In the dependent setting for stationary  $(X_n)$  the existence of an EVL has been shown provided dependence conditions  $D(u_n)$  (mixing condition) and  $D'(u_n)$  (recurrence condition) or similar conditions hold for the system [38, 19]. Freitas et al [17], 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 implies the existence of an EVL and is easier to check in the dynamical setting. We describe more precisely these three conditions below.

There are, however, no general techniques for proving conditions  $D_2(u_n)$  and  $D'(u_n)$  and checking the latter is usually hard.  $D'(u_n)$  is a short returns condition that is not implied by an exponential decay of correlations. However  $D_2(u_n)$  often follows from a suitable rate of decay of correlations. Collet [8] used the rate of decay of correlation of Hölder observations to establish  $D(u_n)$  for certain one-dimensional non-uniformly expanding maps. 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 and in practice is easier to verify.

For completeness we now state conditions  $D(u_n)$ ,  $D_2(u_n)$  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  $M_n$ . We write  $F_{i_1, \dots, i_n}(u)$  for the joint distribution  $F_{i_1, \dots, i_n}(u) = \mu(X_{i_1} \leq u, X_{i_2} \leq u, \dots, X_{i_n} \leq u)$ .

**Condition  $D(u_n)$**  [38] We say condition  $D(u_n)$  holds for the sequence  $X_0, X_1, \dots$ , if for any integers  $i_1 < i_2 < \dots < i_p < j_1 < j_2 < \dots < j_{p'} \leq n$ , for which  $j_1 - i_p > t$  we have

$$|F_{i_1, i_2, \dots, i_p, j_1, j_2, \dots, j_{p'}}(u_n) - F_{i_1, i_2, \dots, i_p}(u_n)F_{j_1, j_2, \dots, j_{p'}}(u_n)| \leq \gamma(n, t)$$

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

**Condition  $D_2(u_n)$**  [17] 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  $n\gamma(n, t_n) \rightarrow 0$  as  $n \rightarrow \infty$  for some sequence  $t_n = o(n)$ ,  $t_n \rightarrow \infty$ .

**Condition**  $D'(u_n)$  [38] We say condition  $D'(u_n)$  holds for the sequence  $X_0, X_1, \dots$ , if

$$\lim_{k \rightarrow \infty} \limsup_n n \sum_{j=1}^{\lfloor n/k \rfloor} \mu(X_0 > u_n, X_j > u_n) = 0. \quad (1.3)$$

Condition  $D'(u_n)$  controls the measure of the set of points of  $(X_0 > u_n)$  which return to the set relatively quickly, and is a condition that rules out “short returns”. It is not a consequence of exponential decay of correlations and usually dynamical and geometric arguments are needed to verify Condition  $D'(u_n)$  in specific cases.

In the dynamical case if the time series of observations  $X_n = \phi \circ T^n$  satisfy  $D(u_n)$  (or  $D_2(u_n)$ ) and  $D'(u_n)$  (or some variation thereof) then an EVL holds. In these results, we have extremal index  $\theta = 1$  for observables of the form  $\phi(x) = f(d(x, p))$ , maximized at generic  $p \in X$  provided  $p$  is non-periodic [10, 16, 20, 24, 32, 36]. For periodic  $p$ , EVLs have been derived for these systems with index  $\theta < 1$  [4, 10, 16, 20, 36, 39].

For statistical estimation and fitting schemes such as block maxima or peak over thresholds methods [12], it is desirable to get a limit along linear sequences of the form  $u_n(y) = y/a_n + b_n$ . Here the emphasis is changed and the sequence  $u_n(y)$  is now required to be linear in  $y$ . For example suppose  $\phi(x) = -\log x$  is an observable on the doubling map of the interval  $[0, 1]$ ,  $Tx = (2x) \bmod 1$ , which preserves Lebesgue measure  $\mu$ . The condition  $n\mu(\phi > u_n(y)) = y$  implies  $u_n(y) = \log n - \log y$ . Furthermore we know that  $n\mu(\phi > u_n(y)) = y$  implies  $\mu(M_n \leq u_n(y)) \rightarrow e^{-y}$ . This is a nonlinear scaling. If we change variables to  $Y = -\log y$  we obtain  $n\mu(-\log x > Y + \log n) \rightarrow e^{-y} = e^{-e^{-Y}}$ , a Gumbel law.

In general if we restrict to linear scalings  $y \in \mathbb{R}$ , we obtain a limit  $n\mu(X_0 > \frac{y}{a_n} + b_n) \rightarrow h(y)$  and hence

$$\mu(a_n(M_n - b_n) \leq y) \rightarrow e^{-h(y)} = G(y), \quad (n \rightarrow \infty).$$

For i.i.d processes, if  $G$  exists and is non-degenerate, then it takes three distinct forms  $G(y) = e^{-h(y)}$  with either:

- (i)  $h(y) = e^{-y}$ ,  $y \in \mathbb{R}$  (Gumbel);
- (ii)  $h(y) = y^{-\alpha}$ ,  $y > 0$  and some  $\alpha > 0$  (Fréchet);
- (iii)  $h(y) = (-y)^\alpha$ ,  $y < 0$  and some  $\alpha > 0$  (Weibull).

These three forms can be combined into a unified *generalized extreme value* (GEV) distribution (up to scale and location  $u \rightarrow \frac{u-\alpha}{\sigma}$ ):

$$G_\xi(y) = \begin{cases} \exp\{-(1 + \xi y)^{-\frac{1}{\xi}}\}, & \text{if } \xi \neq 0; \\ \exp\{-e^{-y}\}, & \text{if } \xi = 0. \end{cases} \quad (1.4)$$

The case  $\xi = 0$  corresponds to the Gumbel distribution,  $\xi > 0$  corresponds to a Fréchet distribution, while  $\xi < 0$  corresponds to a Weibull distribution.

Numerical fitting schemes for the GEV distribution are renormalized under place and scale transformations so that the extremal index (EI) is 1 [9, Theorem 5.2]. Although it is theoretically possible to recover the EI by considering it as a function of these transformations, estimates in this way would have an undetectable level of error. Techniques to directly compute the EI, referred to as *blocks* and *runs* estimators, have been proposed [39, Section 3.4]. Both methods utilize the definition of the EI (outlined above) by numerically estimating the ratio of the number of exceedances in a cluster to the total number of exceedances. Where these differ is in their definitions of a cluster; the blocks estimator splits the data into fixed blocks of size  $k_n$  so that a cluster is defined by the number of exceedances inside each fixed block while the runs estimator introduces a run length of  $q_n$  so that any two exceedances separated by a time gap of less than  $q_n$  belongs to the same cluster. The problem with using these estimators in practical applications is their heavy dependence on choice the of  $k_n$  and  $q_n$ , respectively.

In recent literature, a more systematic approach is taken by first considering the point process describing exceedances which, under certain regularity conditions, follows some Poisson process in the limit [39]. The extremal index can then be seen as the expected value of this process and obtained through likelihood estimation. The issue of dependence in the likelihood approach is addressed in [44], and the resulting estimate of the extremal index is derived as the maximum likelihood estimator of expected value for systems satisfying condition  $D_2(u_n)$ . The setup is as follows.

Let  $q$  be a fixed quantile and  $N$  be the number of recurrences above the chosen quantile. Define  $T_i$  for  $i = 1, \dots, N-1$  as the length of time between each consecutive recurrence,  $S_i = T_i - 1$  and  $N_c = \sum_{i=1}^{N-1} \mathbb{1}_{S_i \neq 0}$ . In other words,  $N_c$  is the number of clusters found by counting the set of recurrences separated by a time gap of at least length 1. Then Süveges estimator [44] is given by,

$$\hat{\theta} = \frac{\sum_{i=1}^{N-1} qS_i + N - 1 + N_c - [(\sum_{i=1}^{N-1} qS_i + N - 1 + N_c)^2 - 8N_c \sum_{i=1}^{N-1} qS_i]^{1/2}}{2 \sum_{i=1}^{N-1} qS_i}. \quad (1.5)$$

In a similar way to [13], we use equation (1.5) to numerically estimate the extremal index of the coupled map and Anosov systems found in Section 3.7.

For dynamical systems, the corresponding problem of finding scaling constants  $a_n, b_n$  depends on both the regularity of  $\mu$  and that of the observable  $\phi(x) = f(d(x, p))$  in the vicinity of the point  $p$ .

For more general dynamical systems, these scaling relations depend on how the invariant measure scales on sets that shrink to  $p$ . This problem has been addressed in the case where  $\mu$  admits a smooth or regularly varying density function  $h$ . However, for general measures (such as Sinai Ruelle Bowen measures) and general observables, estimating  $\mu(X > y/a_n + b_n)$  becomes more delicate, see [21, 35]. However, an extreme law can still be obtained along some non-linear sequence  $u_n(y)$ , with bounds on the growth of  $u_n(y)$ , see [27].

Furthermore, for deterministic dynamical systems the extremal index parameter  $\theta$  may be nontrivial due to periodicity. For the doubling map discussed above, if  $p$  is a periodic point then  $\theta = 1 - \frac{1}{2^q}$  where  $q$  is the period of the period point (see [20, 36]).

In this article, we consider cases where  $\phi$  is maximized on a more general extremal sets  $\mathcal{S}$ . For general  $\mathcal{S}$  we cannot rely on previous methods adapted to observables of the form  $\phi = f(d(x, p))$ .

**1.2. Physical and energy-like observables.** In the study of extreme events in dynamical systems, having in mind applications to weather and climate modeling, the notion of a *physical observable* was introduced and described in [35, 40, 43]. By physical observables we mean those of form  $\phi(x) = x \cdot v$  or  $\phi(x) = x \cdot Ax$ , where  $A$  is  $d \times d$  matrix, and  $v$  a specified vector in  $\mathbb{R}^d$ . The former observable has planar level sets, while the latter has ellipsoidal level sets. In weather applications, these observables correspond to measuring (respectively) the momentum and kinetic energy of the system. The level geometries of  $\phi$  introduced additional technicalities in establishing extreme laws relative to the cases where the level sets are metric balls. These issues are discussed in detail in [35], where  $\mathcal{S}$  had a complicated geometry but its intersection with the attractor of the system was still a single point. In this article, we mainly consider energy-like observables for which the extremal set  $\mathcal{S}$  is achieved on a line segment. We also discuss other extremal sets in Section 4.

**1.3. Organization of the paper.** In Section 2 we describe our main results on: hyperbolic automorphisms of the 2-torus, Sinai dispersing billiard maps, and coupled uniformly expanding maps. We calculate the extreme value distribution, the extremal index and in some cases describe briefly the Poisson return time process. In particular we describe a method for obtaining a nontrivial extremal index which is not due to periodic behavior but rather self-intersection of a set of non-periodic points under the dynamics. Beyond existing approaches, we have to develop arguments that deal with both the geometry of  $\mathcal{S}$ , and the recurrence properties of the dynamical systems under consideration. In our examples the underlying invariant measures have regular densities with respect to Lebesgue measure. This enables us to obtain analytic results on the GEV parameters and the extremal index. We also compare our results to numerical schemes, see Section 3.7. We conclude with a discussion 4 on how the methods we have developed might be applied to general observables whose extremal sets have more complicated geometries.

## 2. STATEMENT OF RESULTS

**2.1. Hyperbolic automorphisms of the 2-torus.** We consider hyperbolic toral automorphisms of the two-dimensional torus  $\mathbb{T}^2$  induced by a matrix

$$T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

with integer entries,  $\det(T) = \pm 1$  and no eigenvalues on the unit circle. We will assume that both eigenvalues are positive in what follows to simplify the discussion and proofs.

Such maps preserve Haar measure  $\mu$  on  $\mathbb{T}^2$ . A well-known example is the Arnold Cat map

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

We consider  $\mathbb{T}^2$  as the unit square with usual identifications with universal cover  $\mathbb{R}^2$ .  $T$  preserves the Haar measure  $\mu$  on  $\mathbb{T}^2$  and has exponential decay of correlations for Lipschitz functions, in the sense that there exists  $\Lambda \in (0, 1)$  such that

$$\left| \int \phi \circ T^n \psi d\mu - \int \phi d\mu \int \psi d\mu \right| \leq C \|\phi\|_{Lip} \|\psi\|_{Lip} \Lambda^n$$

where  $C$  is a constant independent of  $\phi, \psi$  and  $\|\cdot\|_{Lip}$  is the Lipschitz norm [1].

For a set  $D$ , we define  $d_H(x, D) = \inf\{d(x, y) : y \in D\}$  (for Hausdorff distance), where  $d$  is the distance in ambient (usually Euclidean) metric.  $\overline{D}$  denotes the closure of  $D$  and we define  $D_\epsilon = \{x : d_H(x, \overline{D}) \leq \epsilon\}$  is an  $\epsilon$  neighborhood of  $D$ . As outlined in Section 1.2, the observables we consider take the form  $\phi(x) = f(d_H(x, L))$  where  $x = (x_1, x_2) \in \mathbb{T}^2$  and  $L \subset \mathbb{T}$  is a line segment with direction vector  $\hat{L}$  and finite length  $l(L)$ . The function  $f : [0, \infty) \rightarrow \mathbb{R}$  is a smooth monotone decreasing function. We will take  $f(u) = -\log(u)$ . To fix notations, we also need to later consider  $\epsilon$ -tubes around  $\mathcal{S}$ . Thus if  $\mathcal{S}$  is a line, or curve, and  $\epsilon$  is small, then  $\mathcal{S}_\epsilon$  is a thin tube.

The matrix  $DT$  has two unit eigenvectors  $v^+$  and  $v^-$  corresponding to the respective eigenvalues  $\lambda_+ = \lambda > 1$ , and  $\lambda_- = \lambda^{-1} < 1$ . We can write  $\hat{L} = \alpha v^+ + \beta v^-$  for some coefficients  $\alpha, \beta$  and so  $DT^n \hat{L} = \alpha \lambda_+^n v^+ + \beta \lambda_-^n v^-$ . If we let  $v^{(n)}$  denote a unit vector in the direction of  $DT^n \hat{L}$  and  $\alpha \neq 0, \beta = 0$  then  $v^{(n)}$  aligns with the direction  $v^+$  as  $n \rightarrow \infty$ .

If  $L$  is aligned with the unstable direction, we may lift  $L$  to  $\hat{L}$  on a fundamental domain of the cover  $\mathbb{R}^2$  of  $\mathbb{T}^2$  and write  $\hat{L} = \hat{p}_1 + tv^+, t \in [0, l(L)], \hat{p}_1 \in \mathbb{R}^2$ . Thus  $L = \pi(\hat{p}_1 + tv^+)$ ,  $t \in [0, l(L)]$  where  $\pi : \mathbb{R}^2 \rightarrow \mathbb{T}^2$  is the usual projection  $\pi : \mathbb{R}^2 \rightarrow \mathbb{R}^2/\mathbb{Z}^2$ . We write the endpoint of  $\hat{L}$  as  $\hat{p}_2$ , i.e.  $\hat{p}_2 = \hat{p}_1 + l(L)v^+$ . We will also identify the vectors  $\pi\hat{p}_1$  and  $\pi\hat{p}_2$  with the corresponding points  $p_1$  and  $p_2$  in  $\mathbb{T}^2$ . Similarly if  $L$  is aligned with the stable direction, we may lift  $L$  to  $\hat{L}$  on a fundamental domain of the cover  $\mathbb{R}^2$  of  $\mathbb{T}^2$  and write  $L = \pi(\hat{p}_1 + tv^-)$ ,  $t \in [0, l(L)], \hat{p}_1 \in \mathbb{R}^2$ . Again we write the endpoint of  $\hat{L}$  as  $\hat{p}_2$ , i.e.  $\hat{p}_2 = \hat{p}_1 + l(L)v^-$ . We will also identify the vectors  $\hat{p}_1$  and  $\hat{p}_2$  with the corresponding points they project to under  $\pi$ , written  $p_1$  and  $p_2$ .

**Theorem 2.1.** *Let  $T : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  be a hyperbolic toral automorphism with positive eigenvalues  $\lambda^+ = \lambda > 1$ ,  $\lambda^- = \frac{1}{\lambda} < 1$ . Let  $\mu$  denote Haar measure on  $\mathbb{T}^2$ . Let  $L \subset \mathbb{T}^2$  be the projection  $\mathbb{R}^2 \rightarrow \mathbb{T}^2$  of a line segment  $\hat{L}$  with finite length  $l(L)$ . Define  $\phi(x) = -\log(d_H(x, L))$ ,  $\phi : \mathbb{T}^2 \rightarrow \mathbb{R}$ . Define  $M_n(x) = \max\{\phi(x), \phi(Tx), \dots, \phi(T^{n-1}(x))\}$ . Then*

$$\lim_{n \rightarrow \infty} \mu(M_n \leq y + \log(2nl(L))) = \exp\{-\theta e^{-y}\}. \quad (2.1)$$

where the extremal index  $\theta$  is determined by these cases. If:

- (1)  $L$  is not aligned with the stable  $v^-$  or unstable  $v^+$  direction then  $\theta = 1$ .

- (2)  $L$  is aligned with the unstable direction  $v^+$  and  $\pi(\hat{p}_1 + tv^+)$ ,  $-\infty < t < \infty$  contains no periodic points then  $\theta = 1$ .
- (3) If  $L$  is aligned with the stable direction  $v^-$  and  $\pi(\hat{p}_1 + tv^-)$ ,  $-\infty < t < \infty$  contains no periodic points then  $\theta = 1$ .
- (4)  $L$  is aligned with the stable  $v^-$  or unstable  $v^+$  direction and  $L$  contains a periodic point of prime period  $q$  then  $\theta = 1 - \lambda^{-q}$ .
- (5)  $L$  is aligned with the unstable direction  $v^+$ ,  $L$  contains no periodic points but  $\pi(\hat{p}_1 + tv^+)$ ,  $-\infty < t < \infty$  contains a periodic point  $\zeta$  of prime period  $q$ ; then  $L \cap T^q L = \emptyset$  implies  $\theta = 1$ ; otherwise if  $L \cap T^q L \neq \emptyset$  then  $(1 - \lambda^{-q}) \leq \theta \leq 1$  and all values of  $\theta$  in this range can be realized depending on the length and placement of  $L$ ;
- (6)  $L$  is aligned with the stable direction  $v^-$ ,  $L$  contains no periodic points but  $\pi(\hat{p}_1 + tv^-)$ ,  $-\infty < t < \infty$  contains a periodic point of prime period  $q$ ; then  $L \cap T^q L = \emptyset$  implies  $\theta = 1$ ; otherwise if  $L \cap T^q L \neq \emptyset$  then  $(1 - \lambda^{-q}) \leq \theta \leq 1$  and all values of  $\theta$  in this range can be realized depending on the length and placement of  $L$ .

*Remark 2.2.* For cases 5, and 6 we may realize any value of  $\theta$  in the range  $[(1 - \lambda^{-q}), 1]$ . This will be demonstrated in the proof, where the value of  $\theta$  is given as a function of the locations of  $\pi\hat{p}_1$  and  $\pi\hat{p}_2$  relative to the period- $q$  point on a continuation of  $L$ . This formula is difficult to state in an elegant way in full generality.

*Remark 2.3.* In Theorem 2.1 we have focused on the particular case  $f(u) = -\log u$  which gives rise to a Gumbel distribution. For other functional forms, such as  $f(u) = u^{-\alpha}$ , ( $\alpha > 0$ ) we obtain corresponding limit laws.

*Remark 2.4.* Since all periodic points of  $T$  have rational coefficients  $(\frac{p_1}{q_1}, \frac{p_2}{q_2})$  and  $v^+, v^-$  have irrational slopes it follows that if  $\pi(\hat{p}_1 + tv^+)$ ,  $-\infty < t < \infty$  contains a periodic point it contains at most one, and similarly for  $\pi(\hat{p}_1 + tv^-)$ ,  $-\infty < t < \infty$ .

Using exponential decay of correlations of the map, we show that for small  $\epsilon$ -tubes  $L_\epsilon$  around  $L$ , we have (for all  $j$  sufficiently large)  $\mu(T^j L_\epsilon \cap L_\epsilon) \leq C\mu(L_\epsilon)^2$ . This enables us to easily verify the form of the  $D(u_n)$  condition of Leadbetter et al [38] that we use.

The argument in the case that  $L$  is aligned with  $v^+$  turns out to be the most subtle. We need a detailed analysis of how the forward images  $T^j L$  wrap around the torus. It is clear that these forward images are dense, but we need quantitative information on how quickly these images become uniformly distributed. Such considerations are not necessary in the case where  $\mathcal{S}$  is a single point, e.g. as discussed in [10], and furthermore this scenario is not easily captured by axiomatic approaches, as discussed in [5, 30]. The close alignment of  $\mathcal{S}$  with the unstable manifold appears non-generic in this hyperbolic toral automorphism example. However, for general observables one could imagine level set geometries failing transversality conditions generically, e.g. if  $\{\phi > u_n\}$  has a non-trivial boundary, which perhaps coils or accumulates upon itself. These scenarios would have to be treated on a case by case basis.



**2.2. Sinai dispersing billiards maps.** We now consider another setting in which it is natural to have a smooth observable maximized on a line segment. Suppose  $\Gamma = \{\Gamma_i, i = 1 : k\}$  is a family of pairwise disjoint, simply connected  $C^3$  curves with strictly positive curvature on the two-dimensional torus  $\mathbb{T}^2$ . The billiard flow  $B_t$  is the dynamical system generated by the motion of a point particle in  $Q = \mathbb{T}^2 / (\cup_{i=1}^k (\text{convex hull of } \Gamma_i))$  which moves with constant unit velocity inside  $Q$  until it hits  $\Gamma$ , then it undergoes an elastic collision where angle of incidence equals angle of reflection. If each  $\Gamma_i$  is a circle and the system is lifted periodically to  $\mathbb{R}^2$  then this system is called a periodic Lorentz gas and was a model in the pioneering work of Lorentz on electron motion in conductors.

It is often easier to consider the billiard map  $T : \partial Q \rightarrow \partial Q$ , derive statistical properties for it and then deduce corresponding properties for the flow. In this paper we will focus on limit laws for the billiard map. Let  $r$  be the natural one-dimensional coordinate of  $\Gamma$  given by arc-length and let  $n(r)$  be the outward normal to  $\Gamma$  at the point  $r$ . For each  $r \in \Gamma$  the tangent space at  $r$  consists of unit vectors  $v$  such that  $(n(r), v) \geq 0$ . We identify such a unit vector  $v$  with an angle  $\vartheta \in [-\pi/2, \pi/2]$ . The phase space  $M$  is then parametrized by  $M := \partial Q = \Gamma \times [-\pi/2, \pi/2]$ , and  $M$  consists of the points  $(r, \vartheta)$ .  $T : M \rightarrow M$  is the Poincaré map that gives the position and angle  $T(r, \vartheta) = (r_1, \vartheta_1)$  after a point  $(r, \vartheta)$  flows under  $B_t$  and collides again with  $M$ , according to the rule angle of incidence equals angle of reflection. The billiard map preserves a measure  $d\mu = c_M \cos \vartheta dr d\vartheta$  equivalent to 2-dimensional Lebesgue measure  $dm = dr d\vartheta$  with density  $\rho(r, \theta) = c_M \cos \vartheta$ .

For this class of billiards the stable and unstable foliations lie in strict cones  $C^u$  and  $C^s$  in that the graphs  $\vartheta(r)$  of local unstable manifolds have uniform bounds on the slopes of their tangent vectors which lie in the cone  $C^u$ ,  $s_0 \leq \frac{d\vartheta}{dr} \leq s_1$ , and similarly tangents to local stable manifolds lie in a cone  $C^s$ ,  $-t_1 \leq \frac{d\vartheta}{dr} \leq -t_0$ , for some strictly positive constants  $s_0, t_0, s_1, t_1$ .

We will assume a line segment  $L$  with direction vector  $\hat{L}$  is uniformly transverse to  $C^s$  and  $C^u$ . More precisely, we will consider functions maximized on line segments  $L = \{x = (r, \vartheta) : x \cdot v = c\}$ ,  $v = (v_1, v_2)$ , which are transverse to the stable and unstable cone of directions. For example the line segment  $r = r_0$ , which is a position on the table rather than the point  $(r_0, \vartheta_0)$  (which is in phase space). We note that [42] studied distributional and almost sure return time limit laws to the position  $r = r_0$ . In our setting the precise extreme law (Weibull, Fréchet or Gumbell) depends upon the observable we choose but results may be transformed from one observable to another in a standard way. We will take the function  $\phi(r, \vartheta) = 1 - d_H(x, L)$  which because it is bounded will lead to a Weibull distribution. We assume the finite horizon condition, namely that the time of flight of the billiard flow between collisions is bounded above and also away from zero. Under the finite horizon condition Young [46] proved that the billiard map has exponential decay of correlations for Hölder observables. A good reference for background results for this section are the papers [2, 3, 7, 46] and the book [6].

Let  $L$  be a line segment transverse to the stable and unstable cones and  $\phi(r, \vartheta) = 1 - d_H(x, L)$ . Let  $y > 0$ . We define a sequence  $u_n(y) = y/a_n + b_n$  by the requirement  $n\mu\{\phi >$

$u_n(y)\} = y$ . Apart from complication arising from the invariant measure having a cosine term,  $a_n$  scales like  $\frac{1}{n}$ . The set  $\{\phi > u_n\}$  is a rectangle  $U_n$  with center  $L$  roughly of width  $\frac{Cy}{n}$  for some constant  $C$ . Note that we assume  $L$  is not aligned in either the unstable or the stable direction, so the following result is expected from the hyperbolic toral automorphism case.

**Theorem 2.5.** *Let  $T : M \rightarrow M$  be a planar dispersing billiard map with invariant measure  $d\mu = c_M \cos \vartheta dr d\vartheta$ , and assume that there is a finite time horizon between collisions. Suppose  $x = (r, \theta)$  and  $\phi(x) = 1 - d_H(x, L)$  where  $\hat{L}$  is not in the unstable cone  $C^u$  or the stable cone  $C^s$ . Let  $M_n(x) = \max\{\phi(x), \phi \circ T(x), \dots, \phi \circ T^{n-1}(x)\}$ . Then  $\mu(M_n \leq u_n(y)) \rightarrow e^{-y}$  as  $n \rightarrow \infty$ . In particular the extreme index  $\theta = 1$ .*

*Remark 2.6.* We now make some remarks on what we conjecture in the case that a line segment or curve  $L$  is contained in, i.e. a piece of, a local unstable or local stable manifold and  $\phi(x) = 1 - d_H(x, L)$ . If  $L$  is part of a local unstable manifold and  $T^n L$  has no self-intersections with  $L$  then the extremal index is one. The proofs we give in the case of the hyperbolic toral automorphism for this scenario break down but the techniques of the recent preprint [45] probably extend to this case. If  $L$  contains a periodic point  $\zeta$  of period  $q$  then the extremal index would be roughly  $\theta \sim 1 - \frac{1}{|DT_u(\zeta)|^q}$  where  $DT_u(\zeta)$  is the expansion in the unstable direction at  $\zeta$  with a correctional factor due to the conditional measure on the unstable manifold which contains  $L$ . If  $L$  does not contain a periodic point but its continuation in the unstable manifold does contain a periodic point of period  $q$  then as in case (5) of Theorem 2.1, if  $T^q L \cap L = \emptyset$  then  $\theta = 1$ , otherwise we expect  $\theta$  to lie roughly in the range  $1 - \frac{1}{|DT_u(\zeta)|^q} \leq \theta \leq 1$  (with all values of  $\theta$  being realizable depending on the length and placement of  $L$ ). If  $L$  is part of a local stable manifold and  $T^n L$  has no self-intersections with  $L$  then the extremal index  $\theta = 1$ . If  $L$  contains a periodic point  $\zeta$  of period  $q$  then the extremal index would be roughly  $\theta \sim 1 - |DT_s(\zeta)|^q$  where  $DT_s(\zeta)$  is the expansion in the stable direction at  $\zeta$ . If  $L$  does not contain a periodic point but its continuation in the unstable manifold does contain a periodic point of period  $q$  then as in case (6) of Theorem 2.1, if  $T^q L \cap L = \emptyset$  then  $\theta = 1$ , otherwise we expect  $\theta$  to lie roughly in the range  $1 - |DT_s(\zeta)|^q \leq \theta \leq 1$  (with all values of  $\theta$  being realizable depending on the length and placement of  $L$ ).

**2.3. Coupled systems of uniformly expanding maps.** Now we consider a simple class of coupled mixing expanding maps of the unit interval, similar to those examined in [13]. In fact we were motivated by the comprehensive work of [13] (which uses sophisticated transfer operator techniques) to develop in this paper an alternate probabilistic approach in a coupled maps setting. The recent preprint [30] presents similar results to ours in the case of returns to the diagonal  $\{x_1 = x_2 = \dots = x_n\}$ . Let  $T$  be a  $C^2$  uniformly expanding map of  $S^1$  and suppose that  $T$  has an invariant measure  $\mu$  with density  $h$  bounded above and below from zero. In [13] piecewise  $C^2$  expanding maps were considered but we will limit our discussion to smooth maps. We use all-to-all coupling and first discuss the case of two coupled maps for clarity.

Let  $0 < \gamma < 1$  and define

$$F(x, y) = ((1 - \gamma)Tx + \frac{\gamma}{2}(Tx + Ty), (1 - \gamma)Ty + \frac{\gamma}{2}(Tx + Ty)) \quad (2.2)$$

so that  $F : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ . We assume that  $F$  has an invariant measure  $\mu$  on  $\mathbb{T}^2$  with density  $\tilde{h}$  on  $\mathbb{T}^2$  bounded above and also bounded below away from zero almost surely. We will require also that there exists  $\epsilon > 0$  and  $0 < \alpha \leq 1$  such that

$$|\tilde{h}|_\alpha := \operatorname{ess\,sup}_{0 < \epsilon < \epsilon_0, x \in \mathbb{T}^2} \frac{1}{\epsilon^\alpha} \int \operatorname{osc}(h, B_\epsilon(x)) dm < \infty$$

where  $\operatorname{osc}(h, A) = \operatorname{ess\,sup}_{x \in A} h(x) - \operatorname{ess\,inf}_{x \in A} h(x)$  for any measurable set  $A$ . The semi-norm  $|\cdot|_\alpha$  and this notion of regularity was described in [13] and established in several of their examples. An invariant density for  $F$  cannot reasonably be assumed to be continuous or Lipschitz. For example a slight perturbation of the doubling map of the unit circle  $T(x) = (2x) \pmod{1}$  to the map  $T(x) = ((2 + \epsilon)x) \pmod{1}$  gives rise to a map with invariant density which is of bounded variation but not Lipschitz or even continuous.  $|\cdot|_\alpha$  can be completed to a norm  $\|\cdot\|_{\operatorname{osc}, \alpha}$  by defining  $\|\cdot\|_{\operatorname{osc}, \alpha} = |\cdot|_\alpha + \|\cdot\|_1$ . The value of  $\epsilon_0$  and  $\alpha$  does not matter in our subsequent discussion. We note that the bounded variation norm and the quasi-Hölder norm  $\|\cdot\|_{\operatorname{osc}, \alpha}$  are particularly suited to handle dynamical systems with discontinuities or singularities.

We also assume a strong form of exponential decay of correlations in the sense that for all Lipschitz  $\Phi$ ,  $L^\infty$   $\Psi$  on  $\mathbb{T}^2$  there exists  $C_1 > 0$  and  $C_2 > 0$  such that for all  $n$

$$\Theta_n(\Phi, \Psi) := \left| \int \Phi \cdot \Psi \circ F^n d\mu - \int \Phi d\mu \int \Psi d\mu \right| \leq C_1 e^{-C_2 n} \|\Phi\|_{\operatorname{Lip}} \|\Psi\|_\infty, \quad (2.3)$$

where  $\|\cdot\|_{\operatorname{Lip}}$  denotes the Lipschitz norm and  $\|\cdot\|_\infty$  denotes the  $L^\infty$  norm.

The function  $\Theta_n(\Phi, \Psi)$  is called the correlation function.

Let  $\phi(x, y) = -\log|x - y|$ , a function maximized on the line segment (or circle)  $L = \{(x, y) : y = x\}$ . In this setting  $L$  is invariant under  $F$  and the orthogonal direction to  $L$  is uniformly repelling. Note that the projection of  $(x, y)$  onto  $L$  is the point  $(\frac{x+y}{2}, \frac{x+y}{2})$  and the projection on  $L^\perp$  is  $(x - \frac{x+y}{2}, y - \frac{x+y}{2})$ . Close to  $L$  we have uniform expansion away from  $L$  in the  $L^\perp$  direction under  $F$ . This is because  $y - x \mapsto (1 - \gamma)[Ty - Tx]$  under  $F$  so writing  $y - x = \epsilon$  we see  $\epsilon \mapsto (1 - \gamma)[T(x + \epsilon) - Tx] \sim (1 - \gamma)DT(x)\epsilon + O(\epsilon^2)$ . There is uniform repulsion away from the invariant line  $L$ . This observation simplifies many of the geometric arguments we use to establish extreme value laws.

In the more general case of  $m$ -coupled maps we define

$$F(x_1, x_2, \dots, x_m) := (F_1(x_1, x_2, \dots, x_m), \dots, F_m(x_1, x_2, \dots, x_m)),$$

with

$$F_j(x_1, x_2, \dots, x_m) = (1 - \gamma)T(x_j) + \frac{\gamma}{m} \sum_{k=1}^m T(x_k), \quad (2.4)$$

for  $j \in [1, \dots, m]$ . For these maps, we assume:

- (A) there exists a mixing invariant measure  $\mu$  with density  $\tilde{h}$ ,  $\|\tilde{h}\|_{osc,\alpha} < \infty$ , on  $\mathbb{T}^m$  bounded above and below away from zero;
- (B) exponential mixing for Lipschitz functions versus  $L^\infty$  functions as in Equation 2.3.

*Remark 2.7.* Using the spectral analysis of the transfer operator of this system as in [13] and standard perturbation theory it can be shown that (A) and (B) hold if  $\gamma$  is sufficiently small as the uncoupled system is uniformly expanding. An example explicitly given in [13, Example 2.2] is to take  $T(x) = 3x \pmod{1}$  with sufficiently small coupling  $\gamma > 0$ . Note that Lipschitz functions are quasi-Hölder in the sense of [13].

We consider a function maximized on  $L = \{(x_1, x_2, \dots, x_m) : x_1 = x_2 = \dots = x_m\}$ . The component of a point or vector  $x = (x_1, x_2, \dots, x_m)$  orthogonal to  $L$  is  $x^\perp = (x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_m - \bar{x})$  where  $\bar{x} = \frac{1}{m} \sum_{j=1}^m x_j$ . We define  $\|(x_1, x_2, \dots, x_m)\| = \max_j |x_j|$  and define for  $x = (x_1, x_2, \dots, x_m)$

$$\phi(x) = -\log(\|x^\perp\|).$$

The function  $\phi$  is maximized on  $L$ , and large values of  $\phi \circ F^n(x)$  indicate the orbit of  $x$  is close to full synchrony of the coupled systems at time  $n$ . Writing  $p_i = x_i - \bar{x}$  we have  $\sum_{i=1}^m p_i = 0$ . Note if we have a vector  $(\Delta p_1, \Delta p_2, \dots, \Delta p_m)$  orthogonal to  $L$  we have  $\sum_{i=1}^m \Delta p_i = 0$ . Thus in a sufficiently small neighborhood of  $L$  we may write (for  $j \in [1, \dots, m]$ )

$$\begin{aligned} F_j(x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_m - \bar{x}) &= (1 - \gamma)DT\Delta p_j + \frac{\gamma}{m} \sum_{k=1}^m DT\Delta p_k + O(\max_k \Delta p_k)^2, \\ &= (1 - \gamma)DT\Delta p_j + O(\max_k \Delta p_k)^2 \end{aligned}$$

where we have used twice-differentiability and the fact that  $\sum_{i=1}^m \Delta p_i = 0$ . Hence again there is uniform expansion in a sufficiently small neighborhood of  $L$  in the direction of the  $n - 1$  dimensional subspace orthogonal to  $L$ .

For  $y > 0$  define  $u_n(y)$  by  $n\mu(\phi > u_n(y)) = y$ , and  $U_n = \{\phi > u_n(y)\}$ . It can be seen that if  $F$  is a map of  $\mathbb{T}^m$  then  $u_n \sim \frac{1}{m}[\log n - \log y]$ , the precise relation depends upon the density  $\tilde{h}$  of the invariant measure. The precise functional form of  $\phi$  is not important as a different choice of  $\phi$  would lead to a different scaling.

**Theorem 2.8.** *Let  $F : \mathbb{T}^m \rightarrow \mathbb{T}^m$  be a coupled system of expanding maps satisfying (A) and (B). Define  $p^\perp = (x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_m - \bar{x})$  where  $\bar{x} = \frac{1}{m} \sum_{j=1}^m x_j$ . Suppose  $\phi(p) = -\log(\|p^\perp\|)$ . Let  $M_n(x) = \max\{\phi(x), \phi \circ F(x), \dots, \phi \circ F^{n-1}(x)\}$ . Then  $\mu(M_n \leq u_n(y)) \rightarrow e^{-\theta y}$  as  $n \rightarrow \infty$  where*

$$\theta = 1 - \left[ \int_L \frac{1}{[(1 - \gamma)DT(s)]^{m-1}} \tilde{h}(s) ds \right].$$

where  $s$  is the natural co-ordinatization of the 1 dimensional subspace  $L$ .

We may also consider blocks of synchronization, as in [13, Section 7.2] where we take the observable maximized on a set  $L$  consisting of synchrony on subsets of distinct lattice sites, for example of form  $L = \{(x_1, x_2, \dots, x_m) : x_{i_1} = x_{i_2} = \dots = x_{i_k}, x_{j_1} = x_{j_2} = \dots = x_{j_l}\}$ . The main purpose of this section is to illustrate our geometric approach, so we will give one result of this type. Recall that for  $y > 0$  the scaling constants  $u_n(y)$  are defined by the condition  $n\mu(\phi > u_n(y)) = y$ .

**Theorem 2.9.** *Let  $F : \mathbb{T}^m \rightarrow \mathbb{T}^m$  be a coupled system of expanding maps satisfying (A) and (B). Let  $0 < k \leq m$  and choose  $k$  distinct lattice sites  $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ . Define the subspace  $L = \{(x_1, x_2, \dots, x_m) : x_{i_1} = x_{i_2} = \dots = x_{i_k}\}$  of dimension  $m - k + 1$  and  $\bar{x} = \frac{1}{k} \sum_{j=1}^k x_{i_j}$ .*

*Suppose  $\phi(p) = -\log(\max_{j=1, \dots, k} |x_{i_j} - \bar{x}|)$ . Let  $M_n(x) = \max\{\phi(x), \phi \circ F(x), \dots, \phi \circ F^{n-1}(x)\}$ . Then  $\mu(M_n \leq u_n(y)) \rightarrow e^{-\theta y}$  as  $n \rightarrow \infty$  where*

$$\theta = 1 - \left[ \int_L \frac{1}{[(1-\gamma)DT(s)]^{k-1}} \tilde{h}(s) ds \right]$$

*where  $y$  is the natural co-ordinatization of the  $m - k + 1$  dimensional subspace  $L$ .*

### 3. EXTREME VALUE SCHEME OF PROOF

Our proofs are based on ideas from extreme value theory. We will use two conditions, adapted to the dynamical setting, introduced in the important work [22] that are based on  $D(u_n)$  and  $D_2(u_n)$  but also allow a computation of the extremal index.

Let  $X_n = \phi \circ T^n$  and define

$$A_n^{(q)} := \{X_0 > u_n, X_1 \leq u_n, \dots, X_q \leq u_n\}$$

For  $s, l \in \mathbb{N}$  and a set  $B \subset M$ , let

$$\mathcal{W}_{s,l}(B) = \bigcap_{i=s}^{s+l-1} T^{-i}(B^c).$$

Next we describe the two conditions introduced in [22]. In the following recall that  $u_n \equiv u_n(\tau)$  is the sequence defined in equation (1.1).

**Condition  $\mathcal{D}_q(u_n)$ :** We say that  $\mathcal{D}_q(u_n)$  holds for the sequence  $X_0, X_1, \dots$  if, for every  $\ell, t, n \in \mathbb{N}$

$$|\mu(A_n^{(q)} \cap \mathcal{W}_{t,\ell}(A_n^{(q)})) - \mu(A_n^{(q)}) \mu(\mathcal{W}_{0,\ell}(A_n^{(q)}))| \leq \gamma(q, n, t),$$

where  $\gamma(q, n, t)$  is decreasing in  $t$  and there exists a sequence  $(t_n)_{n \in \mathbb{N}}$  such that  $t_n = o(n)$  and  $n\gamma(q, n, t_n) \rightarrow 0$  when  $n \rightarrow \infty$ .

We consider the sequence  $(t_n)_{n \in \mathbb{N}}$  given by condition  $\mathcal{D}_q(u_n)$  and let  $(k_n)_{n \in \mathbb{N}}$  be another sequence of integers such that as  $n \rightarrow \infty$ ,

$$k_n \rightarrow \infty \quad \text{and} \quad k_n t_n = o(n).$$

**Condition  $\mathcal{D}'_q(u_n)$ :** We say that  $\mathcal{D}'_q(u_n)$  holds for the sequence  $X_0, X_1, \dots$  if there exists a sequence  $(k_n)_{n \in \mathbb{N}}$  as above and such that

$$\lim_{n \rightarrow \infty} n \sum_{j=q+1}^{\lfloor n/k_n \rfloor} \mu(A_n^{(q)} \cap T^{-j}(A_n^{(q)})) = 0.$$

We note that, taking  $U_n := \{X_0 > u_n\}$  for  $A_n^{(q)}$ , which corresponds to non-periodic behavior, in condition  $\mathcal{D}'_q(u_n)$  corresponds to condition  $\mathcal{D}'(u_n)$  from [38]. We will abuse notation and consider  $U_n := \{X_0 > u_n\}$  as the case of  $A_n^{(q)}$  with  $q = 0$ .

Now let

$$\theta = \lim_{n \rightarrow \infty} \theta_n = \lim_{n \rightarrow \infty} \frac{\mu(A_n^{(q)})}{\mu(U_n)},$$

whenever this limit exists. In the settings we consider the limit is easily seen to exist.

*Remark 3.1.* In a dynamical setting verifying these two conditions picks up the main underlying periodicity or more generally recurrence properties of the system, for example returns to a periodic point of prime period  $q$ , and determines the extremal index. However, as we demonstrate, other recurrent phenomena may give rise to an extremal index not equal to unity. We show below that the self-intersection of a line segment  $L$ ,  $T(L) \cap L \neq \emptyset$  (none of whose points are periodic) may lead to a nontrivial extremal index for functions maximized on  $L$ . For a more detailed discussion of extremal indices see [22].

From [23, Corollary 2.4], it follows that to establish Theorem 2.1 it suffices to prove conditions  $\mathcal{D}_q(u_n)$  and  $\mathcal{D}'_q(u_n)$  for  $q = 0$  in the non-recurrent case  $\theta = 1$  and for  $q > 0$  corresponding to the ‘period’ of the cases where there is some recurrence phenomena ( $\theta < 1$ ). In both cases

$$\lim_{n \rightarrow \infty} \mu(M_n \leq u_n(y)) = e^{-\theta y}.$$

The scheme of the proof of Condition  $\mathcal{D}_q(u_n)$  is itself somewhat standard [10, 24] and is a consequence of suitable decay of correlation estimates. We outline it for completeness, indicating the modifications that need to be made for the different geometries of  $A_n^{(q)}$ . The main work will be in establishing Condition  $\mathcal{D}'_q(u_n)$ .

**3.1. Proof of Theorem 2.1.** In the first instance we check condition  $\mathcal{D}_q(u_n)$ . We recall some useful statistical properties of hyperbolic toral automorphisms. In the case where  $\Phi$

and  $\Psi$  are Lipschitz continuous functions, it is known for hyperbolic toral automorphisms that there exists  $C > 0$ ,  $\tau_0 \in (0, 1)$  such that

$$\left| \int \Phi(\Psi \circ T^n) d\mu - \int \Phi d\mu \int \Psi d\mu \right| \leq C \tau_0^n \|\Phi\|_{\text{Lip}} \|\Psi\|_{\text{Lip}}, \quad (3.1)$$

Furthermore if  $\Psi$  is constant on local stable leaves corresponding to a Markov partition, then the Lipschitz norm of  $\Psi$  on the right-hand side of equation (3.1) can be replaced by the  $L^\infty$  norm [46, Section 4]. This fact will be useful when checking  $\mathcal{I}_q(u_n)$ , see Proposition 3.2 in Section 3.2 below.

Consider now a set  $D$ , whose boundary  $\partial D$  is a union of a finite number of smooth curves, so that  $\mu(\partial D) = 0$ . Let  $W_1^s(x)$  denote the local stable manifold through  $x$ . We define,

$$H_k(D) := \{x \in D : T^k(W_1^s(x)) \cap \partial D \neq \emptyset\}. \quad (3.2)$$

In Section 3.2 we show roughly that  $\mu(H_k(D))$  decreases exponentially in  $k$ .

**3.2. Checking condition  $\mathcal{I}_q(u_n)$ .** This argument is a minor adjustment of similar estimates in [10, 24]. We state the following proposition.

**Proposition 3.2.** *For every  $\ell, t, n \in \mathbb{N}$ , there exists  $\lambda_0 \in (0, 1)$ , and  $C > 0$  such that*

$$\left| \mu(A_n^{(q)} \cap \mathcal{W}_{t,\ell}(A_n^{(q)})) - \mu(A_n^{(q)}) \mu(\mathcal{W}_{0,\ell}(A_n^{(q)})) \right| \leq C(n^{-2} + n^2 \lambda_0^t).$$

Condition  $\mathcal{I}_q(u_n)$  immediately follows from this. We can take  $t_n = (\log n)^5$  so that  $n\gamma(q, n, t) \rightarrow 0$ .

The proof of Proposition 3.2 is as follows. To check condition  $\mathcal{I}_q(u_n)$  we use decay of correlations. The main problem in estimating the correlation function  $\Theta_n(\Phi, \Psi)$  (recall equation (2.3)) is that the relevant indicator functions  $\Phi = 1_{A_n^{(q)}}$  and  $\Psi = 1_{\mathcal{W}_{0,\ell}(A_n^{(q)})}$  of the sets  $A_n^{(q)}$  and  $\mathcal{W}_{0,\ell}(A_n^{(q)})$  are not Lipschitz continuous. Standard smoothing methods can be used to approximate  $\Phi$ , but  $\Psi$  cannot be uniformly approximated by a Lipschitz function: the level set  $\Psi = 1$  has a geometry that becomes increasingly complex (i.e. with multiple connectivity) as  $\ell$  increases. Fortunately, we can employ a further trick to approximate  $\Psi$ . This is done using a function that is constant on local stable manifolds. This allows us to use a decay of correlations estimate using the  $L^\infty$  norm. As part of this approximation we first estimate  $\mu(H_k(D))$  with  $D = A_n^{(q)}$ . The geometry of the set  $A_n^{(q)}$  will be important in calculating this estimate.

**Lemma 3.3.** *Consider the set  $D = A_n^{(q)}$ . Then there exists  $C > 0$  such that, for all  $k$ ,*

$$\mu(H_k(D)) \leq C \lambda^{-k}, \quad (3.3)$$

where  $\lambda^{-1} < 1$  is the (uniform) contraction rate along the stable manifolds for the hyperbolic toral automorphism.

*Proof.* We follow [10, Proposition 4.1], and consider also the geometrical properties of  $D$ . Since the local stable manifolds contract uniformly there exists  $C_1 > 0$  such that  $\text{dist}(T^n(x), T^n(y)) \leq C_1 |\lambda|^{-n}$  for all  $y \in W_1^s(x)$ . This implies that  $|T^k(W_1^s(x))| \leq C_1 \lambda^{-k}$ . Therefore, for every  $x \in H_k(D)$ , the leaf  $T^k(W_1^s(x))$  lies in an tubular region of width  $2/|\lambda|^k$  around  $\partial D$ . To measure of the size of this tube we note that  $m(D_\epsilon) \leq \epsilon C_D$ , where  $C = \epsilon c_q \ell_D$ . (Again recall the definition of the tubular region  $D_\epsilon$  given in Section 2.1). The constant  $c_q$  depends on the number of connected components of  $A_n^{(q)}$ , (which is bounded), and  $\ell_D$  is the maximum length of a connected component of  $\partial D$ . This is also bounded, since  $\partial D$  is formed of straight lines of bounded length. The lemma follows by taking  $\epsilon = \lambda^{-k}$ .  $\square$

The next lemma also holds for  $\{X_0 > u_n\}$  in place of  $A_n^q$ , and the proof is the same as [10, Lemma 4.2]. Again we give the main steps, indicating the role of Lemma 3.3. The constant  $\tau_1$  in the next lemma comes from the exponential decay of correlations of Lipschitz observables on hyperbolic toral automorphisms.

**Lemma 3.4.** *Suppose  $\Phi : M \rightarrow \mathbb{R}$  is a Lipschitz map and  $\Psi$  is the indicator function*

$$\Psi := \mathbf{1}_{\mathcal{W}_{0,\ell}(A_n^{(q)})}.$$

*Then there exists  $0 < \tau_1 < 1$  such that for all  $j \geq 0$*

$$\left| \int \Phi(\Psi \circ T^j) d\mu - \int \Phi d\mu \int \Psi d\mu \right| \leq C \left( \|\Phi\|_\infty \lambda^{-\lfloor j/2 \rfloor} + \|\Phi\|_{\text{Lip}} \tau_1^{\lfloor j/2 \rfloor} \right). \quad (3.4)$$

*Proof.* Following Lemma [10, Lemma 4.2], we take a version  $\bar{\Psi}$  of  $\Psi$  that is constant on local stable manifolds, for example by taking a distinguished point  $x^*$  on each local stable manifold  $W_1^s(x)$  and defining  $\bar{\Psi}(y) = \Psi(x^*)$  for all  $y \in W_1^s(x)$ . We let  $\Psi_j = \Psi \circ T^j$ , and again denote  $\bar{\Psi}_j$  as the relevant version of  $\Psi_j$  (constant on local stable manifolds). A simple application of the triangle inequality gives the following bound:

$$\Theta_j(\Phi, \Psi) \leq C \left( \|\Phi\|_\infty \mu\{\bar{\Psi}_{j/2} \neq \Psi_{j/2}\} + \|\Phi\|_{\text{Lip}} \tau_0^{j/2} \right), \quad (3.5)$$

(recall that  $\Theta_j$  is defined in equation (2.3)). To estimate  $\mu\{\bar{\Psi}_{j/2} \neq \Psi_{j/2}\}$ , we consider points  $x_1, x_2$  on the same stable manifold, and such that  $x_1 \in \mathcal{W}_{i,\ell}(A_n^{(q)})$ , but  $x_2 \notin \mathcal{W}_{i,\ell}(A_n^{(q)})$ , (for  $i \geq j/2$ ). This set is contained in  $\cup_{k=i}^{i+\ell-1} H_k(A_n^q)$ . Hence

$$\mu\{\bar{\Psi}_{j/2} \neq \Psi_{j/2}\} \leq \sum_{k=j/2}^{\infty} H_k(A_n^{(q)}) \leq C \lambda^{-j/2}.$$

The conclusion of Lemma 3.4 follows.  $\square$

To continue with the proof of Proposition 3.2, and hence verify condition  $\mathcal{D}_q(u_n)$ , we approximate the characteristic function of the set  $A_n^{(q)}$  by a suitable Lipschitz function. The key estimate is to bound the Lipschitz norm of the approximation.



Let  $A_n = A_n^{(q)}$  and  $D_n := \left\{ x \in A_n^{(q)} : d_H(x, \overline{A_n^c}) \geq n^{-2} \right\}$ , where  $\overline{A_n^c}$  denotes the closure of the complement of the set  $A_n$ . Define  $\Phi_n : \mathcal{X} \rightarrow \mathbb{R}$  by

$$\Phi_n(x) = \begin{cases} 0 & \text{if } x \notin A_n \\ \frac{d_H(x, A_n^c)}{d_H(x, A_n^c) + d_H(x, D_n)} & \text{if } x \in A_n \setminus D_n \\ 1 & \text{if } x \in D_n \end{cases} \quad (3.6)$$

Note that  $\Phi_n$  is Lipschitz continuous with Lipschitz constant given by  $n^2$ . Moreover  $\|\Phi_n - \mathbf{1}_{A_n}\|_{L^1(m)} \leq C/n^2$  for some constant  $C$ . It follows that

$$\begin{aligned} & \left| \int \mathbf{1}_{A_n^{(q)}} (\Psi_{\lfloor j/2 \rfloor} \circ T^{j - \lfloor j/2 \rfloor}) d\mu - \mu(A_n^{(q)}) \int \Psi d\mu \right| \\ & \leq \left| \int (\mathbf{1}_{A_n^{(q)}} - \Phi_n) \Psi_{\lfloor j/2 \rfloor} d\mu \right| + \mathcal{C} \left( \|\Phi_n\|_\infty j^2 \lambda^{\lfloor j/4 \rfloor} + \|\Phi_n\|_{\text{Lip}} \tau_1^{\lfloor j/2 \rfloor} \right) \\ & \quad + \left| \int (\mathbf{1}_{A_n^{(q)}} - \Phi_n) d\mu \int \Psi_{\lfloor j/2 \rfloor} d\mu \right|, \end{aligned} \quad (3.7)$$

for some generic constant  $\mathcal{C}$ . Thus

$$\left| \mu(A_n^{(q)} \cap \mathcal{W}_{j,\ell}(A_n^{(q)})) - \mu(A_n^{(q)}) \mu(\mathcal{W}_{0,\ell}(A_n^{(q)})) \right| \leq \gamma(n, j)$$

where

$$\gamma(n, j) = \mathcal{C} \left( n^{-2} + n^2 \lambda_1^{\lfloor j/2 \rfloor} \right),$$

and

$$\lambda_1 = \max \{ \tau_1, \lambda^{-1} \}.$$

Thus if, for instance, we choose  $j = t_n = (\log n)^5$ , then  $n\gamma(n, t_n) \rightarrow 0$  as  $n \rightarrow \infty$ . This completes the proof.

**3.3. Checking condition  $\mathcal{A}'_q(u_n)$ .** We make the following decomposition:

$$\begin{aligned} n \sum_{j=q+1}^{\lfloor n/k_n \rfloor} \mu(A_n^{(q)} \cap T^{-j}(A_n^{(q)})) &= n \sum_{j=q+1}^{R_n} \mu(A_n^{(q)} \cap T^{-j}(A_n^{(q)})) + n \sum_{j=R_n+1}^{(\log n)^5} \mu(A_n^{(q)} \cap T^{-j}(A_n^{(q)})) \\ &\quad + n \sum_{j=(\log n)^5+1}^{\lfloor n/k_n \rfloor} \mu(A_n^{(q)} \cap T^{-j}(A_n^{(q)})), \end{aligned}$$

where the sequence  $R_n \rightarrow \infty$  (as  $n \rightarrow \infty$ ) will be chosen later. Recall that for  $q = 0$ ,  $A_n^{(q)} = U_n$ . By exponential decay of correlations and a suitable Lipschitz approximation the last sum tends to 0 as  $n \rightarrow \infty$ , so it suffices to estimate the two sums where  $1 \leq j \leq (\log n)^5$ .

**Case  $L$  transverse to stable and unstable directions.**

Fix  $y > 0$  and define  $u_n(y)$  by the requirement  $n\mu\{x : \phi(x) \geq u_n(y)\} = y$ . Henceforth we will drop the dependence on  $y$  and write simply  $u_n$  for convenience. We define  $U_n := \{x : \phi(x) \geq u_n\}$ . Geometrically  $U_n$  resembles a parallel strip of width  $\frac{2}{n}$ .

We will verify the short return condition with  $q = 0$ . Consider the set  $T^{-j}U_n \cap U_n = \{x : T^j(x) \in U_n, x \in U_n\}$ .  $T^jU_n$  is a union of parallelogram-like strips corresponding to each winding around the torus and such strip has width  $O(\frac{|\lambda^{-j}|}{n})$  and length  $O(1)$ , the precise constants depending on the angle between  $T^jL$  and  $L$  as  $T^jL$  aligns to the unstable direction. There are approximately  $\lfloor \lambda^j \rfloor$  such parallelogram strips. Each strip intersects  $U_n$  in an area of measure  $O(\lfloor \lambda^{-j} \rfloor n^{-2})$  by transversality. See Figure 1

Hence

$$n \sum_{j=1}^{(\log n)^5} \mu(T^{-j}U_n \cap U_n) = O\left(\frac{(\log n)^5}{n}\right).$$

Thus the extremal index  $\theta = 1$ .

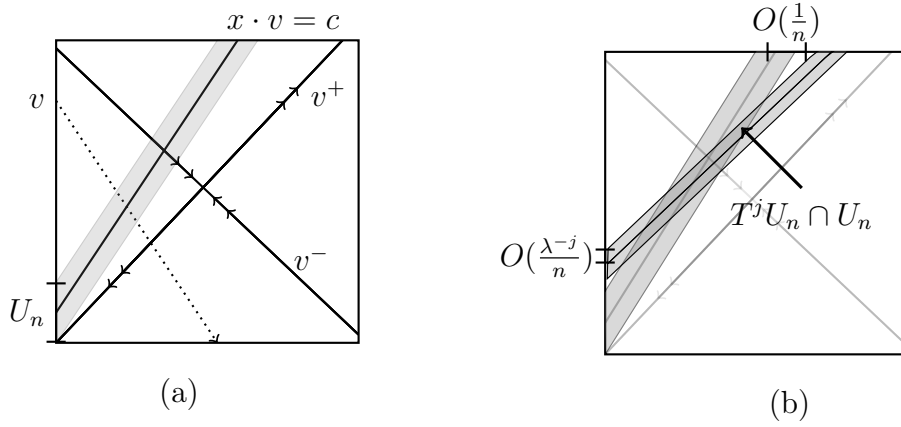


FIGURE 1. (a) The set  $U_n$  and line  $L$  for  $L$  not aligned with  $v^-$  or  $v^+$ . (b) Iterations  $T^jU_n$  and their intersections with  $U_n$ .

### Case $L$ aligned with unstable direction.

We lift  $L$  to  $\hat{L}$  on a fundamental domain of the cover  $\mathbb{R}^2$  of  $\mathbb{T}^2$  and write  $\hat{L} = \hat{p}_1 + tv^+$ ,  $t \in [0, l(L)]$ ,  $\hat{p}_1 \in \mathbb{R}^2$ . We write the endpoint of  $\hat{L}$  as  $\hat{p}_2$ , i.e.  $\hat{p}_2 = \hat{p}_1 + l(L)v^+$ . The points  $\hat{p}_1$  and  $\hat{p}_2$  project to the corresponding points written  $p_1 = \pi\hat{p}_1$  and  $p_2 = \pi\hat{p}_2$ .

There are 2 main cases, with some subcases.

Case (a): First assume that the line  $\hat{p}_1 + tv^+$ ,  $-\infty < t < \infty$  contains no point with rational coordinates. This holds for a measure one set of  $\hat{p}_1$  as the set of points in the plane with rational coordinates is countable. In this case  $T^nL$ ,  $n \geq 1$ , has no intersections with  $L$ . To see this suppose  $p \in L$  and there exists an  $n$  such that  $T^n p = q \in L$ . If we take a line segment  $\tilde{L}$  in direction  $v^+$  of length  $2l(L)$  centered at  $p$  we see by expansion that  $\tilde{L} \subset T^n \tilde{L}$  (since  $d(p, q) \leq l(L)$ ) and hence  $T^n$  restricted to  $\tilde{L}$  has a fixed point  $\tilde{p}$  in  $\tilde{L}$ . However, this implies the lift  $\hat{p}_1 + tv^+$ ,  $-\infty < t < \infty$  contains a point with rational coordinates, which is a contradiction.

Since  $p_1$  is not periodic by assumption and  $\hat{p}_1$  is not in the direction of  $v^+$  (otherwise the point  $(0, 0)$  would be contained in  $\hat{p}_1 + tv^+$ ,  $-\infty < t < \infty$ ) the iterates  $T^j U_n$  are disjoint for large  $n$  for small  $j$  i.e. there exists  $R_n \rightarrow \infty$  such that  $\mu(T^{-j} U_n \cap U_n) = 0$  for  $j < R_n$ . Corollary 2.2 of the recent preprint [45] implies in this case that the extremal index is one. We include an alternate proof for completeness.

For large  $n$  the set  $T^j U_n$  comprises  $\lfloor \lambda^j \rfloor$  parallel rectangles (aligned with the unstable direction) of width  $O(\frac{\lfloor \lambda^{-j} \rfloor}{n})$ . Identifying  $\mathbb{T}^2$  with the unit square the set  $T^j L \cap ([0, 1] \times \{0\})$  consists of  $m(j) \sim \lfloor \lambda^j \rfloor$  points  $x_i^j$ ,  $j = 1, \dots, m(j)$ . If for small iterates  $T^i L$  there is no intersection with  $([0, 1] \times \{0\})$  we extend  $T^i L$  in a straight line so that all  $x_i^j$ ,  $j = 1, \dots, m(j)$  are defined. Let  $\gamma^{-1}$  denote the slope of  $v^+$ . The set  $\{x_i^j\}_{i=1, \dots, m(j)}$  is generated by the relation  $x_1^j + k\gamma \pmod{1}$  for  $k = 1, \dots, m(j)$ .

We now estimate  $\mu(T^{-j} U_n \cap U_n)$ . The set  $T^j U_n$  has approximately  $\lfloor \lambda^j \rfloor$  windings around the torus and we now estimate the fraction of these that intersect  $U_n$ .

Note that  $\gamma$  is a quadratic irrational. This implies that  $\gamma$  has low discrepancy in the sense that there exists a constant  $C > 0$  such that

$$\sup_{0 \leq a < b \leq 1} \{ \# \{x_i^j \in (a, b)\} / \lfloor \lambda^j \rfloor - (b - a) \} \leq C \frac{\log \lfloor \lambda^j \rfloor}{\lfloor \lambda^j \rfloor},$$

see [41]. Hence for  $j > R_n$

$$n \sum_{j=R_n}^{(\log n)^5} \mu(T^{-j} U_n \cap U_n) = O\left(\frac{1}{n} + \frac{\log \lfloor \lambda^{R_n} \rfloor}{\lambda^{R_n}}\right) = o(1).$$

This implies that a standard EVL holds with  $\theta = 1$ . See Figure 2.

Case (b): Assume that  $\hat{p}_1 + tv^+$ ,  $-\infty < t < \infty$  contains a point with rational coordinates, note that it will contain at most one as the slope of  $v^+$  is irrational. Such a point projects to a point  $p_{per}$  periodic under  $T$  with period  $q$  say.

Case (b1): Assume now that  $L$  itself contains  $p_{per}$ , a periodic point of period  $q$ .

There will be only one periodic point in  $L$  as the slope of  $v^+$  is irrational. Without loss of generality we take  $q = 1$  by considering  $T^q$ . It is easy to see that  $\theta = \lim_{n \rightarrow \infty} \frac{\mu(A_n^{(q)})}{\mu(U_n)} = 1 - \frac{1}{\lambda^q}$ . The same discrepancy argument as in the case of no periodic orbits shows that there exists an  $R_n \rightarrow \infty$  such that  $T^{-j} A_n^{(q)} \cap A_n^{(q)} = \emptyset$  for  $j < R_n$  and

$$\sum_{j=R_n}^{(\log n)^5} \mu(T^{-j} A_n^{(q)} \cap A_n^{(q)}) = o\left(\frac{1}{n}\right)$$

Hence  $\theta = 1 - \frac{1}{\lambda^q}$ . See Figure 3.

Case (b2):  $L$  does not contain a periodic point.

We first consider the simplest case where the origin is the fixed point and  $\hat{p}_1$  parallel to  $v^+$  so that  $\hat{p}_1 + tv^+$ ,  $-\infty < t < \infty$  contains the fixed point  $(0,0)$  but  $L$  does not contain  $(0,0)$ . The line  $\hat{p}_1 + tv^+$ ,  $0 \leq t < \infty$  has a natural ordering by distance from the origin  $(0,0)$ . If  $\lambda\hat{p}_1 > \hat{p}_2$  then it is easy to see all iterates of  $T^n L$  on the torus are disjoint and the arguments given in case (a) apply giving  $\theta = 1$ .

Suppose now  $\lambda\hat{p}_1 < \hat{p}_2$ . We take  $q = 1$  and calculate

$$\theta = \mu(A_n^{(1)})/\mu(U_n) = |\hat{p}_2 - \frac{1}{\lambda}\hat{p}_2|/|\hat{p}_2 - \hat{p}_1| = (1 - \frac{1}{\lambda}) \frac{|\hat{p}_2|}{|\hat{p}_2 - \hat{p}_1|},$$

as the stable manifolds are sent strictly into the region of intersection  $U_n \cap TU_n$ . (See Figure 4). The condition  $\lambda\hat{p}_1 < \hat{p}_2$  implies  $1 < \frac{|\hat{p}_2|}{|\hat{p}_2 - \hat{p}_1|} < (1 - \frac{1}{\lambda})^{-1}$ . By varying  $\hat{p}_1$  and  $\hat{p}_2$  we may obtain all values in this range. Hence  $(1 - \frac{1}{\lambda}) \leq \theta \leq 1$ .

In the general case of a periodic point  $p_{per}$  of period  $q$  contained in  $\pi(\hat{p}_1 + tv^+)$ ,  $-\infty < t < \infty$  we consider  $T^q$  and the analysis proceeds in the same way by considering the expansion on the line segment  $[\hat{p}_1 - \hat{p}_{per}, \hat{p}_2 - \hat{p}_{per}]$ . We infer that for general  $q \geq 1$ ,

$$(1 - \frac{1}{\lambda^q}) \leq \theta \leq 1$$

with all values of  $\theta$  in this range being realizable. The verification of condition  $\mathcal{A}'_q(u_n)$  is similar to case (b1).

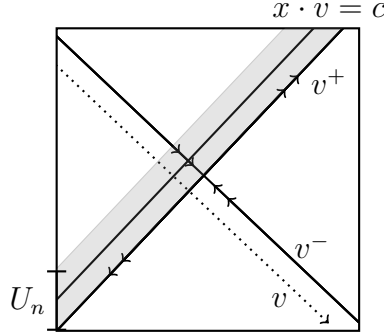


FIGURE 2. The set  $U_n$  and line  $L$  for  $L$  aligned with  $v^-$ .

### Case $L$ is aligned with the stable direction.

Suppose now that  $L$  aligns with the stable direction  $v^-$ . See Figure 5. The analysis is similar to the case where  $L$  is aligned with the unstable direction, and again we consider the lift  $\hat{L} = \hat{p}_1 + tv^-$ ,  $t \in [0, l(L)]$ , with  $\hat{p}_1 \in \mathbb{R}^2$ , and  $\hat{p}_2$  denoting the other endpoint of  $\hat{L}$ , i.e.  $\hat{p}_2 = \hat{p}_1 + l(L)v^-$ . We will make use of the time-reversibility of the system in Case (a) below.

We have the following cases.

Case (a): First assume that the line  $\hat{p}_1 + tv^-$ ,  $-\infty < t < \infty$  contains no point with rational coordinates. Let  $S = T^{-1}$ . Then  $L$  is aligned with the unstable direction for  $S$ . As in the

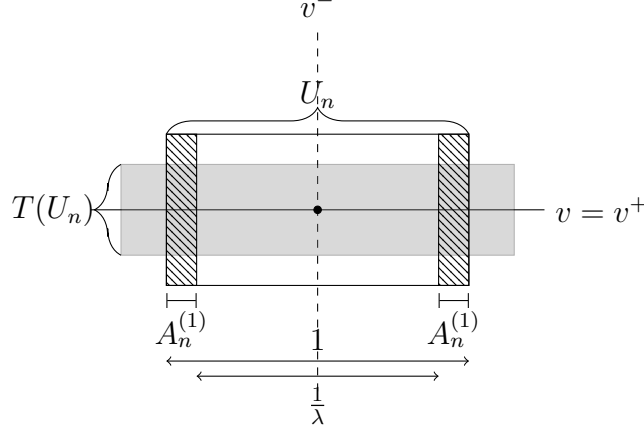


FIGURE 3. Sketch of argument (b1) for  $v$  aligned with the unstable direction and  $L$  contains a periodic orbit showing intersections of  $A_n^{(1)}$  (shown in patterned lines) and  $T(U_n)$  (shown in gray). Estimates of the ratio of  $A_n^{(1)}$  to  $U_n$  (shown in white) give the value of the extremal index.

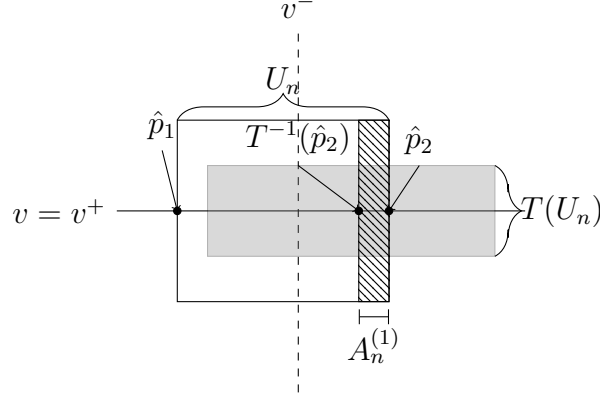


FIGURE 4. Sketch of argument (b2) for  $v$  aligned with the unstable direction and  $L$  does *not* contain a periodic orbit showing intersections of  $A_n^{(1)}$  (shown in patterned lines) and  $T(U_n)$  (shown in gray). Estimates of the ratio of  $A_n^{(1)}$  to  $U_n$  (shown in white) give the value of the extremal index.

case where  $L$  aligned with the unstable direction for  $T$ , it follows again that  $S^n(L)$  has no intersections with  $L$ , for all  $n \geq 1$ . Hence  $T^n(L)$  has no intersections with  $L$  for all  $n \geq 1$ .

Thus all the iterates  $T^j U_n$  are disjoint for small  $j$ , i.e. there exists  $R_n \rightarrow \infty$  such that  $\mu(T^{-j} U_n \cap U_n) = 0$  for  $j < R_n$ . Note that the definition of  $U_n$  is the same for  $T$  and  $S$  and that  $\mu(T^{-j} U_n \cap U_n) = \mu(U_n \cap T^j U_n) = \mu(U_n \cap S^{-j} U_n)$  by measure-preservation. The

argument of Case (a) when  $L$  is aligned with the unstable direction shows that

$$n \sum_{j=R_n}^{(\log n)^5} \mu(S^{-j}U_n \cap U_n) = o(1),$$

and hence

$$n \sum_{j=R_n}^{(\log n)^5} \mu(T^{-j}U_n \cap U_n) = o(1),$$

Thus  $\theta = 1$ .

Case (b): Assume that  $\hat{p}_1 + tv^-$ ,  $-\infty < t < \infty$  contains a point  $\hat{p}_{per}$  with rational coordinates. There will be only one such point as the slope of  $v^-$  is irrational. The point  $\hat{p}_{per}$  projects to a point  $p_{per}$  periodic under  $T$  with period  $q$  say. We cannot use time-reversibility in this case as the set  $A_n^q$  depends upon the consideration of  $T$  or  $T^{-1}$  as the transformation.

Case (b1): Assume now that  $L$  contains the periodic point  $p_{per}$  of period  $q$ .

Without loss of generality we (again) take  $q = 1$  by considering  $T^q$ . We have  $\theta = \lim_{n \rightarrow \infty} \frac{\mu(A_n^{(q)})}{\mu(U_n)} = 1 - \frac{1}{\lambda^q}$ . Geometrically  $A_n^{(q)}$  consists of two strips within  $U_n$ . Both of these have length  $1/n$ , (i.e. the same as  $U_n$ ), but their width relative to  $U_n$  is  $\frac{1}{2}(1 - 1/\lambda^q)$ . See Figure 6. The same argument as in the case of no periodic orbits shows that there exists an  $R_n \rightarrow \infty$  such that  $T^{-j}A_n^{(q)} \cap A_n^{(q)} = \emptyset$  for  $j < R_n$ .

We have uniform expansion of  $A_n^q$  in the unstable direction and the discrepancy argument of Case (b1) of the previous section (alignment with the unstable direction) shows that

$$\sum_{j=R_n}^{(\log n)^5} \mu(T^{-j}A_n^{(q)} \cap A_n^{(q)}) = o\left(\frac{1}{n}\right).$$

We therefore have  $\theta = 1 - \frac{1}{\lambda^q}$ .

Case (b2):  $L$  does not contain a periodic orbit.

Again, we illustrate by considering the simplest case of  $\hat{p}_1$  parallel to  $v^-$  so that  $\hat{p}_1 + tv^-$ ,  $-\infty < t < \infty$  contains the fixed point  $(0, 0)$  but  $L$  does not contain  $(0, 0)$ . For the lifted line  $\hat{p}_1 + tv^-$ ,  $0 \leq t < \infty$ , we use the natural ordering by distance from the origin  $(0, 0)$ . If  $\hat{p}_1 > \lambda^{-1}\hat{p}_2$ , then all iterates of  $T^n L$  on the torus are disjoint and the arguments given in case (a) apply giving  $\theta = 1$ .

Suppose now  $\hat{p}_1 < \lambda^{-1}\hat{p}_2$ . We take  $q = 1$  and calculate

$$\theta = \frac{\mu(A_n^{(1)})}{\mu(U_n)} = \left(1 - \frac{1}{\lambda} \cdot \frac{|\hat{p}_2 - \lambda^{-1}\hat{p}_1|}{\hat{p}_2 - \hat{p}_1}\right).$$

See Figure 7. The general case where  $\hat{p}_1$  is not parallel  $v^-$  proceeds the same way by considering the expansion of  $T$  orthogonal to the segment  $[\hat{p}_1 - \hat{p}_{per}, \hat{p}_2 - \hat{p}_{per}]$ . We infer

that for general  $q \geq 1$ ,

$$(1 - \frac{1}{\lambda^q}) \leq \theta \leq 1$$

with all values of  $\theta$  in this range being realizable. The verification of condition  $\mathcal{I}'_q(u_n)$  is similar to case (b1).

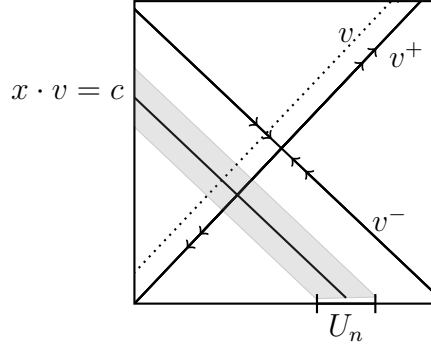


FIGURE 5. The set  $U_n$  and line  $L$  for  $L$  aligned with  $v^+$ .

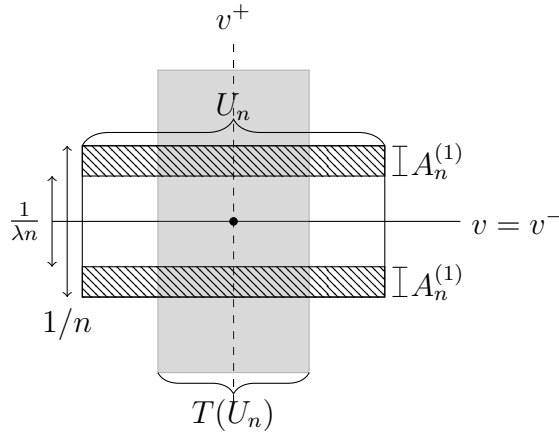


FIGURE 6. Sketch of argument (b1) for  $v$  aligned with the stable direction and  $L$  contains a periodic orbit showing intersections of  $A_n^{(1)}$  (shown in patterned lines) and  $T(U_n)$  (shown in gray). Estimates of the ratio of  $A_n^{(1)}$  to  $U_n$  (shown in white) give the value of the extremal index.

**3.4. Proof of Theorem 2.5.** We will show that conditions  $\mathcal{I}_q(u_n)$  and  $\mathcal{I}'_q(u_n)$  hold with  $q = 0$  so that the extremal index  $\theta = 1$ . We shall drop the subscript  $q$  in this section. The proof of  $\mathcal{I}(u_n)$  follows the same strategy as in the hyperbolic toral automorphism case, the differences necessary in the planar dispersing billiard setting are addressed in [27, Theorem 2.1] and [4, Proposition 5 and Lemma 6]). To simplify the exposition we will consider the case  $L = \{x : r = r_0\}$ .

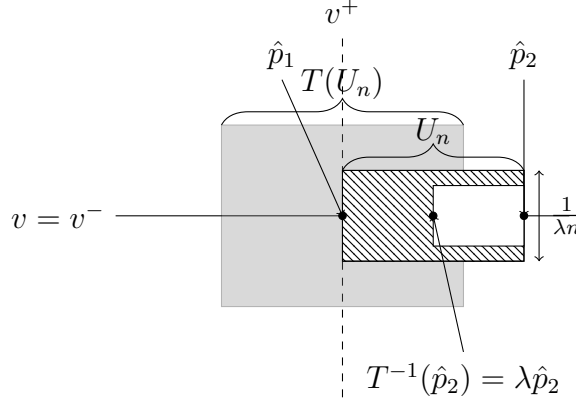


FIGURE 7. Sketch of argument (b2) for  $v$  aligned with the stable direction and  $L$  does *not* contain a periodic orbit. Showing intersections of  $A_n^{(1)}$  (shown in patterned lines) and  $T(U_n)$  (shown in gray). Estimates of the ratio of  $A_n^{(1)}$  to  $U_n$  (shown in white) give the value of the extremal index.

**3.5. Checking condition  $\mathcal{A}'(u_n)$ .** Before checking  $\mathcal{A}'(u_n)$ , we note that as the case of hyperbolic automorphisms of the 2-torus we need only to consider the sum up to time  $(\log n)^{1+\delta}$ , for  $\delta > 0$  because of exponential decay of correlations as in Lemma 3.4. The analogous result to Lemma 3.3. and Lemma 3.4 in the case of Sinai dispersing billiards follows with no essential modification the proofs of Proposition 5 [4] and Lemma 6 [4] respectively (with  $(X_0 > u_n)$  equal to  $A_n^q$ ). Thus the remaining sum

$$n \sum_{j=(\log n)^{1+\delta}}^{\lfloor k_n/n \rfloor} \mu(U_n \cap T^{-j}U_n) \rightarrow 0.$$

(Note here, we work with  $A_n^{(0)} \equiv U_n$ ).

The set  $\{r = r_0\}$  corresponds to a line (call it  $L$ ) which is transverse to the discontinuity set  $S^+$  for  $T$  and the discontinuity  $S^-$  for  $T^{-1}$ . Let  $U_n$  be the rectangle centered at  $L$  with length  $\pi$  and of width roughly  $\frac{\pi}{\pi n}$  corresponding to the set  $\{\phi > u_n\}$  so that  $\mu(U_n) = \frac{\pi}{n}$ .

### Short Returns.

Let  $S_n = \cup_{j=0}^{n-1} T^{-j}S^+$ . The number of smooth connected components of  $S_n$  is bounded above by  $\kappa^n$  for some  $\kappa > 0$ . Let  $C = \frac{1}{4 \log \kappa}$  and then the number of smooth connected components in  $S_{[C \log n]}$  is bounded above by  $n^{1/4}$ . Let  $p_i = (r_0, \vartheta_i) \in L$  be the intersection points  $S_{[C \log n]} \cap L$ , ordered from lowest  $\vartheta$  value to highest and let  $\alpha_i = \vartheta_{i+1} - \vartheta_i$ . Let  $B_1 = \{\alpha_i : \alpha_i < n^{-1/2}\}$ . We estimate  $\sum_{\alpha_i \in B_1} \alpha_i \leq n^{1/4} n^{-1/2} = n^{-1/4}$ . For each  $\alpha_i$  we define the rectangle  $R_i = [r_0 - \frac{1}{n}, r_0 + \frac{1}{n}] \times \alpha_i$  and note that  $\mu(\alpha_i) = O(\frac{\alpha_i}{n})$ . Let  $B = \{R_i : \alpha_i \in B_1\}$ , then  $\mu(B) \leq n^{-1} n^{-1/4} = n^{-5/4}$  and so can be neglected. Let  $G = \{R_i \in B^c\}$ . If  $R_i \in G$  then  $\mu(R_i) \geq n^{-3/2}$  and is of length  $\geq n^{-1/2}$  in the  $\vartheta$  direction



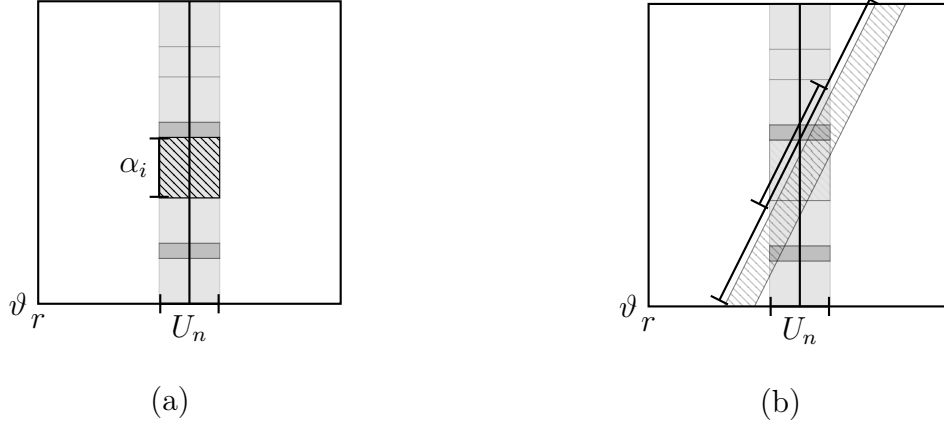


FIGURE 8. (a) Intersection of points with  $r = r_0$  that will not hit a extremal in  $C \log n$  iterates. (b) Expansion of a single rectangle of side-length  $\alpha_i$ . Lines indicate portion that intersects  $U_n$ .

and width  $1/n$  in the  $r$ -direction. If  $R_i \in G$  then  $T^{[C \log n]} R_i$  is a connected ‘rectangle’ which has expanded in the unstable direction, contracted in the stable direction and may wind around the phase space at most once.  $T^{[C \log n]} R_i$  intersects  $U_n$  transversely (since  $L$  is transverse to the unstable cone) in a connected component of measure  $O(n^{-1/2} \mu(R_i))$ . We estimate  $\mu(U_n \cap T^{-j}(U_n)) \leq \mu(R_i \in B) + \sum_{R_i \in B^c} \mu(U_n \cap T^j(R_i)) \leq C n^{-5/4} \mu(U_n)$ . and conclude,

$$\lim_{n \rightarrow \infty} n \sum_{j=1}^{C \log n} \mu(U_n \cap T^{-j}(U_n)) = 0.$$

### Intermediate Returns.

The proof of this section is similar to that in the toral automorphism case but with additional complications due to the presence of discontinuities for  $T$ , causing the unstable manifolds to fragment into small pieces. A scenario which needs to be ruled out is that a large number of small pieces of fragmented unstable manifolds may find themselves again in  $U_n$ . To overcome this we use the following property satisfied by the planar dispersing billiard map:

**One-step expansion.** For  $\alpha \in (0, 1]$ ,

$$\liminf_{\delta \rightarrow 0} \sup_{W: |W| < \delta} \sum_n \left( \frac{|W|}{|V_n|} \right)^\alpha \cdot \frac{|T^{-1}V_n|}{|W|} < 1,$$

where the supremum is taken over regular unstable curves  $W \subset X$ ,  $|W|$  denotes the length of  $W$ , and  $V_n$ ,  $n \geq 1$ , the smooth components of  $T(W)$ ,  $\alpha \in (0, 1]$ . The class of regular curves includes our local unstable manifolds [6].

The expansion by  $DT$  is unbounded and this may lead to different expansion rates at different points on the local unstable manifold  $W^u(x)$ . To overcome this effect and obtain uniform estimates on the densities of conditional SRB measure it is common to define homogeneous local unstable and local stable manifolds. This is the approach adopted in [2, 3, 7, 46]. Fix a large  $k_0$  and define for  $k > k_0$

$$I_k = \{(r, \vartheta) : \frac{\pi}{2} - k^{-2} < \vartheta < \frac{\pi}{2} - (k+1)^{-2}\},$$

$$I_{-k} = \{(r, \vartheta) : -\frac{\pi}{2} + (k+1)^{-2} < \vartheta < -\frac{\pi}{2} + k^{-2}\},$$

and

$$I_{k_0} = \{(r, \vartheta) : -\frac{\pi}{2} + k_0^{-2} < \vartheta < \frac{\pi}{2} - k_0^{-2}\}.$$

In our setting we call a local unstable (stable) manifold  $W^u(x)$ , ( $W^s(x)$ ) homogeneous if for all  $n \geq 0$   $T^n W^u(x)$  ( $T^{-n} W^s(x)$ ) does not intersect any of the line segments in  $\cup_{k > k_0} (I_k \cup I_{-k}) \cup I_{k_0}$ . Homogeneous local unstable manifolds  $W^u(x)$  have almost constant conditional Sinai-Bowen-Ruelle (SRB) densities  $\frac{d\mu_{W^u(x)}}{d\mu_{W^u(y)}}$  in the sense that there exists  $C > 0$  such that  $\frac{1}{C} \leq \frac{d\mu_{W^u(x)}(z_1)}{d\mu_{W^u(x)}(z_2)} \leq C$  for all  $z_1, z_2 \in W^u(x)$  (see [6, Corollary 5.30]). Here  $\frac{d\mu_{W^u(x)}}{d\mu_{W^u(y)}} = \lim_{n \rightarrow \infty} \frac{J_{W^u T^{-n}x}}{J_{W^u T^{-n}y}}$  where  $J_{W^u T^{-n}}$  is the restriction of the map  $T^n$  to the one-dimensional curve  $W^u$  [6, Theorem 5.2]. The conditional density, apriori only defined up to a constant, is then determined uniquely by the requirement that  $\int_{W^u} d\mu_{W^u}(y) dy = 1$ . These conditional densities are the densities of the conditional measures induced on the measurable partition into local unstable manifolds by the invariant measure  $\mu$ . Conditional measures are important in the study of billiard systems as they allow us to go from knowledge of behavior of quantities on unstable manifolds to the whole phase space by Fubini type arguments. For more details of the construction of the conditional measures on local unstable manifolds and their densities we refer the reader to Section 5.2 of [6].

From this point on all the local unstable (stable) manifolds that we consider will be homogeneous. We may as well suppose all such curves are contained in a set  $R_i \in G$  as  $\mu(B) < n^{-5/4}$ .

We now take care of the times  $[C \log n] < j < (\log n)^{1+\delta}$ . If  $W^u(x) \cap U_n \subset R_i \in G$  then  $T^{[C \log n]}$  has expanded  $W^u(x)$  by a factor  $\Lambda^{C \log n} = n^{C \log \Lambda} = n^\beta$  for some  $\beta > 0$  and the iterates of the components of  $W^u(x) \cap U_n$  have not hit an extremal set in the first  $[C \log n]$  iterates. Let  $\gamma_n(x) = W^u(x) \cap U_n$ . By [7, Theorem 5.7]  $\mu(W^u(x) < n^{-1-\beta/2}) < n^{-1-\beta/2}$  so we may require all  $W^u(x) \in \cup_{R_i \in G} R_i$  to satisfy  $|\gamma_n(x)| > n^{-1-\beta/2}$ .

Now we consider  $\mu(U_n \cap T^{-j}(U_n))$  for  $C \log n \leq j \leq (\log n)^{1+\delta}$ . Note that  $T^j(\gamma_n(x))$  consists of a connected curve for  $j \leq C \log n$ . Recall by expansion under the map we have  $|T^j \gamma_n(x)| \geq n^\beta |\gamma_n(x)| > n^{-1+\beta/2}$ . If we iterate this component further such that  $T^{i+j} \gamma_n(x)$ ,  $i > 0$  intersects an extremal line then we may decompose  $T^{i+j} \gamma_n(x)$  into smooth connected components  $V_n$  and their preimages  $Y_n \subset T^j \gamma_n(x)$  so that  $T^i$  maps  $Y_n$  onto  $V_n$  diffeomorphically and with uniformly bounded distortion. Applying one-step expansion to

$T^j\gamma_n(x)$  gives,

$$\sum_n \left( \frac{|T^j\gamma_n(x)|}{|V_n|} \right)^\alpha \left| \frac{Y_n}{T^j\gamma_n(x)} \right| < 1.$$

Fix  $T^j\gamma_n(x)$  and for every point  $p \in T^j\gamma_n(x)$  let  $Y_n(p)$  denote the unique  $Y_n \subset T^j\gamma_n(x)$  such that  $p \in Y_n$  and let  $V_n(p)$  denote  $T^i Y_n(p)$ . Let  $\nu(p) = \frac{|Y_n(p)|}{|T^j\gamma_n(x)|}$  be the probability mass function of a probability measure  $\nu$  on  $T^j\gamma_n(x)$  where  $\|C\|$  is the length of a rectifiable one dimensional curve. Let  $f(p) = \left( \frac{|T^j\gamma_n(x)|}{|V_n(p)|} \right)^\alpha$  be a function on this probability space. By one-step expansion  $\nu(f) \leq 1$ . Now  $\{p \in T^j\gamma_n(x) : |V_n(p)| < n^{-1+\varepsilon\beta/2}\} \subset \{p \in T^j\gamma_n(x) : f(p) > n^{(1-\varepsilon)\beta/2\alpha}\}$ . By Markov's inequality  $\nu\{p \in T^j\gamma_n(x) : |V_n(p)| < n^{-(1+\varepsilon)\beta/2}\} \leq n^{-(1-\varepsilon)\beta/2\alpha}$ .

We choose  $\varepsilon$  sufficiently small so that  $\rho_1 := 1 - (1 - \varepsilon)\beta/2\alpha > 0$  (since  $\beta < 1$ ) and define  $\rho = \min\{\rho_1, \beta/2\alpha\varepsilon\}$ . With our choice of  $\varepsilon$ , if  $|V_n| \geq n^{-1+\varepsilon\beta/2}$  then,

$$\frac{|V_n \cap U_n|}{|V_n|} \leq C_1 n^{-\rho}.$$

By bounded distortion of the map  $T^{i+j}$ , after throwing away the  $V_n$  such that  $|V_n| \leq n^{-1+\varepsilon\beta/2}$  we have

$$\frac{|\gamma_n(x) \cap T^{-i-j}(U_n)|}{|\gamma_n(x)|} \leq C_2 n^{-\rho}.$$

This provides a bound on the length  $|\gamma_n(x) \cap T^{-r}(U_n)|$  for  $c \log n \leq r \leq (\log n)^{1+\delta}$ . We may now use the fact that  $\mu$  decomposes as a product measure of conditional SRB measures on  $\gamma_n(x) \cap U_n$  to estimate

$$\mu(U_n \cap T^{-j}(U_n)) \leq C_4 n^{-1-\rho}.$$

Putting these results together implies,

$$\lim_{n \rightarrow \infty} n \sum_{r=C \log n}^{(\log n)^{1+\delta}} \mu(U_n \cap T^{-r}(U_n)) = 0.$$

Condition  $\mathcal{A}'(u_n)$  follows.

*Remark 3.5.* Using essentially the same analysis it is standard to show that the return time statistics to  $L = \{(r, \vartheta) : r = r_0\}$  is standard simple Poisson. To see this we need verify condition  $D_q^*(u_n)$  of [4, Section 2], but the proof of this is a minor modification of  $\mathcal{A}(u_n)$ . In contrast suppose  $(r_0, \vartheta_0)$  is a periodic point of period  $q$ , then we would obtain a compound Poisson process as given in [4, Theorem 2].

**3.6. Proof of Theorem 2.8 and Theorem 2.9.** We give the proof in detail only for the case of two coupled maps, as the proofs in the other cases are the same with obvious modifications. The uniform expansion away from the invariant subspace plays the same role in each setting. Note that the subspace  $L$  of Theorem 2.9 is invariant, and we will show that there is uniform expansion in the directions orthogonal to  $L$ .

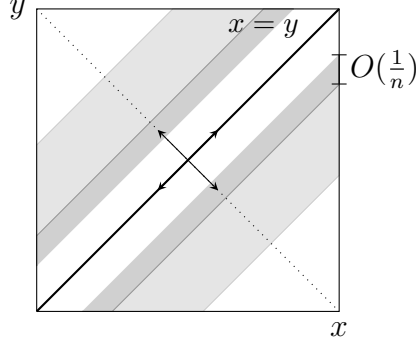


FIGURE 9. Expansion of the set  $A_n$  (given in gray) under the map  $F$ . The thick, black line represents the line of maximization. Arrows indicate uniform expansion under  $F$  in all directions.

Recall  $\phi(x, y) = -\log|x - y|$ , a function maximized on the line segment or circle  $L = \{(x, y) : y = x\}$ . For  $\tau > 0$  define  $u_n(\tau)$  by  $n\mu(\phi > u_n(\tau)) = \tau$ , and  $U_n = \{\phi > u_n(\tau)\}$ . Define  $A_n = \{\phi > u_n, \phi \circ F < u_n\}$  and recall for a set  $B$ ,  $\mathcal{W}_{s,l}(B) = \bigcap_{i=s}^{s+l-1} F^{-i}(B)$ .

Note that the invariant line  $L$  is uniformly repelling in the orthogonal direction  $(1, -1)$  since writing  $y - x = \epsilon$  we see  $\epsilon \rightarrow (1 - \gamma)[T(x + \epsilon) - Tx] \sim (1 - \gamma)DT(x)\epsilon + O(\epsilon^2)$  under the map  $F$ .

Furthermore  $A_n$  is a union of two rectangles and  $A_n \cap F^{-2}A_n = \emptyset$  as a result of uniform expansion away from the invariant line  $L$ .

Condition  $\mathcal{D}(u_n)$  follows easily by an approximation argument using exponential decay of correlations of Lipschitz versus  $L^\infty$  functions taking  $t_n = (\log n)^5$  say.

Now we prove condition  $\mathcal{D}'_q(u_n)$  (for  $q = 1$ ), namely

$$\lim_{n \rightarrow \infty} n \sum_{j=1}^{\lfloor n/k_n \rfloor} \mu(A_n \cap F^{-j}(A_n)) = 0.$$

Note that by uniform repulsion from the invariant line  $L$  there exists  $C_4$  such that for  $j = 1, \dots, C_4 \log n$ ,  $\mu(A_n \cap F^{-j}A_n) = 0$ . This follows since  $F^{-1}A_n \cap A_n = \emptyset$  (by definition) and uniform repulsion from the invariant line ensures also  $F^{-j}A_n \cap A_n = \emptyset$  for a certain number of iterates  $j = 1, \dots, C_4 \log n$  until for all  $(x, y)$  in  $A_n$ ,  $|F^j(x, y)| = O(1)$  (i.e. until the expansion in the  $L^\perp$  direction is  $O(n)$ ).

As  $DF$  is bounded and uniformly expanding, in all directions  $A_n$  has been expanded by the map  $F^{[C_4 \log n]}$  by at least  $n^\alpha$  for some  $0 < \alpha < 1$ . To see this, note that for any expanding map expansion of  $A_n$  by the map  $F^{[C_4 \log n]}$  given by at least  $C_5 |DT|_{\min}^{C_4 \log n} \sim n^\alpha$ .

Choose  $C_3 \geq C_4$  large enough that  $\mu(A_n \cap F^{-j}(A_n)) \leq \frac{1}{n^{3/2}}$ , this is possible by exponential decay of correlations and a Lipschitz approximation to  $1_{A_n}$ .

Thus for  $C_4 \log n \leq j \leq C_3 \log n$ ,  $\mu(A_n \cap F^{-j}(A_n)) \leq \frac{1}{n^{1+\alpha}}$ . For  $1 \leq j \leq C_4 \log n$ ,  $\mu(F^{-j}(A_n) \cap A_n) = 0$  for  $C_4 \log n \leq j \leq C_3 \log n$ ,  $\mu(A_n \cap F^{-j}A_n) \leq \frac{1}{n^{1+\alpha}}$  and for  $j \geq C_4 \log n$ ,  $\mu(A_n \cap F^{-j}(A_n)) \leq \frac{1}{n^{3/2}}$ .

This implies  $\mathcal{I}'_q(u_n)$  for  $q = 1$  (corresponding to the fact that  $L$  is fixed).

Finally we compute the extremal index, changing coordinates to  $v = \frac{x-y}{\sqrt{2}}$ ,  $u = \frac{x+y}{\sqrt{2}}$  we have

$$\theta = \lim_{n \rightarrow \infty} \theta_n = \lim_{n \rightarrow \infty} \frac{\mu(A_n)}{\mu(U_n)}.$$

However

$$\lim_{n \rightarrow \infty} \frac{\mu(A_n)}{\mu(U_n)} = \lim_{n \rightarrow \infty} [1 - \int_0^{\frac{1}{n[Tv]}} \int_0^{\frac{1}{n}} \tilde{h}(u, v) du dv] / \int_0^{\frac{1}{n}} \int_0^{\frac{1}{n}} \tilde{h}(u, v) du dv.$$

Suppose  $m(U_n) = O(\epsilon^{1/4})$ ,  $\epsilon < \epsilon_0$ . Since  $|\tilde{h}|_\alpha < \infty$ ,  $m(x \in U_n : \text{osc}(\tilde{h}, B_\epsilon(x)) > \sqrt{\epsilon}) < \sqrt{\epsilon} = O(m(U_n)^2)$ . We may assume that  $\tilde{h}$  is regularized along the diagonal in the sense that for Lebesgue almost every  $u$ ,  $\tilde{h}(u, u)$  is the average of the limits of  $\tilde{h}(u, v)$  and  $\tilde{h}(u, -v)$  as  $v \rightarrow 0$ . Thus, as expansion along  $v$  at  $v = 0$  is given by  $(1 - \gamma)DT(u)$ , and  $\tilde{h}$  is essentially bounded

$$\theta = 1 - \int_L \frac{\tilde{h}(u, u)}{(1 - \gamma)|DT(u)|} du.$$

*Remark 3.6.* Our techniques allow us to obtain similar results to that of [13] in a simpler setting through a pure probabilistic approach and extend these results to blocks of synchronization discussion in [13, Section 7.2].

**3.7. Numerical Results for the Extremal Index.** In this section we provide numerical estimates for the extremal index to support the theoretical results for the coupled uniformly expanding map and the hyperbolic automorphism of the 2-torus given in Theorems 2.8 and 2.9 and Theorem 2.1, respectively. We begin by verifying that the numerical estimates we obtain from the coupled systems agree with that of [13]. Then, we extend these results to include estimates for the extremal index over blocks of synchronization where each block introduces a new invariant direction and changes the value of the extremal index. We end with a numerical investigation for Arnold's cat map where the alignment of the line  $L$  and the existence of periodic orbits along  $L$  or its continuation determine the value of the extremal index.

**3.8. Coupled systems of uniformly expanding maps.** Numerical barriers in computing trajectories in piecewise uniformly expanding maps are given by the fact that

- (i) The periodic orbits are dense making long trajectories not easily computable.
- (ii) Round off errors may produce unreliable results.

To overcome (i) we employ a numerical technique adapted from [39] to prevent trapping of the orbit near the fixed point by adding a small  $\varepsilon = O(10^{-2})$  perturbation to the trajectory.

Arguments for this technique are typically given in the form of a shadowing lemma which states the existence of a true orbit that is  $\epsilon$ -close to the computed orbit. We first note that [14] proves the existence of an EVL for randomly perturbed piecewise expanding maps provided this perturbation  $\varepsilon > 10^{-4}$ . Further, [13] provides evidence that the extremal index is qualitatively robust under small  $\varepsilon = 10^{-2}$  additive noise. To overcome (ii), in light of our long trajectories ( $t = 10^6$ ), we refer to [15] where the round off error resulting from double precision computation was shown to be equivalent to the addition of random noise of order  $10^{-7}$ .

*Estimating the EI for the coupled map system over the whole extremal set.* We estimate the extremal index in a similar way to that of [13] for  $\phi(\bar{x}) = -\log(\|p^\perp\|)$  using the formula provided by Süveges [44]. The code for this estimate can be found in [39]. From Theorem 2.8 we expect,

$$\theta = 1 - \frac{1}{(1 - \gamma)^{m-1}} \frac{1}{|DT|^{m-1}}.$$

We compute the extremal index for fixed  $m = 2$  and varying values of  $\gamma$ , and varying values of both  $\gamma$  and  $m$ . Our results coincide with that of [13]; higher values of  $m$  and lower values of  $\gamma$  produce an extremal index near 1. Low values of  $\gamma$  give higher weights to the non-coupled components of the map resulting in a system which behaves more independently. Lower values of  $m$  result in a more dependent system since the coupled term is more affected by changes while larger values of  $m$  result in a coupled term which is averaged over a larger number of maps and less affected by individual changes. For results see Figure 10.

*Estimating the EI for the coupled map system over blocks of synchronization.* We provide numerical estimates of the extremal index in a more specific setting of block synchronization where  $L = \{(x_1, x_2, \dots, x_m) : x_1 = x_2 = \dots = x_m\}$  and  $L = \{(x_1, x_2, \dots, x_m) : x_1 = x_2 = \dots = x_{m-1}, x_m\}$ . From Theorem 2.9 we expect,

$$\theta = 1 - \frac{1}{(1 - \gamma)^{m-1}} \frac{1}{|DT|^{m-1}}.$$

for  $L = \{(x_1, x_2, \dots, x_m) : x_1 = x_2 = \dots = x_m\}$  and,

$$\theta = 1 - \frac{1}{(1 - \gamma)^{m-2}} \frac{1}{|DT|^{m-2}}$$

for  $L = \{(x_1, x_2, \dots, x_m) : x_1 = x_2 = \dots = x_{m-1}, x_m\}$ . Defining  $L$  in this way reduces the spatial dimension in which expansion away from  $L$  can occur. This results in a extremal index equivalent to that of an  $m - 1$  coupled system. We give results in the case when  $m = 5$  (see Figure 11) .

We also consider blocks of successive indices in the general setting of block synchronization so that  $L$  can be defined as any combination of block sequences. From Theorem 2.9 we expect the value of the extremal index to be determined by the spatial dimension of expansion for the system. In the following numerical examples we consider  $m = 5$  and

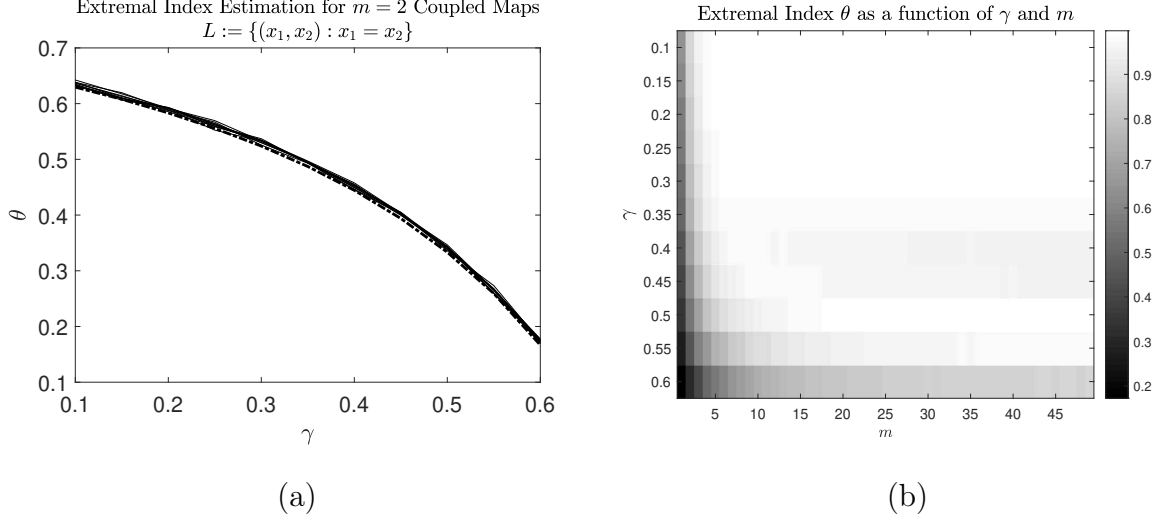


FIGURE 10. Extremal index  $\theta$  estimation for the  $m$ -coupled map  $F$  with  $\varphi(x) = -\log(\|p^\perp\|)$  where the set of maximization  $L = \{(x_1, x_2, \dots, x_m) : x_1 = x_2 = \dots = x_m\}$  for (a) fixed  $m$  and varying  $\gamma$  (10 different realizations  $t = 10^6$ ) and (b) varying  $m$  and  $\gamma$ . (a) The marked line indicates the theoretical value of  $\theta$  given. (b) The colorbar indicates the value of  $0 \leq \theta \leq 1$  for each  $m$  and  $\gamma$  pair.

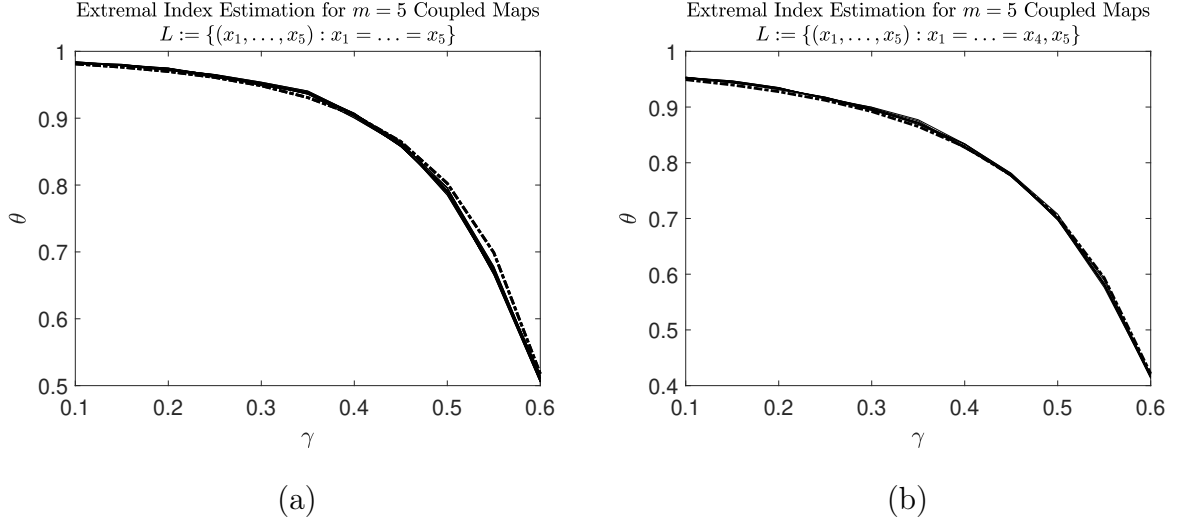


FIGURE 11. Extremal index  $\theta$  estimation (10 different realizations,  $t = 10^6$ ) for the  $m$ -coupled map  $F$  with  $\phi(x) = -\log(\|p^\perp\|)$  where (a)  $L$  is the line  $x_1 = x_2 = \dots = x_m$  and (b)  $L$  is the plane  $x_1 = x_2 = \dots = x_{m-1}, x_m$ . The marked line indicates the theoretical value of  $\theta$  given.

note that the extremal index for that of  $L = \{(x_1, \dots, x_5) : x_1 = x_2 = x_3 = x_4, x_5\}$  (Figure 11 (b)) is equivalent to that of  $L = \{(x_1, \dots, x_5) : x_1 = x_2 = x_3, x_4 = x_5\}$  (Figure 12 (b)). This is expected since they share the same number of non-invariant directions of expansion.

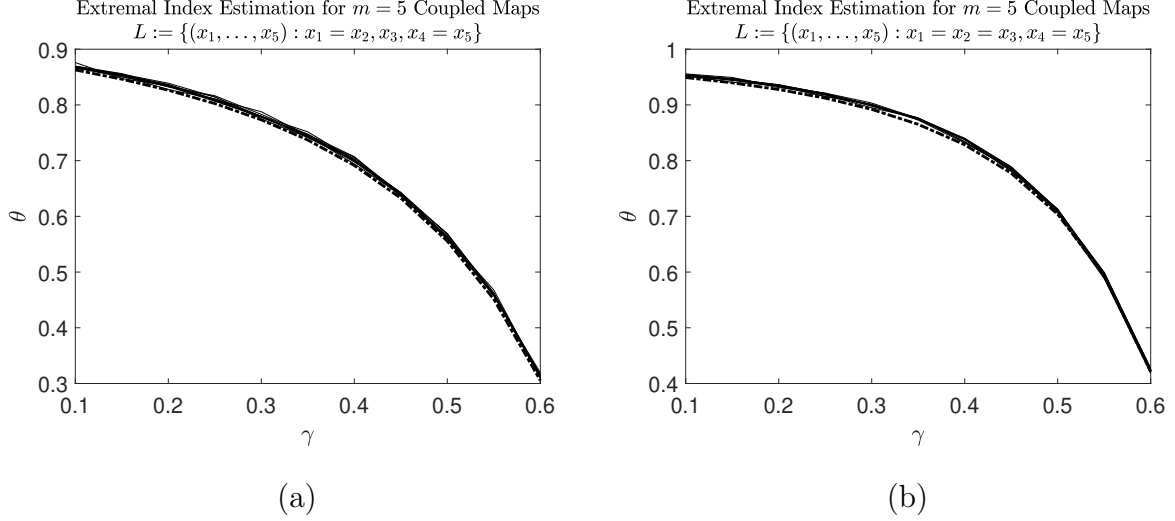


FIGURE 12. Extremal index  $\theta$  estimation (10 different realizations,  $t = 10^6$ ) for the  $m$ -coupled map  $F$  with  $\phi(x) = -\log(\|p^\perp\|)$  where (a)  $L$  is the set of two planes  $x_1 = x_2$  and  $x_4 = x_5$  so that  $\theta = 1 - \frac{1}{(1-\gamma)^2|DT|^2}$  (b)  $L$  is the set of planes  $x_1 = x_2 = x_3$  and  $x_4 = x_5$ ,  $\theta = 1 - \frac{1}{(1-\gamma)^3|DT|^3}$ . The marked line indicates the theoretical value of  $\theta$  given.

**3.9. Hyperbolic toral automorphisms.** We compute trajectories for increasing time intervals of Arnold's cat map given by,

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mod 1.$$

The uniformly hyperbolic structure of this map allows us to calculate long trajectories without the risk of points being trapped in a few time steps. The stability of this map ensures that the qualitative behavior is unaffected by small perturbations. We use this to argue the accuracy of the calculated orbit up to  $t = 10^4$  under double precision.

From Theorem 2.1 we expect the value of the extremal index  $\theta$  to depend on both the alignment of  $v$  in the observable  $\phi(x) = -\log d(x, L)$ , with  $d$  the usual Euclidean metric, and the existence of a periodic orbit along  $L$ . Figure 13 (a) shows the extremal index estimation given by [44] for 10 different initial values where  $v$  aligns with the unstable direction and contains a 2-periodic point. Hence,  $\theta = 1 - \frac{1}{\lambda^2}$ . Figure 13(b) shows the extremal index estimation for 10 different initial values where  $v$  is not aligned with the stable or unstable direction. In this setting we expect  $\theta = 1$ . The variation from the expected value for each realization is at most  $O(10^{-2})$ .



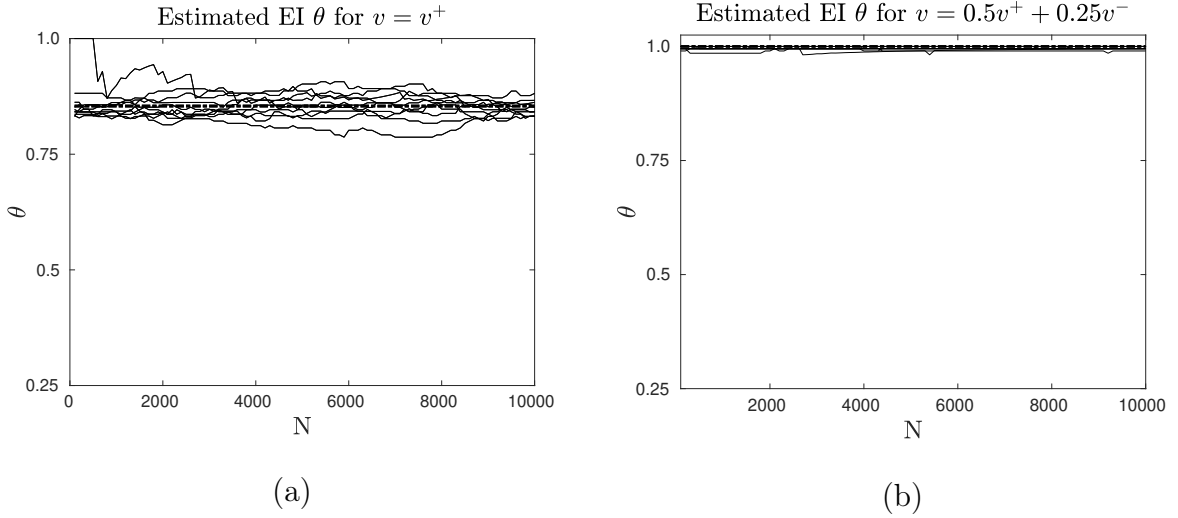


FIGURE 13. Extremal index  $\theta$  estimation (10 different realizations,  $t = 10^4$ ) for Arnold's cat map with  $\varphi(x) = -\log d(x, L)$  where (a)  $L = v^+$  and (b)  $L = 0.5v^+ + 0.25v^-$ . The marked line indicates the theoretical value of  $\theta$  given.

#### 4. DISCUSSION: TOWARDS MORE GENERAL OBSERVABLES AND NON-UNIFORMLY HYPERBOLIC SYSTEMS

In this article we have focused on hyperbolic systems with invariant measures absolutely continuous with respect to Lebesgue measure and considered observables whose level sets  $\mathcal{S}_\epsilon$  shrink to a non-trivial extremal set  $\mathcal{S}$ , such as a line segment. We recall that  $\mathcal{S}_\epsilon = \{x \in X : d_H(x, \mathcal{S}) \leq \epsilon\}$ , and  $d_H(x, \mathcal{S})$  is the Hausdorff distance from  $x$  to  $\mathcal{S}$ . Thus if  $\mathcal{S}$  is a smooth curve, then for this metric  $d_H$  we see that  $\mathcal{S}_\epsilon$  is a thin tube of width  $\epsilon$  around  $\mathcal{S}$ . The observable  $\phi : X \rightarrow \mathbb{R}$  we have assumed to be given by  $\phi(x) = f(d_H(x, \mathcal{S}))$ , for some smooth function  $f : [0, \infty) \rightarrow \mathbb{R}$ , maximised at 0, e.g.  $f(u) = -\log u$ .

As explained in Section 2, our methods apply to the case where  $\mathcal{S}$  is a line segment. We conjecture that our methods extend to the case where  $\mathcal{S}$  is a smooth curve, assuming some transversality conditions of  $\mathcal{S}$  relative to the global stable/unstable manifolds of the system. We have also considered seemingly non-generic geometrical cases, e.g. where the line segment  $\mathcal{S}$  aligns precisely with the global stable/unstable manifolds. For hyperbolic automorphisms of the 2-torus, we established the limit laws that arise in these scenarios. More generally, it is natural to consider observables whose extremal set  $\mathcal{S}$  is no longer (strictly) transverse to the global stable/unstable manifolds, i.e. there exist points of tangency between  $\mathcal{S}$  and the global manifolds.

For (non-uniformly) hyperbolic systems  $(f, \Lambda, \mu)$  where  $\Lambda$  is an attractor the Sinai-Ruelle-Bowen (SRB) measure  $\mu$  may not be equivalent to Lebesgue. These systems include Axiom A systems, or Hénon-like attractors whose statistical properties (such as mixing rates) are

established in [46]. As outlined in Section 1.1, there is an established literature on extreme value theory in the non-uniformly hyperbolic setting for observables whose extremal set  $\mathcal{S}$  is a point. Recently some progress has been made on more complicated geometries for  $\mathcal{S}$  [30] but in a very axiomatic way. In the case where  $\mathcal{S}$  is a line (or in higher dimensions a planar set), then we expect  $\mathcal{S}$  to (generically) intersect a fractal attractor  $\Lambda$  in a Cantor-like set. For such a set, there are various difficulties that arise when trying to find the limit extreme value distribution, in the sense of establishing (1.2), or in particular the limit law given by (1.4). If we suspect that a limit law of the form given in equation (1.2) is going to exist, then finding the scaling sequence  $u_n$  is a first problem. For a specified observable  $\phi$  (i.e. through specifying  $f$ ), the properties of the sequence  $u_n$  depend on the asymptotic properties of  $\mu(\mathcal{S}_\epsilon)$  as  $\epsilon \rightarrow 0$ . To estimate this measure, we cannot use local dimension estimates, and finer arguments are required based on the geometric properties of  $\mu$ . Furthermore, existence of a GEV limit of the form (1.4) is not guaranteed, as this requires  $\mu(\mathcal{S}_\epsilon)$  to satisfy conditions of regular variation in  $\epsilon$  (as  $\epsilon \rightarrow 0$ ), see [39, Chapter 3]. Axiomatic approaches, e.g. [5, 30, 34] suggest that once we've found these scaling laws then an extreme value law holds in the sense of equation (1.2). However, verification of these axioms still requires fine analysis. This includes verification of axiomatic conditions involving transversality of  $\mathcal{S}$  with  $\Lambda$ , and conditions involving how  $\mu$  behaves on certain shrinking sets (such as thin annuli) on a case-by-case basis.

## REFERENCES

- [1] R. Bowen. Equilibrium states and the ergodic theory of Anosov diffeomorphisms. Springer-Verlag, Berlin. Lecture Notes in Mathematics, Vol. 470. (1975)
- [2] L.A. Bunimovich, Ya. Sinai, and N. Chernov. *Markov partitions for two-dimensional billiards*. Russ. Math. Surv. **45**, (1990), 105–152.
- [3] L.A. Bunimovich, Ya. Sinai, and N. Chernov, *Statistical properties of two-dimensional hyperbolic billiards*. Russ. Math. Surv. **46**, (1991), 47–106.
- [4] M. Carney, M. Nicol and H. K. Zhang. *Compound Poisson law for hitting times to periodic orbits in two-dimensional hyperbolic systems*. J. Stat. Phys. 169, (4), (2017), 804–823.
- [5] J.-R. Chazottes and P. Collet. *Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems*. Ergodic Theory and Dynamical Systems. **33**, (2013), 49–80.
- [6] N. Chernov and R. Markarian. *Chaotic billiards*. Math. Surv. Monographs, **127**, AMS, Providence, RI, (2006), 316 pp.
- [7] N. Chernov and R. Markarian. *Dispersing billiards with cusps: slow decay of correlations*. Commun. Math. Phys., **270**, (2007), 727–758.
- [8] P. Collet. *Statistics of closest return for some non-uniformly hyperbolic systems*. Ergodic Theory and Dynamical Systems. **21**, (2001), 401–420.
- [9] S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer, 2004.
- [10] M. Carvalho, A. C. M. Freitas, J. M. Freitas, M. Holland and M. Nicol. *Extremal dichotomy for hyperbolic toral automorphisms*. Dynamical Systems: An International Journal, **30**, (4), (2015), 383–403.
- [11] M. Denker, M. Gordin, and A. Sharova. *A Poisson limit theorem for toral automorphisms*. Illinois J. Math, **48**, (1), (2004), 1–20.
- [12] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events for insurance and finance*. Applications of Mathematics (New York), **33**, Springer-Verlag, Berlin, 1997.
- [13] D. Faranda, H. Ghoudi, P. Guirard and S. Vaienti. *Extreme value theory for synchronization of coupled map lattices*, Nonlinearity, **31**, (7), (2018), 3326–3358
- [14] D. Faranda, J. Freitas, V. Lucarini, G. Turchetti and S. Vaienti. *Extreme value statistics for dynamical systems with noise*. Nonlinearity, **26**, (2013), 2597.
- [15] D. Faranda, M. Mestre, and G. Turchetti. *Analysis of round off errors with reversibility test as a dynamical indicator*. International Journal of Bifurcation and Chaos, **22**, (9), (2012), 1250215.
- [16] A. Ferguson and M. Pollicott. *Escape Rates for Gibbs measures*. Ergodic Theory and Dynamical Systems, **32**, (3), (2012), 961–988.
- [17] 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.
- [18] A. Freitas, J. Freitas, F. Rodrigues, and J. Soares. Rare events for Cantor target sets, Preprint 2019, arXiv:1903.07200.
- [19] J. Freitas, A. Freitas and M. Todd. *Hitting Times and Extreme Value Theory*. Probab. Theory Related Fields, **147**, (3), (2010). 675–710.
- [20] A. Freitas, J. Freitas, and M. Todd. Extremal Index. *Hitting Time Statistics and periodicity*. Adv. Math., **231**, (5), (2012), 2626–2665.
- [21] A. Freitas, F. Freitas and M. Todd. *Extreme value laws in dynamical systems for non-smooth observations*. J. Stat. Phys. **142**, (1), (2011), 108–126.
- [22] A. C. M. Freitas, J. M. Freitas, and M. Todd. Speed of convergence for laws of rare events and escape rates. Stochastic Process. Appl., **125**, (4), (2015), 1653–1687.
- [23] A. C. M. Freitas, J. M. Freitas, and M. Todd. *The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics*. Comm. Math. Phys., **321**, (2), (2013), 483–527.
- [24] J. Freitas, N. Haydn and M. Nicol. *Convergence of rare events point processes to the Poisson for billiards*. Nonlinearity, **27**, (2014), 1669–1687.

- [25] J. Galambos. *The Asymptotic Theory of Extreme Order Statistics*. John Wiley and Sons, 1978.
- [26] C. Gupta. *Extreme-value distributions for some classes of non-uniformly partially hyperbolic dynamical systems*. Ergodic Theory and Dynamical Systems. **30**, (3), (2011), 757–771.
- [27] C. Gupta, M. Holland and M. Nicol. *Extreme value theory and return time statistics for dispersing billiard maps and flows, Lozi maps and Lorenz-like maps*. Ergodic Theory and Dynamical Systems. **31**, (2011), 1363–1390.
- [28] C. Gupta, M. Nicol and W. Ott. *A Borel-Cantelli lemma for non-uniformly expanding dynamical systems*. Nonlinearity **23**, (8), (2010), 1991–2008.
- [29] N. Haydn, M. Nicol, T. Persson and S. Vaienti. *A note on Borel-Cantelli lemmas for non-uniformly hyperbolic dynamical systems*. Ergodic Theory and Dynamical Systems **33**, (2), (2013), 475–498.
- [30] N. Haydn and S. Vaienti. *Limiting entry times distribution for arbitrary null sets*, arXiv:1904.08733v1, April 18 2019.
- [31] M. Hirata. *Poisson Limit Law for Axiom A diffeomorphisms*. Ergodic Theory and Dynamical Systems. **13**, (3), (1993), 533–556.
- [32] M. Holland, M. Nicol, A. Török. *Extreme value theory for non-uniformly expanding dynamical systems*. Trans. Am. Math. Soc., **364**, (2012), 661–688.
- [33] M. Holland, M. Nicol, A. Török. *Almost sure convergence of maxima for chaotic dynamical systems*. Stochastic Process. Appl. **126**, (10), (2016), 3145–3170.
- [34] M. P. Holland, P. Rabassa, A. E. Sterk. *Quantitative recurrence statistics and convergence to an extreme value distribution for non-uniformly hyperbolic dynamical systems*. Nonlinearity, **29**, (8), (2016).
- [35] M. Holland, R. Vitolo, P. Rabassa, A. E. Sterk, H. Broer. *Extreme value laws in dynamical systems under physical observables*. Phys. D, **241**, (2012), 497–513.
- [36] G. Keller. *Rare events, exponential hitting times and extremal indices via spectral perturbation*. Dynamical Systems: An International Journal. **27**, (1), (2012), 11–27.
- [37] D. Kim. *The dynamical Borel-Cantelli lemma for interval maps*. Discrete Contin. Dyn. Syst. **17**, (4), (2007), 891–900.
- [38] Leadbetter, M. R., Lindgren, G., and Rootzen, H. *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York, 1983.
- [39] V. Lucarini, D. Faranda, A.C. Freitas, J.M. Freitas, M. P. Holland, T. Kuna, M. Nicol, M. Todd, S. Vaienti, *Extremes and Recurrence in Dynamical Systems*, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs, and Tracts, 2016.
- [40] V. Lucarini, D. Faranda, J. Wouters J, and T. Kuna. *Towards a General Theory of Extremes for Observables of Chaotic Dynamical Systems*. Journal of Statistical Physics, **154**, (3), (2014), 723–750.
- [41] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, Soc. Industrial and Applied Mathematics, 1992.
- [42] F. Pène and B. Saussol. *Back to balls in billiards*. Comm. Math. Phys. **293**, (3), (2010), 837–866.
- [43] A. E. Sterk, M. P. Holland, P. Rabassa, H. W. Broer, R. Vitolo. *Predictability of extreme values in geophysical models*. Nonlinear Processes in Geophysics, **19**, (2012), 529–539.
- [44] M. Süveges, *Likelihood estimation of the extremal index*. Extremes, Springer, **10**, (1-2), (2007), 41–55.
- [45] Fan Yang. *Rare event processes and entry times distribution for arbitrary null sets on compact manifolds*, arxiv:1905.09956.
- [46] L.-S. Young. *Statistical properties of dynamical systems with some hyperbolicity*. Ann. Math. **147**, (1998), 585–650.

MEAGAN CARNEY, MAX PLANCK INSTITUTE FOR PHYSICS OF COMPLEX SYSTEMS, NONLINEAR DYNAMICS AND TIME SERIES ANALYSIS, DRESDEN, GERMANY

*E-mail address:* `meagan@pks.mpg.de`

MARK HOLLAND, DEPARTMENT OF MATHEMATICS (CEMPS), HARRISON BUILDING (327), NORTH PARK ROAD, EXETER, EX4 4QF, UK

*E-mail address:* `M.P.Holland@exeter.ac.uk`

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

MATTHEW NICOL, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HOUSTON, HOUSTON, TX 77204, USA

*E-mail address:* `nicol@math.uh.edu`

*URL:* `http://www.math.uh.edu/~nicol/`