Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2017, Vol. 53, No. 3, 1341–1370
DOI: 10.1214/16-AIHP757

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Extreme Value Laws for non stationary processes generated by sequential and random dynamical systems

Ana Cristina Moreira Freitas[a], Jorge Milhazes Freitas[b] and Sandro Vaienti[c]

[a]*Centro de Matemática & Faculdade de Economia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal. E-mail: amoreira@fep.up.pt; url: http://www.fep.up.pt/docentes/amoreira/*
[b]*Centro de Matemática & Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal. E-mail: jmfreita@fc.up.pt; url: http://www.fc.up.pt/pessoas/jmfreita/*
[c]*Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France and Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France. E-mail: vaienti@cpt.univ-mrs.fr; url: http://www.cpt.univ-mrs.fr/~vaienti/*

Received 1 December 2015; revised 4 April 2016; accepted 6 April 2016

Abstract. We develop and generalise the theory of extreme value for non-stationary stochastic processes, mostly by weakening the uniform mixing condition that was previously used in this setting. We apply our results to non-autonomous dynamical systems, in particular to *sequential dynamical systems*, given by uniformly expanding maps, and to a few classes of *random dynamical systems*. Some examples are presented and worked out in detail.

Résumé. Nous développons et généralisons la théorie des valeurs extrêmes pour des processus stochastiques non-stationnaires, en affaiblissant la condition de mélange uniforme qui avait été utilisée auparavant. Nous appliquons nos résultats à des systèmes dynamiques non autonomes, en particulier aux systèmes dynamiques séquentiels engendrés par des applications dilatantes et à une large classe de systèmes dynamiques aléatoires. Quelques exemples sont présentés et calculés en détail.

MSC: 37A50; 60G70; 37B20; 37A25

Keywords: Non-stationarity; Extreme value theory; Hitting Times; Sequential dynamical systems; Random dynamical systems

1. Introduction

1.1. *The motivation and the dynamical setting*

One of the most successful directions of ergodic theory in the last decades was the application of probabilistic tools to characterise the asymptotic evolution of a given dynamical system. There is now a well established domain known as *statistical properties of dynamical systems*, which attempts to prove limit theorems under different degrees of mixing. Mixing is the way to restore asymptotic independence and, in this way, mimic independent and identically distributed (i.i.d.) sequences of random variables. A common distribution for the time series arising from the dynamical systems is acquired from the existence of an invariant measure for such systems. In some sense, the existence of such a measure is what defines a dynamical system. Relaxing this assumption gives rise to non-autonomous dynamical systems for which the study of limit theorems is just at the beginning. In this paper, we will focus on one of those statistical properties, namely on asymptotic extreme value distribution laws. Our first goal will be to improve and generalise the previous results by Hüsler (see below), which held for non-identically distributed random variables but under a uniform mixing condition, to the mixing situations typical in dynamical systems. Then we will apply our theoretical results to two important examples of non-stationary processes arising in dynamical systems.

The first example is given by *sequential dynamical systems*; they were introduced by Berend and Bergelson [6], as a non-stationary system in which a concatenation of maps is applied to a given point in the underlying space, and the probability is taken as a conformal measure, which is conformal for all maps considered and allows the use of the transfer operator (Perron–Fröbenius) as a useful tool to quantify the loss of memory of any prescribed initial observable. The theory of sequential systems was later developed in the fundamental paper by Conze and Raugi [9], where a few limit theorems, in particular the Central Limit Theorem, were proved for concatenations of one-dimensional dynamical systems, each possessing a transfer operator with a quasi-compact structure on a suitable Banach space. For the same systems and others, even in higher dimensions, the Almost Sure Invariance Principle was subsequently shown [18]; we will refer to the large class of systems investigated in [18] as concrete examples to which the non-stationary extreme value theory presented in this article applies.

The second example pertains to *random transformations*, which are constructed on a skew-system whose base is an invertible and hyperbolic system which codes a map on the second factor (this second factor could be seen as *fibers*, which are all copy of the same set). On these fibers live a family of *sample measures*, each of them corresponding to different ways to code the orbit of a given point. These sample measures will be taken as the probability measures that describe the statistical properties along the factor and they do not give rise to stationary processes (although they satisfy an interesting property when they move from one fiber to the other). Averaging along a sample measure means to fix the particular initial fiber which supports it; the dynamics will transport this measure from one fiber to the other, and this non-stationary process could be assimilated to a quenched process, where the map changes step by step according to a given realisation. We defer to the books by L. Arnold [3] and Y. Kifer [23,24] for a detailed account of these transformations, in particular for their ergodic properties. Limit theorems, in particular the CLT, were investigated in [25]. There are a few attempts to investigate recurrence in the framework of random transformations: see for instance [4,26,30–32].

1.2. *Extreme Value Laws for general non-stationary processes*

As mentioned in [10], the class of non-stationary stochastic processes is rather large and an extreme value theory for such a general class does not exist. In [21,22], Hüsler developed the first approach to the subject. Under convenient conditions, one can recover the usual extremal behaviour seen for i.i.d. or stationary sequences under Leadbetter's conditions. Of course the degree of freedom involved is so large that it is not difficult to give examples with pathological behaviour (see [22, Section 3] or [10, Example 9.4.4]). However, for appropriate subclasses, such as for stochastic processes of the form $X_i = a_i + b_i Y_i$, with trend values a_i, scaling values b_i and a stationary (or i.i.d.) stochastic process $Y_0, Y_1, \ldots$, one can study them and obtain the expected behaviour (see [28]).

The existing theory of extreme values for non-stationary sequences (which is still mostly based on Hüsler's results, see [10]) is not applicable in a dynamical setting because it is built over a uniform mixing condition obtained by adjusting to the non-stationary setting, Leadbetter's $D(u_n)$ condition for stationary processes. As was seen in the stationary setting in [8,11], this type of condition is not appropriate for stochastic processes arising from dynamical systems since it does not follow from usual properties regarding the loss of memory of chaotic systems, which are usually formulated in terms of decay of correlations. See discussion in Section 2 of [15] and Remarks 2.1 and 3.5 of the same paper.

Hence, the first goal of this paper is to develop a more general theory of extreme values for non-stationary stochastic processes, which enables the study of the extremal behaviour of the non-stationary systems discussed in the preceding Section. The major highlights of this generalisation are: the use of a much weaker mixing condition, motivated by an idea of Collet (in [8]) and further developed in [11,14,15], that we will adapt to the non-stationary setting and denote by a cyrilic D, i.e., Д, as in [15]; and a much more sophisticated way of dealing with clustering and the appearance of an Extremal Index less than 1, which is based on an idea introduced in [14] and further developed in [15], which basically says that when dealing with clustering due to the presence of a periodic phenomenon we can replace the role of the occurrence of exceedances (which in the dynamical setting correspond to hits to target ball sets) by that of the occurrence of escapes (which in the dynamical setting can be associated with hits to annuli target sets).

While in [21,22], Hüsler built on the existing theory of extreme values for stationary sequences developed by Leadbetter and others, here we will follow Hüsler's approach but adapt to the non-stationary setting the more refined [15].

2. A general result for Extreme Value Laws for non-stationary processes

In this section will try to keep as much as possible the notations used in [15,21,22].

Let $X_0, X_1, \dots$ be a stochastic process, where each r.v. $X_i : \mathcal{Y} \to \mathbb{R}$ is defined on the measure space $(\mathcal{Y}, \mathcal{B}, \mathbb{P})$.

We assume that $\mathcal{Y}$ is a sequence space with a natural product structure so that each possible realisation of the stochastic process corresponds to a unique element of $\mathcal{Y}$ and there exists a measurable map $\mathcal{T} : \mathcal{Y} \to \mathcal{Y}$, the time evolution map, which can be seen as the passage of one unit of time, so that

$$X_{i-1} \circ \mathcal{T} = X_i, \quad \text{for all } i \in \mathbb{N}.$$

The σ-algebra $\mathcal{B}$ can also be seen as a product σ-algebra adapted to the X_i's. For the purpose of this paper, $X_0, X_1, \dots$ is possibly non-stationary. Stationarity would mean that $\mathbb{P}$ is $\mathcal{T}$-invariant. Note that $X_i = X_0 \circ \mathcal{T}_i$, for all $i \in \mathbb{N}_0$, where $\mathcal{T}_i$ denotes the i-fold composition of $\mathcal{T}$, with the convention that $\mathcal{T}_0$ denotes the identity map on $\mathcal{Y}$. In the applications below to sequential dynamical systems, we will have that $\mathcal{T}_i = T_i \circ \cdots \circ T_1$ will be the concatenation of i possibly different transformations $T_1, \dots, T_i$.

Each random variable X_i has a marginal distribution function (d.f.) denoted by F_i, i.e., $F_i(x) = \mathbb{P}(X_i \leq x)$. Note that the F_i, with $i \in \mathbb{N}_0$, may all be distinct from each other. For a d.f. F we let $\bar{F} = 1 - F$. We define $u_{F_i} = \sup\{x : F_i(x) < 1\}$ and let $F_i(u_{F_i}-) := \lim_{h \to 0, h>0} F_i(u_{F_i} - h) = 1$ for all i.

Our main goal is to determine the limiting law of

$$\mathbf{P}_n = \mathbb{P}(X_0 \leq u_{n,0}, X_1 \leq u_{n,1}, \dots, X_{n-1} \leq u_{n,n-1})$$

as $n \to \infty$, where $\{u_{n,i}, i \leq n-1, n \geq 1\}$ is considered a real-valued boundary. We assume throughout the paper that

$$\bar{F}_{n,\max} := \max\{\bar{F}_i(u_{n,i}), i \leq n-1\} \to 0 \quad \text{as } n \to \infty, \tag{2.1}$$

which is equivalent to

$$u_{n,i} \to u_{F_i} \quad \text{as } n \to \infty, \text{ uniformly in } i.$$

Let us denote $F_n^* := \sum_{i=0}^{n-1} \bar{F}_i(u_{n,i})$, and assume that there is $\tau > 0$ such that

$$F_n^* := \sum_{i=0}^{n-1} \bar{F}_i(u_{n,i}) \to \tau, \quad \text{as } n \to \infty. \tag{2.2}$$

To simplify the notation let $u_i := u_{n,i}$.

In what follows, for every $A \in \mathcal{B}$, we denote the complement of A as $A^c := \mathcal{Y} \setminus A$.

Let $\mathbb{A} := (A_0, A_1, \dots)$ be a sequence of events such that $A_i \in \mathcal{T}_i^{-1}\mathcal{B}$. For some $s, \ell \in \mathbb{N}_0$, we define

$$\mathcal{W}_{s,\ell}(\mathbb{A}) = \bigcap_{i=s}^{s+\ell-1} A_i^c. \tag{2.3}$$

We will write $\mathcal{W}_{s,\ell}^c(\mathbb{A}) := (\mathcal{W}_{s,\ell}(\mathbb{A}))^c$.

For some $j \in \mathbb{N}_0$, we consider

$$\mathbb{A}_n^{(j)} := \left(A_{n,0}^{(j)}, A_{n,1}^{(j)}, \dots\right),$$

where the event $A_{n,i}^{(j)}$ is defined for $j \in \mathbb{N}$ as

$$A_{n,i}^{(j)} := \{X_i > u_{n,i}, X_{i+1} \leq u_{n,i+1}, \dots, X_{i+j} \leq u_{n,i+j}\}$$

and, for $j = 0$, we simply define $A_{n,i}^{(0)}(u_{n,i}) := \{X_i > u_{n,i}\}$.

For each $i \in \mathbb{N}_0$ and $n \in \mathbb{N}$, let $R_{n,i}^{(j)} = \min\{r \in \mathbb{N} : A_{n,i}^{(j)} \cap A_{n,i+r}^{(j)} \neq \emptyset\}$. We assume that there exists $q \in \mathbb{N}_0$ such that:

$$q = \min\Big\{ j \in \mathbb{N}_0 : \lim_{n\to\infty} \min_{i\le n}\big\{R_{n,i}^{(j)}\big\} = \infty \Big\}. \tag{2.4}$$

When $q = 0$ then $A_{n,i}^{(0)}(u_{n,i})$ corresponds to an exceedance of the threshold $u_{n,i}$ and we expect no clustering of exceedances.

When $q > 0$, heuristically one can think that there exists an underlying periodic phenomenon creating short recurrence, i.e., clustering of exceedances, when exceedances occur separated by no more than $q - 1$ units of time then they belong to the same cluster. Hence, the sets $A_{n,i}^{(q)}(u_{n,i})$ correspond to the occurrence of exceedances that escape the periodic phenomenon and are not followed by another exceedance in the same cluster. We will refer to the occurrence of $A_{n,i}^{(q)}(u_{n,i})$ as the occurrence of an escape at time i, whenever $q > 0$.

The following result adapts to the non-stationary setting an idea introduced in [14] and further developed in [15, Proposition 2.7], which essentially says the asymptotic distribution of $\mathbf{P}_n$ coincides with that of $\mathscr{W}_{0,n}(\mathbb{A}_n^{(q)})$, which motivates the special role played by $\mathbb{A}_n^{(q)}$ and the conditions we propose next.

Proposition 2.1. *Given events $B_0, B_1, \ldots \in \mathcal{B}$, let $r, q, n \in \mathbb{N}$ be such that $q < n$ and define $\mathbb{B} = (B_0, B_1, \ldots)$, $A_r = B_r \setminus \bigcup_{j=1}^{q} B_{r+j}$ and $\mathbb{A} = (A_0, A_1, \ldots)$. Then*

$$\big|\mathbb{P}\big(\mathscr{W}_{0,n}(\mathbb{B})\big) - \mathbb{P}\big(\mathscr{W}_{0,n}(\mathbb{A})\big)\big| \le \sum_{j=1}^{q} \mathbb{P}\big(\mathscr{W}_{0,n}(\mathbb{A}) \cap (B_{n-j} \setminus A_{n-j})\big).$$

Now, we introduce a mixing condition which is specially designed for the application to the dynamical setting, contrary to the existing ones in the literature.

***Condition** ($\text{Д}_q(u_{n,i})$). We say that $\text{Д}_q(u_n)$ holds for the sequence $X_0, X_1, \ldots$ if for every $\ell, t, n \in \mathbb{N}$,*

$$\big|\mathbb{P}\big(A^{(q)}{}_{n,i} \cap \mathscr{W}_{i+t,\ell}\big(\mathbb{A}_n^{(q)}\big)\big) - \mathbb{P}\big(A^{(q)}{}_{n,i}\big)\mathbb{P}\big(\mathscr{W}_{i+t,\ell}\big(\mathbb{A}_n^{(q)}\big)\big)\big| \le \gamma_i(q, n, t), \tag{2.5}$$

where $\gamma_i(q, n, t)$ is decreasing in t for each n and each i and there exists a sequence $(t_n^*)_{n\in\mathbb{N}}$ such that $t_n^* \bar{F}_{n,\max} \to 0$ and $\sum_{i=0}^{n-1} \gamma_i(q, n, t_n^*) \to 0$ when $n \to \infty$.

***Remark 2.2.** Condition $\text{Д}_q(u_{n,i})$ is a sort of mixing condition resembling Hüsler's adjustment of Leadbetter's condition $D(u_n)$ but with the great advantage that it can be checked for non-stationary dynamical systems, as we will see in Sections 4.2 and 5.1, contrary to Hüsler's $D(u_{n,i})$. This advantage resides on the fact that the event $A_{n,i}^{(q)}(u_{n,i})$ depends only on a finite number of random variables, making $\text{Д}_q(u_{n,i})$ a much weaker requirement in terms of uniformity when compared to Hüsler's $D(u_{n,i})$. Recall that Hüsler's $D(u_{n,i})$ required an uniform bound for all possible i and all possible numbers of random variables of the process on which the first event depended.*

In order to prove the existence of a distributional limit for $\mathbf{P}_n$ we use as usual a blocking argument that splits the data into k_n blocks separated by time gaps of size larger than t_n^*, which are created by simply disregarding the observations in the time frame occupied by the gaps. The precise construction of the blocks is given in Section 2.2 but we briefly describe below some of the properties of this construction.

In the stationary context, one takes blocks of equal size, which in particular means that the expected number of exceedances within each block is $n\mathbb{P}(X_0 > u_n)/k_n \sim \tau/k_n$. Here the blocks may have different sizes, which we will denote by $\ell_{n,1}, \ldots, \ell_{n,k_n}$ but, as in [21,22], these are chosen so that the expected number of exceedances is again $\sim\tau/k_n$. Also, for $i = 1, \ldots, k_n$, let $\mathcal{L}_{n,i} = \sum_{j=1}^{i} \ell_{n,j}$ and $\mathcal{L}_{n,0} = 0$.

The time gaps are created by disregarding the last observations in each block so that the true blocks become the remaining part. To do that, we have to balance the facts that we want the gaps to be big enough so that they are larger than t_n^* but on the other hand we also want the gaps to be sufficiently small so that the information disregarded does not compromise the computations. This is achieved by choosing the number of blocks, which correspond to the sequence $(k_n)_{n\in\mathbb{N}}$ diverging but slowly enough so that the weight of the gaps is negligible when compared to that of the true blocks.

As usual in extreme value theory, in order to guarantee the existence of a distributional limit one needs to impose some restrictions on the speed of recurrence.

For $q \in \mathbb{N}_0$ given by (2.4), consider the sequence $(t_n^*)_{n\in\mathbb{N}}$, given by condition $Д_q(u_n)$ and let $(k_n)_{n\in\mathbb{N}}$ be another sequence of integers such that

$$k_n \to \infty \quad \text{and} \quad k_n t_n^* \bar{F}_{n,\max} \to 0 \tag{2.6}$$

as $n \to \infty$.

Condition **($Д'_q(u_{n,i})$).** *We say that $Д'_q(u_{n,i})$ holds for the sequence $X_0, X_1, X_2, \ldots$ if there exists a sequence $(k_n)_{n\in\mathbb{N}}$ satisfying (2.6) and such that*

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} \sum_{j=0}^{\ell_i-1} \sum_{r>j}^{\ell_i-1} \mathbb{P}\big(A^{(q)}{}_{\mathcal{L}_{i-1}+j} \cap A^{(q)}{}_{\mathcal{L}_{i-1}+r}\big) = 0. \tag{2.7}$$

Condition $Д'_q(u_{n,i})$ precludes the occurrence of clustering of escapes (or exceedances, when $q = 0$).

Remark 2.3. *Note that condition $Д'_p(u_{n,i})$ is an adjustment of a similar condition $Д'_p(u_n)$ in [15] in the stationary setting, which is similar to (although slightly weaker than) condition $D^{(p+1)}(u_n)$ in the formulation of [7, Equation (1.2)]*

When $q = 0$, observe that $Д'_q(u_{n,i})$ is very similar to $D'(u_{n,i})$ from Hüsler, which prevents clustering of exceedances, just as $D'(u_n)$ introduced by Leadbetter did in the stationary setting.

When $q > 0$, we have clustering of exceedances, i.e., the exceedances have a tendency to appear aggregated in groups (called clusters). One of the main ideas in [14] that we use here is that the events $A^{(q)}{}_{n,i}$ play a key role in determining the limiting EVL and in identifying the clusters. In fact, when $Д'_q(u_{n,i})$ holds we have that every cluster ends with an entrance in $A^{(q)}{}_{n,i}$, meaning that the inter cluster exceedances must appear separated at most by q units of time.

In this approach, it is rather important to observe the prominent role played by condition $Д'_q(u_{n,i})$. In particular, note that if condition $Д'_q(u_{n,i})$ holds for some particular $q = q_0 \in \mathbb{N}_0$, then condition $Д'_q(u_{n,i})$ holds for all $q \geq q_0$. Then, q as defined in (2.4) is indeed the natural candidate to try to show the validity of $Д'_q(u_n)$.

We now give a way of defining the Extremal Index (EI) using the sets $A^{(q)}{}_{n,i}$. For $q \in \mathbb{N}_0$ given by (2.4), we also assume that there exists $0 \leq \theta \leq 1$, which will be referred to as the EI, such that

$$\lim_{n\to\infty} \max_{i=1,\ldots,k_n} \left\{ \left| \theta k_n \sum_{j=\mathcal{L}_{n,i-1}}^{\mathcal{L}_{n,i}-1} \bar{F}(u_{n,j}) - k_n \sum_{j=\mathcal{L}_{n,i-1}}^{\mathcal{L}_{n,i}-1} \mathbb{P}\big(A_{n,j}^{(q)}\big) \right| \right\} = 0. \tag{2.8}$$

The following is the main theorem of this section.

Theorem 2.4. *Let $X_0, X_1, \ldots$ be a stationary stochastic process and suppose (2.1) and (2.2) hold for some $\tau > 0$. Let $q \in \mathbb{N}_0$ be as in (2.4) and assume that (2.8) holds. Assume also that conditions $Д(u_{n,i})$ and $Д'_q(u_{n,i})$ are satisfied. Then*

$$\lim_{n\to\infty} \mathbf{P}_n = \mathrm{e}^{-\theta\tau}.$$

The rest of this section is devoted to the proof of Theorem 2.4.

To simplify notation, we will drop the index $n \in \mathbb{N}$ and write: $u_i := u_{n,i}$, $A_i^{(q)} := A_{n,i}^{(q)}$, $\mathbb{A}^{(q)} := \mathbb{A}_n^{(q)}$, $\ell_i := \ell_{n,i}$, $\mathcal{L}_i := \mathcal{L}_{n,i}$.

2.1. *Preliminaries to the argument*

We begin by proving the crucial observation stated in Proposition 2.1.

Proof of Proposition 2.1. Since $A_r \subset B_r$, then clearly $\mathcal{W}_{0,n}(\mathbb{B}) \subset \mathcal{W}_{0,n}(\mathbb{A})$. Hence, we have to estimate the probability of $\mathcal{W}_{0,n}(\mathbb{A}) \setminus \mathcal{W}_{0,n}(\mathbb{B})$.

Let $x \in \mathcal{W}_{0,n}(\mathbb{A}) \setminus \mathcal{W}_{0,n}(\mathbb{B})$. We will see that there exists $j \in \{1,\dots,q\}$ such that $x \in B_{n-j}$. In fact, suppose that no such j exists. Then let $\ell = \max\{i \in \{1,\dots,n-1\} : x \in B_i\}$. Then, clearly, $\ell < n-q$. Hence, if $x \notin B_j$, for all $i = \ell+1,\dots,n-1$, then we must have that $x \in A_\ell$ by definition of A. But this contradicts the fact that $x \in \mathcal{W}_{0,n}(\mathbb{A})$. Consequently, we have that there exists $j \in \{1,\dots,q\}$ such that $x \in B_{n-j}$ and since $x \in \mathcal{W}_{0,n}(\mathbb{A})$ then we can actually write $x \in B_{n-j} \setminus A_{n-j}$.

This means that $\mathcal{W}_{0,n}(\mathbb{A}) \setminus \mathcal{W}_{0,n}(\mathbb{B}) \subset \bigcup_{j=1}^{q}(B_{n-j} \setminus A_{n-j}) \cap \mathcal{W}_{0,n}(\mathbb{A})$ and then

$$\begin{aligned}
\left|\mathbb{P}\big(\mathcal{W}_{0,n}(\mathbb{B})\big) - \mathbb{P}\big(\mathcal{W}_{0,n}(\mathbb{A})\big)\right| &= \mathbb{P}\big(\mathcal{W}_{0,n}(\mathbb{A}) \setminus \mathcal{W}_{0,n}(\mathbb{B})\big) \\
&\le \mathbb{P}\left(\bigcup_{j=1}^{q}(B_{n-j} \setminus A_{n-j}) \cap \mathcal{W}_{0,n}(\mathbb{A})\right) \\
&\le \sum_{j=1}^{q} \mathbb{P}\big(\mathcal{W}_{0,n}(\mathbb{A}) \cap (B_{n-j} \setminus A_{n-j})\big),
\end{aligned}$$

as required. □

We prove next some lemmata that pave the way for Proposition 2.7, which is the cornerstone of the argument leading to the proof of Theorem 2.4

Lemma 2.5. *For any fixed* $\mathbb{A} = (A_0, A_1, \dots)$, $A_i \in \mathcal{B}$ *for* $i = 0, 1, \dots,$ *and integers* a, s, t, m, *with* $a < s$, *we have*:

$$\left|\mathbb{P}\big(\mathcal{W}_{a,s+t+m}(\mathbb{A})\big) - \mathbb{P}\big(\mathcal{W}_{a,s}(\mathbb{A}) \cap \mathcal{W}_{a+s+t,m}(\mathbb{A})\big)\right| \le \sum_{j=s}^{s+t-1} \mathbb{P}(A_{a+j}).$$

Proof.

$$\begin{aligned}
\mathbb{P}\big(\mathcal{W}_{a,s}(\mathbb{A}) \cap \mathcal{W}_{a+s+t,m}(\mathbb{A})\big) - \mathbb{P}\big(\mathcal{W}_{a,s+t+m}(\mathbb{A})\big) &= \mathbb{P}\big(\mathcal{W}_{a,s}(\mathbb{A}) \cap \mathcal{W}^c_{a+s,t}(\mathbb{A}) \cap \mathcal{W}_{a+s+t,m}(\mathbb{A})\big) \\
&\le \mathbb{P}\big(\mathcal{W}^c_{a+s,t}(\mathbb{A})\big) = \mathbb{P}\left(\bigcup_{j=s}^{s+t-1}(A_{a+j})\right) \\
&\le \sum_{j=s}^{s+t-1} \mathbb{P}(A_{a+j}).
\end{aligned}$$

□

Lemma 2.6. *For any fixed* $\mathbb{A} = (A_0, A_1, \dots)$, $A_i \in \mathcal{B}$ *for* $i = 0, 1, \dots,$ *and integers* a, s, t, m, *with* $a < s$, *we have*:

$$\begin{aligned}
&\left|\mathbb{P}\big(\mathcal{W}_{a,s}(\mathbb{A}) \cap \mathcal{W}_{a+s+t,m}(\mathbb{A})\big) - \mathbb{P}\big(\mathcal{W}_{a+s+t,m}(\mathbb{A})\big)\left(1 - \sum_{j=0}^{s-1}\mathbb{P}(A_{a+j})\right)\right| \\
&\quad \le \left|\sum_{j=0}^{s-1}\mathbb{P}(A_{a+j})\mathbb{P}\big(\mathcal{W}_{a+s+t,m}(\mathbb{A})\big) - \sum_{j=0}^{s-1}\mathbb{P}\big(A_{a+j} \cap \mathcal{W}_{a+s+t,m}(\mathbb{A})\big)\right| + \sum_{j=0}^{s-1}\sum_{i>j}^{s-1}\mathbb{P}(A_{a+i} \cap A_{a+j}).
\end{aligned}$$

Proof. Observe that

$$\begin{aligned}
&\left|\mathbb{P}\big(\mathscr{W}_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)-\mathbb{P}\big(\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)\left(1-\sum_{j=0}^{s-1}\mathbb{P}(A_{a+j})\right)\right| \\
&\quad\le\left|\sum_{j=0}^{s-1}\mathbb{P}(A_{a+j})\mathbb{P}\big(\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)-\sum_{j=0}^{s-1}\mathbb{P}\big(A_{a+j}\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)\right| \\
&\qquad+\left|\mathbb{P}\big(\mathscr{W}_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)-\mathbb{P}\big(\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)+\sum_{j=0}^{s-1}\mathbb{P}\big(A_{a+j}\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)\right|.
\end{aligned}$$

Regarding the second term on the right, we have

$$\mathbb{P}\big(\mathscr{W}_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)=\mathbb{P}\big(\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)-\mathbb{P}\big(\mathscr{W}^c_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big).$$

Now, since $\mathscr{W}^c_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})=\bigcup_{i=0}^{s-1}(A_{a+i}\cap\mathscr{W}_{a+s+t,m}(\mathbb{A}))$, we have

$$\mathbb{P}\big(\mathscr{W}^c_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)\le\sum_{i=0}^{s-1}\big(A_{a+i}\cap\mathbb{P}\big(\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)\big)$$

and so,

$$\begin{aligned}
0&\le\sum_{j=0}^{s-1}\mathbb{P}\big(A_{a+j}\cap\mathscr{W}_{s+t,m}(\mathbb{A})\big)-\mathbb{P}\big(\mathscr{W}^c_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big) \\
&\le\sum_{j=0}^{s-1}\sum_{i>j}^{s-1}\mathbb{P}\big(A_{a+i}\cap A_{a+j}\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big).
\end{aligned}$$

Hence, using these last computations we get:

$$\begin{aligned}
&\left|\mathbb{P}\big(\mathscr{W}_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)-\mathbb{P}\big(\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)+\sum_{j=0}^{s-1}\mathbb{P}\big(A_{a+j}\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)\right| \\
&\quad=\left|-\mathbb{P}\big(\mathscr{W}^c_{a,s}(\mathbb{A})\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)+\sum_{j=0}^{s-1}\mathbb{P}\big(A_{a+j}\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big)\right| \\
&\quad\le\sum_{j=0}^{s-1}\sum_{i>j}^{s-1}\mathbb{P}\big(A_{a+i}\cap A_{a+j}\cap\mathscr{W}_{a+s+t,m}(\mathbb{A})\big) \\
&\quad\le\sum_{j=0}^{s-1}\sum_{i>j}^{s-1}\mathbb{P}(A_{a+i}\cap A_{a+j}).
\end{aligned}$$

□

2.2. *The construction of the blocks*

The construction of the blocks here, contrary to the stationary case, in which the blocks have equal size, is designed so that the expected number of exceedances in each block is the same. We follow closely the construction in [21,22].

For each $n \in \mathbb{N}$ we split the random variables $X_0, \dots, X_{n-1}$ into k_n initial blocks, where k_n is given by (2.6), of sizes $\ell_1, \dots, \ell_{k_n}$ defined in the following way. Let as before $\mathcal{L}_i = \sum_{j=1}^{i} \ell_i$ and $\mathcal{L}_0 = \ell_0 = 0$. Assume that $\ell_1, \dots, \ell_{i-1}$ are already defined. Take ℓ_i to be the largest integer such that:

$$\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_{i-1}+\ell_i-1} \bar{F}(u_{n,i}) \le \frac{F_n^*}{k_n}.$$

The final working blocks are obtained by disregarding the last observations of each initial block, which will create a time gap between each final block. The size of the time gaps must be balanced in order to have at least a size t_n^* but such that its weight on the average number of exceedances is negligible when compared to that of the final blocks. For that purpose we define

$$\varepsilon(n) := \left(t_n^* + 1\right)\bar{F}_{\max}\frac{k_n}{F_n^*}.$$

Note that by (2.2) and (2.6), it follows immediately that $\lim_{n\to\infty} \varepsilon(n) = 0$. Now, for each $i = 1, \dots, k_n$ let t_i be the largest integer such that

$$\sum_{j=\mathcal{L}_i-t_i}^{\mathcal{L}_i-1} \bar{F}(u_{n,i}) \le \varepsilon(n)\frac{F_n^*}{k_n}.$$

Hence, the final working blocks correspond to the observations within the time frame $\mathcal{L}_{i-1} + 1, \dots, \mathcal{L}_i - t_i$, while the time gaps correspond to the observations in the time frame $\mathcal{L}_i - t_i + 1, \dots, \mathcal{L}_i$, for all $i = 1, \dots, k_n$.

Note that $t_n^* \le t_i < \ell_i$, for each $i = 1, \dots, k_n$. The second inequality is trivial. For the first inequality note that by definition of t_i we have

$$\varepsilon(n)\frac{F_n^*}{k_n} \le \sum_{j=\mathcal{L}_i-t_i}^{\mathcal{L}_i-1} \bar{F}(u_{n,i}) + \bar{F}(u_{n,\mathcal{L}_i-t_i-1}) \le (t_i+1)\bar{F}_{\max}.$$

The first inequality follows easily now by definition of $\varepsilon(n)$.

Proposition 2.7. *For every, $n \in \mathbb{N}$, let $\mathbb{A} := \mathbb{A}_n^{(q)}$ for q defined by (2.4). Consider the construction of the k_n blocks above, the respective sizes $\ell_1, \dots, \ell_{k_n}$ and time gaps $t_1, \dots, t_{k_n}$. Recall that $\mathcal{L}_i = \sum_{j=1}^{i} \ell_i$. Assume that $n \in \mathbb{N}$ is large enough so that $F_n^*/k_n < 2$. We have*:

$$\begin{aligned}
&\left|\mathbb{P}\big(\mathscr{W}_{0,n}(\mathbb{A})\big) - \prod_{i=1}^{k_n}\left(1 - \sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-t_i-1} \mathbb{P}(A_j)\right)\right| \\
&\quad \le \sum_{i=1}^{k_n} \sum_{j=\mathcal{L}_{i-1}-t_i}^{\mathcal{L}_i-1} \mathbb{P}\big(A_j^{(q)}\big) + \sum_{j=\mathcal{L}_{k_n}}^{n-1} \mathbb{P}\big(A_j^{(q)}\big) \\
&\qquad + \sum_{i=1}^{k_n}\left|\sum_{j=0}^{\ell_i-t_i-1} \big(\mathbb{P}(A_{\mathcal{L}_{i-1}+j})\mathbb{P}\big(\mathscr{W}_{\mathcal{L}_i,\mathcal{L}_{k_n}-\mathcal{L}_i}(\mathbb{A})\big) - \mathbb{P}\big(A_{\mathcal{L}_{i-1}+j} \cap \mathscr{W}_{\mathcal{L}_i,\mathcal{L}_{k_n}-\mathcal{L}_i}(\mathbb{A})\big)\big)\right| \\
&\qquad + \sum_{i=1}^{k_n} \sum_{j=0}^{\ell_i-1} \sum_{r>j}^{\ell_i-1} \mathbb{P}(A_{\mathcal{L}_{i-1}+j} \cap A_{\mathcal{L}_{i-1}+r}).
\end{aligned}$$

Proof. Using Lemma 2.5, we have:

$$\left|\mathbb{P}\big(\mathcal{W}_{0,n}(\mathbb{A})\big) - \mathbb{P}\big(\mathcal{W}_{0,\mathcal{L}_{k_n}}(\mathbb{A})\big)\right| \le \sum_{j=\mathcal{L}_{k_n}}^{n-1} \mathbb{P}\big(A_j^{(q)}\big). \tag{2.9}$$

To simplify the notation let $\bar{\mathcal{L}}_i = \mathcal{L}_{k_n} - \mathcal{L}_{i-1} = \sum_{j=i}^{k_n} \ell_j$. It follows by using (2.6) that

$$\begin{aligned}
&\left|\mathbb{P}\big(\mathcal{W}_{\mathcal{L}_{i-1},\bar{\mathcal{L}}_i}(\mathbb{A})\big) - \left(1 - \sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i - t_i - 1} \mathbb{P}(A_j)\right)\mathbb{P}\big(\mathcal{W}_{\mathcal{L}_i,\bar{\mathcal{L}}_{i+1}}(\mathbb{A})\big)\right| \\
&\quad \le \left|\mathbb{P}\big(\mathcal{W}_{\mathcal{L}_{i-1},\bar{\mathcal{L}}_i}(\mathbb{A})\big) - \mathbb{P}\big(\mathcal{W}_{\mathcal{L}_{i-1},\ell_i - t_i}(\mathbb{A}) \cap \mathcal{W}_{\mathcal{L}_i,\bar{\mathcal{L}}_{i+1}}(\mathbb{A})\big)\right| \\
&\qquad + \left|\mathbb{P}\big(\mathcal{W}_{\mathcal{L}_{i-1},\ell_i - t_i}(\mathbb{A}) \cap \mathcal{W}_{\mathcal{L}_i,\bar{\mathcal{L}}_{i+1}}(\mathbb{A})\big) - \left(1 - \sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i - t_i - 1} \mathbb{P}(A_j)\right)\mathbb{P}\big(\mathcal{W}_{\mathcal{L}_i,\bar{\mathcal{L}}_{i+1}}(\mathbb{A})\big)\right| \\
&\quad \le \sum_{j=\mathcal{L}_{i-1}-t_i}^{\mathcal{L}_i - 1} \mathbb{P}(A_j) + \left|\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i - t_i - 1} \big(\mathbb{P}(A_j)\mathbb{P}\big(\mathcal{W}_{\mathcal{L}_{i-1},\bar{\mathcal{L}}_i}(\mathbb{A})\big) - \mathbb{P}\big(A_j \cap \mathcal{W}_{\mathcal{L}_{i-1},\bar{\mathcal{L}}_i}(\mathbb{A})\big)\right| \\
&\qquad + \sum_{j=0}^{\ell_i - 1} \sum_{r>j}^{\ell_i - 1} \mathbb{P}(A_{\mathcal{L}_{i-1}+j} \cap A_{\mathcal{L}_{i-1}+r}).
\end{aligned} \tag{2.10}$$

Let

$$\begin{aligned}
\Upsilon_i :={}& \sum_{j=\mathcal{L}_{i-1}-t_i}^{\mathcal{L}_i - 1} \mathbb{P}(A_j) + \left|\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i - t_i - 1} \mathbb{P}(A_j)\mathbb{P}\big(\mathcal{W}_{\mathcal{L}_{i-1},\bar{\mathcal{L}}_i}(\mathbb{A})\big) - \mathbb{P}\big(A_j \cap \mathcal{W}_{\mathcal{L}_{i-1},\bar{\mathcal{L}}_i}(\mathbb{A})\big)\right| \\
&+ \sum_{j=0}^{\ell_i - 1} \sum_{r>j}^{\ell_i - 1} \mathbb{P}(A_{\mathcal{L}_{i-1}+j} \cap A_{\mathcal{L}_{i-1}+r}).
\end{aligned}$$

Note that, for $i = k_n$ in (2.10), $|\mathcal{W}_{\mathcal{L}_{k_n-1},\bar{\mathcal{L}}_{k_n}}(\mathbb{A})) - (1 - \sum_{j=\mathcal{L}_{k_n-1}}^{\mathcal{L}_{k_n} - t_{k_n} - 1} \mathbb{P}(A_j))| \le \Upsilon_{k_n}$.

Since $\frac{F_n^*}{k_n} < 2$ and, by construction, for all $i = 1, \dots, k_n$, it is clear that $\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i - t_i - 1} \mathbb{P}(A_j) \le \frac{F_n^*}{k_n}$, then $|1 - \sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i - t_i - 1} \mathbb{P}(A_j)| < 1$, for all $i = 1, \dots, k_n$.

Now, we use (2.10) recursively and obtain

$$\left|\mathbb{P}\big(\mathcal{W}_{0,\mathcal{L}_{k_n}}(\mathbb{A})\big) - \prod_{i=1}^{k_n}\left(1 - \sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i - t_i - 1} \mathbb{P}(A_j)\right)\right| \le \sum_{i=1}^{k_n} \Upsilon_i. \tag{2.11}$$

The result follows now at once from (2.9) and (2.11). $\square$

2.3. *Final argument*

We are now in a position to prove Theorem 2.4.

Proof of Theorem 2.4. The theorem follows if we show that all the error terms in Proposition 2.7 converge to 0, as $n \to \infty$.

For the first term, by choice of the t_i's, we have

$$\sum_{i=1}^{k_n}\sum_{j=\mathcal{L}_{i-1}-t_i}^{\mathcal{L}_i-1}\mathbb{P}\big(A_j^{(q)}\big)\leq\sum_{i=1}^{k_n}\sum_{j=\mathcal{L}_{i-1}-t_i}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})\leq k_n\varepsilon(n)\frac{F_n^*}{k_n}=\varepsilon(n)F_n^*,$$

which tends to 0 as $n\to\infty$, by (2.2) and definition of $\varepsilon(n)$.

Regarding the second term observe first that

$$\sum_{j=\mathcal{L}_{k_n}}^{n-1}\mathbb{P}\big(A_j^{(q)}\big)\leq\sum_{j=\mathcal{L}_{k_n}}^{n-1}\bar{F}(u_{n,j}).$$

Since, by choice of ℓ_i, we have $\frac{F_n^*}{k_n}\leq\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})+\bar{F}(u_{n,\mathcal{L}_i})\leq\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})+\bar{F}_{\max}$, then it follows that

$$\frac{F_n^*}{k_n}-\bar{F}_{\max}\leq\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})\leq\frac{F_n^*}{k_n}.\tag{2.12}$$

From the first inequality we get $F_n^*-k_n\bar{F}_{\max}\leq\sum_{i=1}^{k_n}\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})$, which implies that

$$\sum_{j=\mathcal{L}_{k_n}}^{n-1}\bar{F}(u_{n,j})=F_n^*-\sum_{i=1}^{k_n}\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})\leq k_n\bar{F}_{\max},$$

which goes to 0 as $n\to\infty$ by (2.6).

For the third term, recalling that, for each n and i, $\gamma_i(q,n,t)$ from condition $Д_q(u_{n,i})$ is decreasing in t, we have:

$$\sum_{i=1}^{k_n}\left|\sum_{j=0}^{\ell_i-t_i-1}\Big(\mathbb{P}\big(A^{(q)}{}_{\mathcal{L}_{i-1}+j}\big)\mathbb{P}\big(\mathscr{W}_{\mathcal{L}_i,\mathcal{L}_{k_n}-\mathcal{L}_i}(\mathbb{A})\big)-\mathbb{P}\big(A^{(q)}{}_{\mathcal{L}_{i-1}+j}\cap\mathscr{W}_{\mathcal{L}_i,\mathcal{L}_{k_n}-\mathcal{L}_i}(\mathbb{A})\big)\Big)\right|\leq\sum_{i=0}^{n-1}\gamma_i(q,n,t_n),$$

which tends to 0 as $n\to\infty$ by condition $Д_q(u_{n,i})$.

By condition $Д'(u_n)$, we have that the fourth term goes to 0 as $n\to\infty$.

Now, we will see that

$$\left|\prod_{i=1}^{k_n}\left(1-\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-t_i-1}\mathbb{P}\big(A^{(q)}{}_j\big)\right)-\mathrm{e}^{-\theta\tau}\right|\xrightarrow[n\to\infty]{}0.$$

By (2.8) we have that $k_n\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\mathbb{P}(A^{(q)}{}_j)=k_n\theta\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})+o(1)$. Then

$$\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\mathbb{P}\big(A^{(q)}{}_j\big)=\theta\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})+o\big(k_n^{-1}\big).$$

Since by (2.6), we have $\bar{F}_{\max}=o(k_n^{-1})$, then, by (2.12), it follows that

$$\sum_{j=\mathcal{L}_{i-1}}^{\mathcal{L}_i-1}\bar{F}(u_{n,j})+o\big(k_n^{-1}\big)=\frac{F_n^*}{k_n}+o\big(k_n^{-1}\big).$$

Also note that

$$\sum_{j=\mathcal{L}_i-t_i}^{\mathcal{L}_i-1} \mathbb{P}\big(A^{(q)}{}_j\big) \leq \sum_{j=\mathcal{L}_i-t_i}^{\mathcal{L}_i-1} \bar{F}(u_{n,j}) \leq \varepsilon(n)\frac{F_n^*}{k_n} = o\big(k_n^{-1}\big).$$

Hence, for all $i = 1, \ldots, k_n$ we have

$$\sum_{j=\mathcal{L}_i-t_i}^{\mathcal{L}_i-t_i-1} \mathbb{P}\big(A^{(q)}{}_j\big) = \theta\frac{F_n^*}{k_n} + o\big(k_n^{-1}\big).$$

Finally, by (2.2), we have

$$\prod_{i=1}^{k_n}\left(1 - \sum_{j=\mathcal{L}_i-t_i}^{\mathcal{L}_i-t_i-1} \mathbb{P}\big(A^{(q)}{}_j\big)\right) \sim \left(1 - \theta\frac{F_n^*}{k_n} + o\big(k_n^{-1}\big)\right)^{k_n} \underset{n\to\infty}{\longrightarrow} \mathrm{e}^{-\theta\tau}.$$

Finally, by Proposition 2.1 we have

$$\begin{aligned}\big|P_n - \mathbb{P}\big(\mathcal{W}_{0,n}\big(\mathbb{A}^{(q)}\big)\big)\big| &\leq \sum_{j=1}^{q} \mathbb{P}\big(\mathcal{W}_{0,n}\big(\mathbb{A}^{(q)}\big) \cap \big(\{X_{n-j} > u_{n,n-j}\} \setminus \big\{A_{n-j}^{(q)}\big\}\big)\big) \\ &\leq \sum_{j=1}^{q} \mathbb{P}\big(\{X_{n-j} > u_{n,n-j}\} \setminus \big\{A_{n-j}^{(q)}\big\}\big) \\ &\leq \sum_{j=1}^{q}\big(1 - F_{n-j}(u_{n,n-j})\big), \end{aligned} \tag{2.13}$$

which converges to 0 as $n \to \infty$.

Note that when $q = 0$ both sides of inequality (2.13) equal 0. □

3. Sequential dynamical systems

3.1. *General presentation*

In this section we will give a first example of a non-stationary process, by considering families $\mathcal{F}$ of non-invertible maps defined on compact subsets X of $\mathbb{R}^d$ or on the torus $\mathbb{T}^d$ (still denoted with X in the following), and non-singular with respect to the Lebesgue or the Haar measure, i.e. $m(A) \neq 0 \Longrightarrow m(T(A)) \neq 0$. Such measures will be defined on the Borel sigma algebra $\mathcal{B}$. We will be mostly concerned with the case $d = 1$. A countable sequence of maps $\{T_k\}_{k\geq 1} \in \mathcal{F}$ defines a *sequential dynamical system*. A *sequential orbit* of $x \in X$ will be defined by the concatenation

$$\mathcal{T}_n(x) := T_n \circ \cdots \circ T_1(x), \quad n \geq 1. \tag{3.1}$$

We denote by P_j the Perron–Fröbenius (transfer) operator associated to T_j defined by the duality relation

$$\int_X P_j f g\, dm = \int_X f g \circ T_j\, dm, \quad \text{for all } f \in L_m^1, g \in L_m^\infty.$$

Note that here the transfer operator P_j is defined with respect to the reference Lebesgue measure m.

Similarly to (3.1), we define the composition of operators as

$$\Pi_n := P_n \circ \cdots \circ P_1, \quad n \geq 1. \tag{3.2}$$

It is easy to check that duality persists under concatenation, namely

$$\int_X g(\mathcal{T}_n) f \, dm = \int_X g(T_n \circ \cdots \circ T_1) f \, dm = \int_X g(P_n \circ \cdots \circ P_1 f) \, dm = \int_X g(\Pi_n f) \, dm. \tag{3.3}$$

In [9] the authors begin a systematic study of the statistical properties of sequential dynamical systems by proving in particular the law of large numbers and the central limit theorem. In [18], it was shown that the Almost Sure Invariance Principle still holds. In order to establish such results a few assumptions are needed and some of them are also relevant for the extreme value theory. We will recall them in this section and then we will provide a list of examples which will go beyond the β transformations, which was the prototype case investigated by Conze and Raugi.

We first need to choose a suitable couple of adapted spaces in order to get and exploit the quasi-compactness of the transfer operator. We will consider in particular a Banach space $\mathcal{V} \subset L^1_m$ $(1 \in \mathcal{V})$ of functions over X with norm $\|\cdot\|_\alpha$, such that $\|\phi\|_\infty \le C\|\phi\|_\alpha$.

For example, we could let $\mathcal{V}$ be the Banach space of bounded variation functions over X with norm $\|\cdot\|_{\mathrm{BV}}$ given by the sum of the L^1_m norm and the total variation $|\cdot|_{\mathrm{BV}}$, or we could take $\mathcal{V}$ to be the space of quasi-Hölder functions with a suitable norm which we will define later on.

One of the basic assumption is the following:

Uniform Doeblin–Fortet–Lasota–Yorke inequality (DFLY): There exist constants $A, B < \infty$, $\rho \in (0,1)$, such that for any n and any sequence of operators $P_n, \dots, P_1$ associated to transformations in $\mathcal{F}$ and any $f \in \mathcal{V}$ we have

$$\|P_n \circ \cdots \circ P_1 f\|_\alpha \le A\rho^n \|f\|_\alpha + B\|f\|_1. \tag{3.4}$$

At this point one would like to dispose of a sort of quasi-compactness argument which would allow to get exponential decay for the composition of operators. In all the examples we will present, the class $\mathcal{F}$ will be constructed *around* (this will be made clear in a moment) a given map T_0 for which the corresponding operator P_0 will satisfy quasi-compactness. Namely we require:

Exactness property: The operator P_0 has a spectral gap, which implies that there are two constants $C_1 < \infty$ and $\gamma_0 \in (0,1)$ so that

$$\left\|P_0^n f\right\|_\alpha \le C_1 \gamma_0^n \|f\|_\alpha \tag{3.5}$$

for all $f \in \mathcal{V}$ of zero (Lebesgue) mean and $n \ge 1$.

The next step is to consider the following distance between two operators P and Q associated to maps in $\mathcal{F}$ and acting on $\mathcal{V}$:

$$d(P,Q) = \sup_{f \in \mathcal{V}, \|f\|_\alpha \le 1} \|Pf - Qf\|_1.$$

A very useful criterion is given in Proposition 2.10 in [9], and in our setting it reads: *if P_0 verifies the exactness property, then there exists $\delta_0 > 0$, such that the set $\{P \in \mathcal{F}; d(P, P_0) < \delta_0\}$ satisfies the decorrelation* (DEC) *condition*, where

Property (DEC): Given the family $\mathcal{F}$ there exist constants $\hat{C} > 0$, $\hat{\gamma} \in (0,1)$, such that for any n and any sequence of transfer operators $P_n, \dots, P_1$ corresponding to maps chosen from $\mathcal{F}$ and any $f \in \mathcal{V}$ of zero (Lebesgue) mean,[1] we have

$$\|P_n \circ \cdots \circ P_1 f\|_\alpha \le \hat{C}\hat{\gamma}^n \|f\|_\alpha. \tag{3.6}$$

By induction on the Doeblin–Fortet–Lasota–Yorke inequality for compositions we immediately have

$$d\big(P_r \circ \cdots \circ P_1, P_0^r\big) \le M \sum_{j=1}^{r} d(P_j, P_0), \tag{3.7}$$

[1] Actually, the definition of the (DEC) property in [9] is slightly more general since it requires the above property for functions in a suitable subspace, not necessarily that of functions with zero expectation.

with $M = 1 + A\rho^{-1} + B$.

According to [9, Lemma 2.13], (3.5) and (3.7) imply that there exists a constant C_2 such that

$$\|P_n \circ \cdots \circ P_1\phi - P_0^n\phi\|_1 \le C_2\|\phi\|_{\mathrm{BV}}\left(\sum_{k=1}^{p} d(P_{n-k+1}, P_0) + (1-\gamma_0)^{-1}\gamma_0^p\right)$$

for all integers $p \le n$ and all functions $\phi \in \mathcal{V}$. We will use this bound to get a quantitative rate of the exponential decay for composition of operators in the L_m^1 norm when we relate it to the following two assumptions:

Lipschitz continuity property: Assume that the maps (and their transfer operators) are parametrised by a sequence of numbers ε_k, $k \in \mathbb{N}$, such that $\lim_{k\to\infty} \varepsilon_k = \varepsilon_0$ ($P_{\varepsilon_0} = P_0$). We assume that there exists a constant C_3 so that

$$d(P_{\varepsilon_k}, P_{\varepsilon_j}) \le C_3|\varepsilon_k - \varepsilon_j|, \quad \text{for all } k, j \ge 0.$$

We will restrict in the following to the subclass $\mathcal{F}_{\mathrm{exa}}$ of maps, and therefore of operators, for which

$$\mathcal{F}_{\mathrm{exa}} := \left\{P_{\varepsilon_k} \in \mathcal{F};\, |\varepsilon_k - \varepsilon_0| < C_3^{-1}\delta_0\right\}.$$

The maps in $\mathcal{F}_{\mathrm{exa}}$ will therefore verify the (DEC) condition, but we will sometimes need something stronger, namely:

Convergence property: We require algebraic convergence of the parameters, that is, there exist a constant C_4 and $\kappa > 0$ so that

$$|\varepsilon_n - \varepsilon_0| \le \frac{C_4}{n^\kappa} \quad \forall n \ge 1.$$

With these last assumptions, we get a polynomial decay for (3.7) of the type $O(n^{-\kappa})$ and in particular we obtain the same algebraic convergence in L_m^1 of $P_n \circ \cdots \circ P_1\phi$ to $h\int \phi\, dm$, where h is the density of the absolutely continuous mixing measure of the map T_0.

3.2. *Stochastic processes for sequential systems*

Similarly to [12] (in the context of stationary deterministic systems), we consider that the time series $X_0, X_1, \ldots$ arises from these sequential systems simply by evaluating a given observable $\varphi : X \to \mathbb{R} \cup \{\pm\infty\}$ along the sequential orbits.

$$X_n = \varphi \circ \mathcal{T}_n, \quad \text{for each } n \in \mathbb{N}. \tag{3.8}$$

Note that, contrary to the setup in [12], the stochastic process $X_0, X_1, \ldots$ defined in this way is not necessarily stationary.

We assume that the r.v. $\varphi : X \to \mathbb{R} \cup \{\pm\infty\}$ achieves a global maximum at $\zeta \in X$ (we allow $\varphi(\zeta) = +\infty$) being of following form:

$$\varphi(x) = g\big(\mathrm{dist}(x, \zeta)\big), \tag{3.9}$$

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

Type g_1: there exists some strictly positive function $h : W \to \mathbb{R}$ such that for all $y \in \mathbb{R}$

$$\lim_{s\to g_1(0)} \frac{g_1^{-1}(s + yh(s))}{g_1^{-1}(s)} = \mathrm{e}^{-y}; \tag{3.10}$$

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

$$\lim_{s\to+\infty} \frac{g_2^{-1}(sy)}{g_2^{-1}(s)} = y^{-\beta}; \tag{3.11}$$

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

$$\lim_{s\to 0} \frac{g_3^{-1}(D - sy)}{g_3^{-1}(D - s)} = y^\gamma. \tag{3.12}$$

It may be shown that no non-degenerate limit applies if $\int_0^{g_1(0)} g_1^{-1}(s)\, ds$ is not finite. Hence, an appropriate choice of h in the Type 1 case is given by $h(s) = \int_s^{g_1(0)} g_1^{-1}(t)\, dt / g_1^{-1}(s)$ for $s < g_1(0)$.

Examples of each one of the three types are as follows: $g_1(x) = -\log x$ (in this case (3.10) is easily verified with $h \equiv 1$), $g_2(x) = x^{-1/\alpha}$ for some $\alpha > 0$ (condition (3.11) is verified with $\beta = \alpha$) and $g_3(x) = D - x^{1/\alpha}$ for some $D \in \mathbb{R}$ and $\alpha > 0$ (condition (3.12) is verified with $\gamma = \alpha$).

3.3. *Examples*

We now give a few examples of sequential systems satisfying the preceding assumptions. The family of maps $\mathcal{F}$ will be parametrised by a small positive number ε (or a vector with small positive components) and we will tacitly suppose that we restrict to $\mathcal{F}_{\text{exa}}$ having previously proved that the transfer operator P_0 for a reference map T_0 is exact. This will impose restrictions on the choice of ε (less than a constant times δ_0, see above), and in this case we will use the terminology *for ε small enough*. The verification of the DFLY condition, which in turn will imply the analogous condition for the unperturbed operator P_0 will usually follow from standard arguments and the exactness of P_0 will be proved by assuming the existence of a unique mixing absolutely continuous invariant measure (for instance by adding further properties to the map T_0), or alternatively by restricting to one of the finitely many mixing components prescribed by the quasi-compactness of P_0.

The following examples have already been introduced and treated in [18], but in the latter paper a much stronger condition was required, namely that there exists $\delta > 0$ such that for any sequence $P_n, \ldots, P_1$ in $\mathcal{F}$ we have the uniform lower bound

$$\inf_{x\in M} P_n \circ \cdots \circ P_1 1(x) \geq \delta, \quad \forall n \geq 1. \tag{3.13}$$

We do not need that property in the context of EVT.

3.3.1. *β transformation*

Let $\beta > 1$ and denote by $T_\beta(x) = \beta x \bmod 1$ the β-transformation on the unit circle. Similarly, for $\beta_k \geq 1 + c > 1$, $k = 1, 2, \ldots$, we have the transformations T_{β_k} of the same kind, $x \mapsto \beta_k x \bmod 1$. Then $\mathcal{F} = \{T_{\beta_k} : k\}$ is the family of transformations we want to consider here. The property (DEC) was proved in [9, Theorem 3.4(c)] and continuity (Lip) is precisely the content of Section 5 still in [9].

3.3.2. *Random additive noise*

In this second example we consider piecewise uniformly expanding maps T on the unit interval $M = [0, 1]$ which preserve a unique absolutely continuous invariant measure μ which is also mixing. We denote by $A_k, k = 1, \ldots, m$ the m open intervals of monotonicity of the map T which give a partition mod-0 of the unit interval. The map T is C^2 over the A_k and with a C^2 extension on the boundaries. We put $\min_{x\in M} |DT(x)| \geq \lambda > 1$; $\max_{x\in M} |DT(x)| \leq \Lambda$; $\sup_{x\in M} |\frac{D^2T_\varepsilon(x)}{DT_\varepsilon(x)}| \leq C_1 < \infty$. We will perturb with additive noise, namely we will consider a family of maps $\mathcal{F}$ given by $T_\varepsilon(x) = T(x) + \varepsilon$, where $\varepsilon \in U$ and such that $\forall \varepsilon \in U$ we have the images $T_\varepsilon A_k, k = 1, \ldots, m$ *strictly* included in $[0, 1]$. We will also suppose that $\exists A_w$ such that $\forall T_\varepsilon \in \mathcal{F}$ and $k = 1, \ldots, m : T_\varepsilon A_k \supset A_w$; moreover there exists $1 \geq L' > 0$ such that $\forall k = 1, \ldots, m$ and $\forall T_\varepsilon \in \mathcal{F}$, $|T_\varepsilon(A_w) \cap A_k| > L'$. These conditions are useful in obtaining distortion bounds. We note that our assumptions are satisfied if we consider C^2 uniformly expanding maps on the circle and again perturbed with additive noise, without, this time, any restriction of the values of ε. In particular, the intervals of local injectivity $A_k, k = 1, \ldots, m$, of T_ε are now independent of ε. The functional space $\mathcal{V}$ will coincide with the functions of bounded variation with norm $\|\cdot\|_{\text{BV}}$.

The (DFLY) inequality follows easily with standard arguments. The next step is to show that two operators are close when the relative perturbation parameters are close: we report here for completeness the short proof already given in [18]. We thus consider the difference $\|\hat{P}_{\varepsilon_1} f - \hat{P}_{\varepsilon_2} f\|_1$, with f in BV. We have

$$\hat{P}_{\varepsilon_1} f(x) - \hat{P}_{\varepsilon_2} f(x) = \sum_{l=1}^{m} f \cdot \mathbf{1}_{U_n^c}\big(T_{\varepsilon_1,l}^{-1} x\big)\left[\frac{1}{DT_{\varepsilon_1}(T_{\varepsilon_1,l}^{-1} x)} - \frac{1}{DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1} x)}\right]$$

$$+ \sum_{l=1}^{m} \frac{1}{DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1} x)}\big[f \cdot \mathbf{1}_{U_n^c}\big(T_{\varepsilon_1,l}^{-1} x\big) - f \cdot \mathbf{1}_{U_n^c}\big(T_{\varepsilon_2,l}^{-1} x\big)\big] = E_2(x) + E_3(x).$$

In the formula above we considered, without restriction, the derivative positive and moreover we discarded those points x which have only one pre-image in each interval of monotonicity. After integration this will give an error (E_1) as $E_1 \le 4m|\varepsilon_1 - \varepsilon_2| \|\hat{P}_\varepsilon f\|_\infty$. But $\|\hat{P}_\varepsilon f\|_\infty \le \|f\|_\infty \sum_{l=1}^{m} \frac{DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1} x')}{DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1} x)} \frac{1}{DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1} x')}$, where x' is the point where $DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1} x')|A_l| \ge \eta$, being η the minimum length of $T(A_k)$, $k = 1, \dots, m$. But the first ratio in the previous sum is simply bounded by the distortion constant $D_c = \Lambda\lambda^{-1}$; therefore we get

$$E_1 \le 4m|\varepsilon_1 - \varepsilon_2| \|f\|_\infty \frac{D_c}{\eta} \sum_{l=1}^{m} |A_l| \le 4m|\varepsilon_1 - \varepsilon_2| \|f\|_\infty \frac{D_c}{\eta}.$$

We now bound E_2. The term in the square bracket and for given l (we drop this index in the derivatives in the next formulas), will be equal to $\frac{D^2 T(\xi)}{[DT(\xi)]^2} |T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)|$, being ξ a point in the interior of A_l. The first factor is uniformly bounded by C_1. Since $x = T_{\varepsilon_1}(T_{\varepsilon_1}^{-1}(x)) = T(T_{\varepsilon_1}^{-1}(x)) + \varepsilon_1 = T(T_{\varepsilon_2}^{-1}(x)) + \varepsilon_2 = T_{\varepsilon_2}(T_{\varepsilon_2}^{-1}(x))$, we have that $|T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)| = |\varepsilon_1 - \varepsilon_2| |DT(\xi')|^{-1}$, where ξ' is in A_l. Replacing ξ' by $T_{\varepsilon_1,l}^{-1} x$, because of distortion, we get

$$\int \big|E_2(x)\big|\, dx \le |\varepsilon_1 - \varepsilon_2| C_1 D_c \int \left[\sum_{l=1}^{m} \big|f\big(T_{\varepsilon_1,l}^{-1}\big)\big| \frac{1}{DT_{\varepsilon_1}(T_{\varepsilon_1,l}^{-1} x)}\right] dx$$

$$= |\varepsilon_1 - \varepsilon_2| C_1 D_c \int P_{\varepsilon_1}\big(|f|\big)(x)\, dx = |\varepsilon_1 - \varepsilon_2| C_1 D_c \|f\|_1.$$

To bound the last term we use the formula (3.11), in [9],

$$\int \sup_{|y-x| \le t} \big|f(y) - f(x)\big|\, dx \le 2t \operatorname{Var}(f),$$

by observing again that $|T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)| = |\varepsilon_1 - \varepsilon_2| |DT(\xi')|^{-1}$, where ξ' is in A_l. By integrating $E_3(x)$ we get

$$\int \big|E_3(x)\big|\, dx \le 2m\lambda^{-2} |\varepsilon_1 - \varepsilon_2| \operatorname{Var}(f \mathbf{1}_{U_n^c})$$

$$\le 10m\lambda^{-2} |\varepsilon_1 - \varepsilon_2| \operatorname{Var}(f).$$

Putting together the three errors we finally get that there exists a constant $\tilde{C}$ such that

$$\|\hat{P}_{\varepsilon_1} f - \hat{P}_{\varepsilon_2} f\|_1 \le \tilde{C} |\varepsilon_1 - \varepsilon_2| \|f\|_{\mathrm{BV}},$$

and we can complete the argument as in the first example of β transformations.

3.3.3. *Multidimensional maps*

We give here a multidimensional version of the maps considered in the preceding section; these maps were extensively investigated in [2,4,19,20,33] and we defer to those papers for more details. Let M be a compact subset of $\mathbb{R}^N$ which is the closure of its non-empty interior. We take a map $T : M \to M$ and let $\mathcal{A} = \{A_i\}_{i=1}^{m}$ be a finite family of disjoint open sets such that the Lebesgue measure of $M \setminus \bigcup_i A_i$ is zero, and there exist open sets $\tilde{A}_i \supset \overline{A_i}$ and $C^{1+\alpha}$ maps $T_i : \tilde{A}_i \to \mathbb{R}^N$, for some real number $0 < \alpha \le 1$ and some sufficiently small real number $\varepsilon_1 > 0$, such that

(1) $T_i(\tilde{A}_i) \supset B_{\varepsilon_1}(T(A_i))$ for each i, where $B_\varepsilon(V)$ denotes a neighbourhood of size ε of the set V. The maps T_i are the local extensions of T to the $\tilde{A}_i$.
(2) there exists a constant C_1 so that for each i and $x, y \in T(A_i)$ with $\operatorname{dist}(x, y) \le \varepsilon_1$,

$$\left|\det DT_i^{-1}(x) - \det DT_i^{-1}(y)\right| \le C_1 \left|\det DT_i^{-1}(x)\right| \operatorname{dist}(x, y)^\alpha;$$

(3) there exists $s = s(T) < 1$ such that $\forall x, y \in T(\tilde{A}_i)$ with $\operatorname{dist}(x, y) \le \varepsilon_1$, we have

$$\operatorname{dist}\left(T_i^{-1}x, T_i^{-1}y\right) \le s \operatorname{dist}(x, y);$$

(4) each ∂A_i is a codimension-one embedded compact piecewise C^1 submanifold and

$$s^\alpha + \frac{4s}{1-s} Z(T) \frac{\gamma_{N-1}}{\gamma_N} < 1, \tag{3.14}$$

where $Z(T) = \sup_x \sum_i \#\{\text{smooth pieces intersecting } \partial A_i \text{ containing } x\}$ and γ_N is the volume of the unit ball in $\mathbb{R}^N$.

Given such a map T, we define locally on each A_i the map $T_\varepsilon \in \mathcal{F}$ by $T_\varepsilon(x) := T(x) + \varepsilon$, where now ε is an n-dimensional vector with all the components of absolute value less than one. As in the previous example the translation by ε is allowed if the image $T_\varepsilon A_i$ remains in M: in this regard, we could play with the sign of the components of ε or not move the map at all. As in the one dimensional case, we shall also make the following assumption on $\mathcal{F}$. We assume that there exists a set $A_w \in \mathcal{A}$ satisfying:

(i) $A_w \subset T_\varepsilon A_k$ for all $\forall T_\varepsilon \in \mathcal{F}$ and for all $k = 1, \ldots, m$.
(ii) TA_w is the whole M, which in turn implies that there exists $1 \ge L' > 0$ such that $\forall k = 1, \ldots, q$ and $\forall T_\varepsilon \in \mathcal{F}$, $\operatorname{diameter}(T_\varepsilon(A_w) \cap A_k) > L'$.

As $\mathcal{V} \subset \mathcal{L}^1(m)$ we use the space of quasi-Hölder functions, for which we refer again to [20,33]. On this space, the transfer operator satisfies a Doeblin–Fortet–Lasota–Yorke inequality. Finally, Lipschitz continuity has been proved for additive noise in Proposition 4.3 in [4].

3.3.4. *Covering maps*: *A general class*

We now present a more general class of examples which were introduced in [5] to study metastability for randomly perturbed maps. As before, the family $\mathcal{F}$ will be constructed around a given map T which is again defined on the unit interval M. We therefore begin by introducing such map T.

(A1) There exists a partition $\mathcal{A} = \{A_i : i = 1, \ldots, m\}$ of M, which consists of pairwise disjoint intervals A_i. Let $\bar{A}_i := [c_{i,0}, c_{i+1,0}]$. We assume there exists $\delta > 0$ such that $T_{i,0} := T|_{(c_{i,0}, c_{i+1,0})}$ is C^2 and extends to a C^2 function $\bar{T}_{i,0}$ on a neighbourhood $[c_{i,0} - \delta, c_{i+1,0} + \delta]$ of $\bar{A}_i$;
(A2) There exists $\beta_0 < \frac{1}{2}$ so that $\inf_{x \in I \setminus \mathcal{C}_0} |T'(x)| \ge \beta_0^{-1}$, where $\mathcal{C}_0 = \{c_{i,0}\}_{i=1}^m$.

We note that Assumption (A2), more precisely the fact that β_0^{-1} is strictly bigger than 2 instead of 1, is sufficient to get the uniform Doeblin–Fortet–Lasota–Yorke inequality (3.17) below, as explained in Section 4.2 of [17]. We now construct the family $\mathcal{F}$ by choosing maps $T_\varepsilon \in \mathcal{F}$ close to $T_{\varepsilon=0} := T$ in the following way:

Each map $T_\varepsilon \in \mathcal{F}$ has m branches and there exists a partition of M into intervals $\{A_{i,\varepsilon}\}_{i=1}^m$, $A_{i,\varepsilon} \cap A_{j,\varepsilon} = \varnothing$ for $i \ne j$, $\bar{A}_{i,\varepsilon} := [c_{i,\varepsilon}, c_{i+1,\varepsilon}]$ such that

(i) for each i one has that $[c_{i,0} + \delta, c_{i+1,0} - \delta] \subset [c_{i,\varepsilon}, c_{i+1,\varepsilon}] \subset [c_{i,0} - \delta, c_{i+1,0} + \delta]$; whenever $c_{1,0} = 0$ or $c_{q+1,0} = 1$, we *do not move* them with δ. In this way, we have established a one-to-one correspondence between the unperturbed and the perturbed extreme points of A_i and $A_{i,\varepsilon}$. (The quantity δ is from Assumption (A1) above.)

(ii) the map T_ε is locally injective over the closed intervals $\overline{A_{i,\varepsilon}}$, of class C^2 in their interiors, and expanding with $\inf_x |T'_\varepsilon x| > 2$. Moreover there exists $\sigma > 0$ such that $\forall T_\varepsilon \in \mathcal{F}$, $\forall i = 1, \dots, m$ and $\forall x \in [c_{i,0} - \delta, c_{i+1,0} + \delta] \cap \overline{A_{i,\varepsilon}}$ where $c_{i,0}$ and $c_{i,\varepsilon}$ are two (left or right) *corresponding points*, we have:

$$|c_{i,0} - c_{i,\varepsilon}| \le \sigma \tag{3.15}$$

and

$$\left|\bar{T}_{i,0}(x) - T_{i,\varepsilon}(x)\right| \le \sigma. \tag{3.16}$$

Under these assumptions and by taking, with obvious notations, a concatenation of n transfer operators, we have the uniform Doeblin–Fortet–Lasota–Yorke inequality, namely there exist $\eta \in (0, 1)$ and $B < \infty$ such that, for all $f \in \mathrm{BV}$, all n and all concatenations of n maps of $\mathcal{F}$, we have

$$\|P_{\varepsilon_n} \circ \cdots \circ P_{\varepsilon_1} f\|_{\mathrm{BV}} \le \eta^n \|f\|_{\mathrm{BV}} + B\|f\|_1. \tag{3.17}$$

About the continuity (Lip): looking carefully at the proof of the continuity for the expanding map of the intervals, one sees that it extends to the actual case if one gets the following bounds:

$$\left.\begin{aligned} |T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)| \\ |DT_{\varepsilon_1}(x) - DT_{\varepsilon_2}(x)| \end{aligned}\right\} = O\big(|\varepsilon_1 - \varepsilon_2|\big), \tag{3.18}$$

where the point x is in the same domain of injectivity of the maps T_{ε_1} and T_{ε_2}, the comparison of the *same* functions and derivative in two *different* points being controlled by the condition (3.15). The bounds (3.18) follow easily by adding to (3.15), (3.16) the further assumptions that $\sigma = O(\varepsilon)$ and requiring a continuity condition for derivatives like (3.16) and with σ again being of order ε.

4. EVT for the sequential systems: An example of uniformly expanding map

In this section, we will give a detailed analysis of the application of the general result obtained in Section 2 to a particular sequential system. It is constructed with β transformations; similar approach and technique can be used to treat the other examples of sequential systems introduced above with suitable adaptations and modifications. We point out that in this example we will take $u_{n,i} = u_n$, where $(u_n)_{n\in\mathbb{N}}$ satisfies $n\mu(U_n) = n\mu(X_0 > u_n) \to \tau$, as $n \to \infty$ for some $\tau > 0$, where μ is the invariant measure of the original map T_β.

Consider the family of maps on the unit circle $S^1 = [0, 1]$, with the identification $0 \sim 1$, given by $T_\beta(x) = \beta x \bmod 1$ for $\beta > 1 + c$, with $c > 0$. Note that for many such β, we have that $T_\beta(1) \ne 1$ and, by the identification $0 \sim 1$, this means that T_β as a map on S^1 is not continuous at $\zeta = 0 \sim 1$. For simplicity we assume that $T_\beta(0) = 0$ but consider that the orbit of 1 is still defined to be $T_\beta(1), T_\beta^2(1), \dots$ although, strictly speaking, $1 \sim 0$ should be considered a fixed point. In what follows m denotes Lebesgue measure on $[0, 1]$.

Theorem 4.1. *Consider an unperturbed map T_β corresponding to some $\beta = \beta_0 > 1 + c$, with invariant absolutely continuous probability $\mu = \mu_\beta$. Consider a sequential system acting on the unit circle and given by $\mathcal{T}_n = T_n \circ \cdots \circ T_1$, where $T_i = T_{\beta_{i-1}}$, for all $i = 1, \dots, n$ and $|\beta_n - \beta| \le n^{-\xi}$ holds for some $\xi > 1$. Let $X_1, X_2, \dots$ be defined by (3.8), where the observable function φ, given by (3.9), achieves a global maximum at a chosen $\zeta \in [0, 1]$. Let $(u_n)_{n\in\mathbb{N}}$ be such that $n\mu(X_0 > u_n) \to \tau$, as $n \to \infty$ for some $\tau \ge 0$. Then, there exists $0 < \theta \le 1$ such that*

$$\lim_{n\to\infty} m(X_0 \le u_n, X_1 \le u_n, \dots, X_{n-1} \le u_n) = \mathrm{e}^{-\theta\tau}.$$

The value of θ is determined by the behaviour of ζ under the original dynamics T_β, namely,

- *If the orbit of ζ by T_β never hits $0 \sim 1$ and ζ is periodic of prime period p*[2] *then $\theta = 1 - \beta^{-p}$;*
- *If the orbit of ζ by T_β never hits $0 \sim 1$ and ζ is not periodic then $\theta = 1$;*

[2] $T_\beta^p(\zeta) = \zeta$ and p is the minimum integer with such property.

- *If* $\zeta = 0 \sim 1$ *and* 1 *is not periodic*,[3] *then* $\theta = \frac{d\mu}{dm}(0)(1-\beta^{-1}) + \frac{d\mu}{dm}(1)$;
- *If* $\zeta = 0 \sim 1$ *and* 1 *is periodic of prime period* p *then* $\theta = \frac{d\mu}{dm}(0)(1-\beta^{-1}) + \frac{d\mu}{dm}(1)(1-\beta^{-p})$.

We remark that if the decay rate of $|\beta_n - \beta|$ is slower than in the statement of the theorem then the observed extremal index for the sequential system at periodic points of the original dynamics may be 1 as shown in Section 4.5.

4.1. *Preliminaries*

As we said above, we let μ denote the invariant measure of the original map T_β and let $h = \frac{d\mu}{dm}$ be its density.

We assume throughout this subsection that there exists $\xi > 1$ such that

$$|\beta_n - \beta| \le \frac{1}{n^\xi}. \tag{4.1}$$

Also let $0 < \gamma < 1$ be such that $\gamma\xi > 1$. In what follows P denotes the transfer operator associated to the unperturbed map T_β. Recall that $\Pi_i = P_i \circ \cdots \circ P_1$, where P_i is the transfer operator associated to $T_i = T_{\beta_i}$, while P^i is the corresponding concatenation for the unperturbed map T_β. Note that by [9, Lemma 3.10], we have

$$\left\| \Pi_i(g) - \int g\, dmh \right\|_1 \le C_1 \frac{\log i}{i^\xi} \|g\|_{\mathrm{BV}}. \tag{4.2}$$

Consider a measurable set $A \subset [0,1]$. Then

$$\begin{aligned} m\big(\mathcal{T}_j^{-1}(A)\big) &= \int \mathbf{1}_A \circ T_j \circ \cdots \circ T_1\, dm = \int \mathbf{1}_A \Pi_j(1)\, dm \\ &= \int \mathbf{1}_A h\, dm + \int \mathbf{1}_A\big(\Pi_j(1) - h\big)\, dm. \end{aligned}$$

By (4.2), if $j \ge n^\gamma$ (recall that $\gamma\xi > 1$) then we have $\int |\Pi_j(1) - h|\, dm \le C_1 \frac{\log i}{i^\xi} = o(n^{-1})$, which allows us to write:

$$m\big(\mathcal{T}_j^{-1}(A)\big) = \mu(A) + o\big(n^{-1}\big). \tag{4.3}$$

4.1.1. *Verification of condition* (2.2), *i.e.*, $\lim_{n\to\infty} \sum_{i=0}^{n-1} m(X_i > u_n) = \tau$

We start with the following lemma.

Lemma 4.2. *We have that*

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} \int_{U_n} P^i(1)\, dm = \tau.$$

Proof. By hypothesis, for all $j \in \mathbb{N}$ and $g \in \mathrm{BV}$ we have $P^j(g) = h\int g \cdot h\, dm + Q^j(g)$, where $\|Q^j(g)\|_\infty \le \alpha^j \|g\|_{\mathrm{BV}}$, for some $\alpha < 1$. Then we can write:

$$\begin{aligned} \sum_{i=0}^{n-1} \int_{U_n} P^i(1)\, dm &= \sum_{i=0}^{n-1} \int h\left(\int 1 \cdot h\, dm\right) \mathbf{1}_{U_n}\, dm + \sum_{i=0}^{n-1} \int Q^i(1)\mathbf{1}_{U_n}\, dm \\ &= \sum_{i=0}^{n-1} \int_{U_n} h\, dm + \sum_{i=0}^{n-1} \int Q^i(1)\mathbf{1}_{U_n}\, dm \\ &= n\mu(U_n) + \sum_{i=0}^{n-1} \int Q^i(1)\mathbf{1}_{U_n}\, dm. \end{aligned}$$

[3] $T_\beta^n(1) \ne 0 \sim 1$ for all n.

The result follows if we show that the second term on the r.h.s. goes to 0, as $n \to \infty$. This follows easily since

$$\sum_{i=0}^{n-1} \int Q^i(1)\mathbf{1}_{U_n}\,dm \le \sum_{i=0}^{n-1} \alpha^i \int \mathbf{1}_{U_n}\,dm = \frac{1-\alpha^n}{1-\alpha} m(U_n) \underset{n\to\infty}{\longrightarrow} 0.$$

□

Since

$$\sum_{i=0}^{n-1} m(X_i > u_n) = \sum_{i=0}^{n-1} \int_{U_n} \Pi_i(1)\,dm = \sum_{i=0}^{n-1} \int_{U_n} P^i(1)\,dm + \sum_{i=0}^{n-1} \int_{U_n} \Pi_i(1) - P^i(1)\,dm,$$

then condition (2.2) holds if we prove that the second term on the r.h.s. goes to 0 as $n \to \infty$.

Let $\varepsilon > 0$ be arbitrary. Now, since $\xi > 1$ then $\sum_{i\ge 0} \frac{\log i}{i^\xi} < \infty$, so there exists $N \in \mathbb{N}$ such that $C_0 \sum_{i\ge N} \frac{\log i}{i^\xi} < \varepsilon/2$.

On the other hand, using the Lasota–Yorke inequalities for both Π and P, we have that there exists some $C > 0$ such that $|\Pi_i(1) - P^i(1)| \le C$, for all $i \in \mathbb{N}$. Let n be sufficiently large so that $CNm(U_n) < \varepsilon/2$. Then

$$\begin{aligned}\sum_{i=0}^{n-1} \int_{U_n} \Pi_i(1) - P^i(1)\,dm &= \sum_{i=0}^{N-1} \int_{U_n} \Pi_i(1) - P^i(1)\,dm + \sum_{i=N}^{\infty} \int_{U_n} \Pi_i(1) - P^i(1)\,dm \\ &\le CNm(U_n) + C_0 \sum_{i\ge N} \frac{\log i}{i^\xi} < \varepsilon/2 + \varepsilon/2 = \varepsilon.\end{aligned}$$

4.2. *Verification of* $Д_q(u_n)$

We start by proving the following statement about decay of correlations, which is just a slightly more general statement then the one proved in [9, Section 3].

Proposition 4.3. *Let $\phi \in \mathrm{BV}$ and $\psi \in L^1(m)$. Then for the β transformations $T_n = T_{\beta_n}$ we have that*

$$\left| \int \phi \circ \mathcal{T}_i \psi \circ \mathcal{T}_{i+t}\,dm - \int \phi \circ \mathcal{T}_i\,dm \int \psi \circ \mathcal{T}_{i+t}\,dm \right| \le B\lambda^t \|\phi\|_{\mathrm{BV}} \|\psi\|_1,$$

for some $\lambda < 1$ and $B > 0$ independent of ϕ and ψ.

Remark 4.4. *Note that as it can be seen in [9, Section 3], Proposition 4.3 holds for any sequence $T_{\beta_1}, T_{\beta_2}, \ldots$ of β transformations and not necessarily only for the ones that satisfy condition (4.1).*

Proof of Proposition 4.3. Using the adjoint property, write

$$\begin{aligned} DC(\phi,\psi,i,t) &:= \int \phi \circ \mathcal{T}_i \psi \circ \mathcal{T}_{i+t}\,dm - \int \phi \circ \mathcal{T}_i\,dm \int \psi \circ \mathcal{T}_{i+t}\,dm \\ &= \int \psi P_{i+t} \cdots P_{i+1}\big(\phi \Pi_i(1)\big)\,dm - \int \phi \Pi_i(1)\,dm \int \psi \Pi_{i+t}(1)\,dm. \end{aligned}$$

Using the fact that the Perron–Fröbenius operators preserve integrals we have

$$\int \phi \Pi_i(1)\,dm \int \psi \Pi_{i+t}(1)\,dm = \int\int \psi \Pi_{i+t}(1)\,dm P_{i+t} \cdots P_{i+1}\big(\phi \Pi_i(1)\big)\,dm.$$

By linearity we also have

$$\int \phi \Pi_i(1)\,dm \int \psi \Pi_{i+t}(1)\,dm = \int \psi P_{i+t} \cdots P_{i+1}\left(\int \phi \Pi_i(1)\,dm \Pi_i(1)\right)dm.$$

Again linearity and preservation of the integrals allow us to write:

$$\int \phi \Pi_i(1)\,dm \int \psi \Pi_{i+t}(1)\,dm = \iint \psi \Pi_{i+t}(1)\,dm\, P_{i+t}\cdots P_{i+1}\left(\int \phi \Pi_i(1)\,dm\, \Pi_i(1)\right) dm.$$

Consequently we have

$$\begin{aligned} DC(\phi,\psi,i,t) &= \int \psi P_{i+t}\cdots P_{i+1}\big(\phi \Pi_i(1)\big)\,dm - \iint \psi \Pi_{i+t}(1)\,dm\, P_{i+t}\cdots P_{i+1}\big(\phi\Pi_i(1)\big)\,dm \\ &\quad - \int \psi P_{i+t}\cdots P_{i+1}\left(\int \phi \Pi_i(1)\,dm\,\Pi_i(1)\right) dm \\ &\quad + \iint \psi \Pi_{i+t}(1)\,dm\, P_{i+t}\cdots P_{i+1}\left(\int \phi\Pi_i(1)\,dm\,\Pi_i(1)\right) dm \\ &= \int \left(\psi - \int \psi \Pi_{i+t}(1)\,dm\right) P_{i+t}\cdots P_{i+1}\left(\Pi_i(1)\left(\phi - \int \phi \Pi_i(1)\,dm\right)\right). \end{aligned}$$

Let $\tilde\phi = \phi - \int \phi \Pi_i(1)\,dm$. Observe that $\int \Pi_i(1)\tilde\phi\,dm = 0$. This means that the observable function $\Pi_i(1)\tilde\phi \in \mathcal{V}_0$, where $\mathcal{V}_0$ is the set of functions with 0 integral that was defined in [9, Lemma 2.12]. Moreover, by (DFLY), there exists a constant C_0 independent of ϕ and ψ such that $\|\Pi_i(1)\tilde\phi\|_{\mathrm{BV}} \le 3C_0\|\phi\|_{\mathrm{BV}}$.

As it has been shown in [9, Section 3], condition (Dec) of the same paper is satisfied for any sequence of β transformations as considered here. It follows that for all $g \in \mathcal{V}_0$ and $i \in \mathbb{N}$ we have that $\|P_{i+t}\cdots P_{i+1}(g)\|_{\mathrm{BV}} \le K\lambda^t\|g\|_{\mathrm{BV}}$, for some $K>0$ and $\lambda<1$ independent of g, which applied to $\Pi_i(1)\tilde\phi$ gives:

$$\big\|P_{i+t}\cdots P_{i+1}\big(\Pi_i(1)\tilde\phi\big)\big\|_{\mathrm{BV}} \le 3KC_0\lambda^t\|\phi\|_{\mathrm{BV}}. \tag{4.4}$$

Let $\tilde\psi = \psi - \int \psi\Pi_{i+t}(1)\,dm$. Again, by [9, (2.4)], we have $\|\tilde\psi\|_1 \le 2C_0\|\psi\|_1$. Hence, using (4.4) we obtain

$$\begin{aligned} \big|DC(\phi,\psi,i,t)\big| &= \left|\int \tilde\psi P_{i+t}\cdots P_{i+1}\big(\Pi_i(1)\tilde\phi\big)\,dm\right| \\ &\le \big\|P_{i+t}\cdots P_{i+1}\big(\Pi_i(1)\tilde\phi\big)\big\|_{\mathrm{BV}}\int|\tilde\psi|\,dm \\ &\le 6KC_0^2\lambda^t\|\phi\|_{\mathrm{BV}}\|\psi\|_1. \end{aligned}$$

□

Condition $Д_q(u_{n,i})$ follows from Proposition 4.3 by taking for each $i\in\mathbb{N}$,

$$\phi_i = \mathbf{1}_{D^{(q)}_{n,i}} \quad\text{and}\quad \psi_i = \mathbf{1}_{D^{(q)}_{n,i+t}}\cdot \mathbf{1}_{D^{(q)}_{n,i+t+1}}\circ T_{i+t+1}\cdot\dots\cdot \mathbf{1}_{D^{(q)}_{n,i+t+\ell}}\circ T_{i+t+\ell}\circ\dots\circ T_{i+t+1},$$

where for every $j\in\mathbb{N}$ we define

$$D^{(q)}_{n,j} = U_n \cap T^{-1}_{j+1}\big(U_n^c\big)\cap\dots\cap T^{-1}_{j+q}\big(U_n^c\big). \tag{4.5}$$

Since we assume that (4.1) holds, there exists a constant $C>0$ depending on q but not on i such that $\|\phi_i\|_{\mathrm{BV}} < C$. Moreover, it is clear that $\|\psi_i\|\le 1$. Hence,

$$\begin{aligned} &\big|\mathbb{P}\big(A^{(q)}{}_{n,i}\cap \mathcal{W}_{i+t,\ell}\big(\mathbb{A}_n^{(q)}\big)\big) - \mathbb{P}\big(A^{(q)}{}_{n,i}\big)\mathbb{P}\big(\mathcal{W}_{i+t,\ell}\big(\mathbb{A}_n^{(q)}\big)\big)\big| \\ &\quad = \left|\int \phi_i\circ\mathcal{T}_i\psi_i\circ\mathcal{T}_{i+t}\,dm - \int\phi_i\circ\mathcal{T}_i\,dm\int\psi_i\circ\mathcal{T}_{i+t}\,dm\right| \le \mathrm{const}\,\lambda^t. \end{aligned}$$

Thus, if we take $\gamma_i(q,n,t) = \mathrm{const}\,\lambda^t$ and $t_n = (\log n)^2$ condition $Д_q(u_{n,i})$ is trivially satisfied.

4.3. *Verification of condition* $Д'_q(u_n)$

We start by noting that we may neglect the first n^γ random variables of the process $X_0, X_1, \dots,$ where γ is such that $\gamma\xi > 1$, for ξ given as in (4.1).

In fact, by Lemma 2.5 and (DFLY) we have

$$m\big(\max\{X_{n^\gamma}, \dots, X_{n-1}\} \le u_n\big) - m(M_n \le u_n) \le \sum_{i=0}^{n^\gamma-1} m(X_i > u_n) = \sum_{i=0}^{n^\gamma-1} \int \mathbf{1}_{U_n} \Pi_i(1)\, dm$$
$$\le C_0 n^\gamma m(U_n) \underset{n\to\infty}{\longrightarrow} 0.$$

This way, we simply disregard the n^γ random variables of $X_0, X_1, \dots$ and start the blocking procedure, described in Section 2.2, in X_{n^γ} by taking $\mathcal{L}_0 = n^\gamma$. We split the remaining $n - n^\gamma$ random variables into k_n blocks as described in Section 2.2. Our goal is to show that

$$S'_n := \sum_{i=1}^{k_n} \sum_{j=0}^{\ell_i-1} \sum_{r>j}^{\ell_i-1} m\big(A^{(q)}{}_{\mathcal{L}_{i-1}+j} \cap A^{(q)}{}_{\mathcal{L}_{i-1}+r}\big)$$

goes to 0.

We define for some $i, n, q \in \mathbb{N}_0$,

$$R^{(q)}_{n,i} := \min\big\{j > i : \mathbf{1}_{A^{(q)}_i} \cdot \mathbf{1}_{A^{(q)}_j}(x) > 0 \text{ for some } x \in [0,1]\big\},$$
$$\tilde{R}^{(q)}_n := \tilde{R}^{(q)}_n\big(n^\gamma\big) = \min\big\{R^{(q)}_{n,i}, i = n^\gamma, \dots, n\big\},$$
$$L_n = \max\{\ell_{n,i}, i = 1, \dots, k_n\}.$$

We have

$$S'_n \le \sum_{i=n^\gamma}^{n} \sum_{j>i+R^{(q)}_{n,i}}^{L_n} m\big(A^{(q)}_i \cap A^{(q)}_j\big) = \sum_{i=n^\gamma}^{n} \sum_{j>i+R^{(q)}_{n,i}}^{L_n} \int \mathbf{1}_{D^{(q)}_{n,i}} \circ \mathcal{T}_i \cdot \mathbf{1}_{D^{(q)}_{n,j}} \circ \mathcal{T}_j\, dm,$$

where $D^{(q)}_{n,i}$ and $D^{(q)}_{n,j}$ are given as in (4.5). Using Proposition 4.3, with $\phi = \mathbf{1}_{D^{(q)}_{n,i}}$ and $\psi = \mathbf{1}_{D^{(q)}_{n,j}}$ and the adjoint property of the operators, it follows that

$$\int \mathbf{1}_{D^{(q)}_{n,i}} \circ \mathcal{T}_i \cdot \mathbf{1}_{D^{(q)}_{n,j}} \circ \mathcal{T}_j\, dm \le \int \mathbf{1}_{D^{(q)}_{n,i}} \Pi_i(1)\, dm \int \mathbf{1}_{D^{(q)}_{n,j}} \Pi_j(1)\, dm + B\lambda^{j-i} \|\mathbf{1}_{D^{(q)}_{n,i}}\|_{\mathrm{BV}} \|\mathbf{1}_{D^{(q)}_{n,j}}\|_1.$$

Using (DFLY) and since there exists some $C_2 > 0$ (independent of n) such that $\|\mathbf{1}_{D^{(q)}_{n,i}}\|_{\mathrm{BV}} \le C_2$, we have

$$\int \mathbf{1}_{D^{(q)}_{n,i}} \circ \mathcal{T}_i \cdot \mathbf{1}_{D^{(q)}_{n,j}} \circ \mathcal{T}_j\, dm \le C_0^2 m(U_n)^2 + BC_2\lambda^{j-1} m(U_n).$$

Hence,

$$S'_n \le \sum_{i=n^\gamma}^{n} \sum_{j\ge i+R^{(q)}_{n,i}}^{L_n} \big(C_0^2 m(U_n)^2 + BC_2\lambda^{j-1} m(U_n)\big) \le C_0^2 n L_n m(U_n)^2 + BC_2 m(U_n) n \sum_{k\ge \tilde{R}^{(q)}_n}^{L_n} \lambda^k$$
$$\le C_0^2 n L_n m(U_n)^2 + BC_2 m(U_n) n \lambda^{\tilde{R}^{(q)}_n} \frac{1}{1-\lambda}.$$

Now we show that

$$L_n = \frac{n}{k_n}\big(1+o(1)\big). \tag{4.6}$$

To see this, observe that each ℓ_{n_i} is defined, in this case, by the largest integer ℓ_n such that $\sum_{j=s}^{s+\ell_n-1} m(X_j > U_n) \le \frac{1}{k_n}\sum_{j=n^\gamma}^{n-1} m(X_j > u_n)$. Using (4.3), it follows that $\ell_n\mu(U_n)(1+o(1)) \le \frac{n-n^\gamma}{k_n}\mu(U_n)(1+o(1))$. On the other hand, by definition of ℓ_n we must have $\sum_{j=s}^{s+\ell_n-1} m(X_j > U_n) > \frac{1}{k_n}\sum_{j=n^\gamma}^{n-1} m(X_j > u_n) - m(X_{s+\ell_n} > u_n)$. Using (4.3) again, we have $\ell_n\mu(U_n)(1+o(1)) > \frac{n-n^\gamma}{k_n}\mu(U_n)(1+o(1)) - \mu(U_n)(1+o(1))$. Together with the previous inequality, (4.6) follows at once.

Using estimate (4.6), the fact that $\lim_{n\to\infty} n\mu(U_n) = \tau$ and $h \in \mathrm{BV}$, we have that there exists some positive constant C such that

$$S_n' \le C\left(\frac{1}{k_n} + \lambda^{\tilde{R}_n^{(q)}}\right).$$

In order to prove that $⨿_q'(u_n)$ holds, we need to show that $\tilde{R}_n^{(q)} \to \infty$, as $n \to \infty$, for all $q \in \mathbb{N}_0$. To do that we have to split the proof in several cases. First, we have to consider the cases when the orbit of ζ hits 1 or not. Then for each of the previous two cases, we have to consider if ζ is periodic or not.

We will consider that the maps T_i, for all $i \in \mathbb{N}_0$, are defined in S^1 by using the usual identification $0 \sim 1$. Observe that the only point of discontinuity of such maps is $0 \sim 1$. Moreover, $\lim_{x\to 0^+} T_i(x) = 0$ and $\lim_{x\to 1^-} T_i(x) = \beta_i - \lfloor\beta_i\rfloor$.

4.3.1. *The orbit of ζ by the unperturbed T_β map does not hit* 1
We mean that for all $j \in \mathbb{N}_0$ we have $T^j(\zeta) \ne 1$.

4.3.1.1. *The orbit of ζ is not periodic*. In this case, for all $j \in \mathbb{N}$, we have that $T^j(\zeta) \ne \zeta$, we take $q = 0$ and in particular $D_{n,i}^{(q)} = U_n$, for all $i \in \mathbb{N}_0$. Let $J \in \mathbb{N}$.

We will check that for n sufficiently large $\tilde{R}_n^{(q)} > J$. Since ζ is not periodic, there exists some $\varepsilon > 0$ such that $\min_{j=1,\dots,J} \mathrm{dist}(T^j(\zeta),\zeta) > \varepsilon$. Let $N_1 \in \mathbb{N}$ be sufficiently large so that for all $i \ge N_1$, we have

$$\min_{j=1,\dots,J} \mathrm{dist}\big(T_{i+j}\circ\cdots\circ T_i(\zeta), T^j(\zeta)\big) < \varepsilon/4.$$

Let $N_2 \in \mathbb{N}$ be sufficiently large so that for all $i \ge N_2$ we have

$$\mathrm{diam}\big(T_{i+J}\circ\cdots\circ T_i(U_n)\big) < \varepsilon/4.$$

This way for all $i \ge \max\{N_1, N_2\}$, for all $x \in U_n$ and for all $j \le J$ we have

$$\mathrm{dist}\big(T_{i+j}\circ\cdots\circ T_i(x), \zeta\big) > \varepsilon/2.$$

Hence, as long as $n^\gamma > \max\{N_1, N_2\}$ we have $\tilde{R}_n^{(q)} > J$.

Note that for this argument to work we only need that $\beta_n \to \beta$ and the stronger restriction imposed by (4.1) is not necessary.

4.3.1.2. *The orbit of ζ is periodic*. In this case, there exists $p \in \mathbb{N}$, such that $T^j(\zeta) \ne \zeta$ for all $j < p$ and $T^p(\zeta) = \zeta$. We take $q = p$.

Let

$$\varepsilon_n := |\beta_{n^\gamma} - \beta|. \tag{4.7}$$

By (4.1) and choice of γ, we have that $\varepsilon_n = o(n^{-1})$. Also let $\delta > 0$, be such that $B_\delta(\zeta)$ is contained on a domain of injectivity of all T_i, with $i \ge n^\gamma$.

Let $J \in \mathbb{N}$ be chosen. Using a continuity argument, we can show that there exists $C := C(J, p) > 0$ such that

$$\operatorname{dist}\big(T_{i+j} \circ \cdots \circ T_{i+1}(\zeta), T^j(\zeta)\big) < C\varepsilon_n, \quad \text{for all } i = 1, \dots, J$$

and moreover $U_n \cap T_{i+j} \circ \cdots \circ T_{i+1}(U_n) = \emptyset$, for all $j \le J$ such that $j/p - \lfloor j/p \rfloor > 0$.

We want to check that if $x \in A_i^{(q)}$ for some $i \ge n^\gamma$, i.e., $\mathcal{T}_i(x) \in D_{n,i}^{(q)}$, then $x \notin A_{i+j}^{(q)}$, for all $j = 1, \dots, J$, i.e., $\mathcal{T}_{i+j}(x) \notin D_{n,i+j}^{(q)} \subset U_n$, for all such j. By the assumptions above, we only need to check the latter for all $j = 1, \dots, J$ such that $j/p - \lfloor j/p \rfloor = 0$, i.e., for all $j = sp$, where $s = 1, \dots, \lfloor J/p \rfloor$.

By definition of $A_i^{(q)}$ the statement is clearly true when $s = 1$. Let us consider now that $s > 1$ and let $x \in A_i^{(q)}$. We may write

$$\operatorname{dist}\big(\mathcal{T}_{i+sp}(x), T_{i+sp} \circ \cdots \circ T_{i+p+1}(\zeta)\big) > (\beta - \varepsilon_n)^{(s-1)p} \operatorname{dist}\big(\mathcal{T}_{i+p}(x), \zeta\big).$$

On the other hand,

$$\operatorname{dist}\big(T_{i+sp} \circ \cdots \circ T_{i+p+1}(\zeta), \zeta\big) \le C\varepsilon_n.$$

Hence,

$$\begin{aligned}
\operatorname{dist}\big(\mathcal{T}_{i+sp}(x), \zeta\big) &\ge \operatorname{dist}\big(\mathcal{T}_{i+sp}(x), T_{i+sp} \circ \cdots \circ T_{i+p+1}(\zeta)\big) - \operatorname{dist}\big(T_{i+sp} \circ \cdots \circ T_{i+p+1}(\zeta), \zeta\big) \\
&\ge (\beta - \varepsilon_n)^{(s-1)p} \operatorname{dist}\big(\mathcal{T}_{i+p}(x), \zeta\big) - C\varepsilon_n \\
&\ge (\beta - \varepsilon_n)^{(s-1)p} \frac{m(U_n)}{2} - C\varepsilon_n, \quad \text{since } x \in A_i^{(q)} \Rightarrow \mathcal{T}_{i+p}(x) \notin U_n \\
&> \frac{m(U_n)}{2}, \quad \text{for } n \text{ sufficiently large, since } \varepsilon_n = o\big(n^{-1}\big).
\end{aligned}$$

This shows that $\mathcal{T}_{sp+i}(x) \notin U_n$, which means that $\mathcal{T}_{sp+i}(x) \notin D_{n,i}^{(q)}$ and hence $x \notin A_{i+sp}^{(q)}$.

4.3.2. $\zeta = 0 \sim 1$

In this case we proceed in the same way as in [4, Section 3.3], which basically corresponds considering two versions of the same point: $\zeta^+ = 0$ and $\zeta^- = 1$. Note that ζ^+ is a fixed point for all maps considered and ζ^- may or not be periodic. So we split again into two cases.

4.3.2.1. 1 *is not periodic*. This means that $T^i(1) \neq \zeta$ for all $i \in \mathbb{N}$. Note that U_n can be divided into U_n^+ which corresponds to the bit having 0 at its left border and U_n^- which corresponds to the interval with 1 as its endpoint. In this case, $q = 1$ and $D_{n,i}^{(1)}$ has two connected components one of them being U_n^-. Let $J \in \mathbb{N}$ be fixed as before. A continuity argument as the one used in Section 4.3.1.1, allows us to show that the points of U_n^- do not return before J iterates. An argument similar to the one used in Section 4.3.1.2 would allow us to show also that the points of the other connected component of $D_{n,i}^{(1)}$ do not return to U_n before time J, also.

4.3.2.2. 1 *is periodic*. This means that there exists $p \in \mathbb{N}$ such that $T^i(1) \neq \zeta$ for all $i < p$ and $T^p(1) = \zeta$. In this case, we need to take $q = p$ and observe that $D_{n,i}^{(q)}$ has again two connected components, one to the right of 0 and the other to the left of 1, where none of the two points belongs to the set. The argument follows similarly as in the previous paragraph, except that this time both sides require mimicking the argument used in Section 4.3.1.2. Note that, the maps are orientation preserving so there is no switching as described in [4, Section 3.3].

4.4. *Verification of condition* (2.8)

We only need to verify (2.8), when ζ has some sort of periodic behaviour. Let ε_n be defined as in (4.7). Let δ_n be such that $U_n = B_{\delta_n}(\zeta)$. For simplicity, we assume that we are using the usual Riemannian metric so that we have a symmetry of the balls, which means that $|U_n| = m(U_n) = 2\delta_n$.

Let us assume first that ζ is a periodic point of prime period p with respect to the unperturbed map $T = T_\beta$ and the orbit of ζ does not hit $0 \sim 1$. In this case, we take $q = p$, $\theta = 1 - \beta^{-p}$ and check (2.8).

Using a continuity argument we can show that there exists $C := C(J, p) > 0$ such that

$$\mathrm{dist}\big(T_{i+p} \circ \cdots \circ T_{i+1}(\zeta), \zeta\big) < C\varepsilon_n.$$

We define two points ξ_u and ξ_l of $B_{\delta_n}(\zeta)$ on the same side with respect to ζ such that $\mathrm{dist}(\xi_u, \zeta) = (\beta - \varepsilon_n)^{-p}\delta_n + C\varepsilon_n$ and $\mathrm{dist}(\xi_l, \zeta) = (\beta + \varepsilon_n)^{-p}\delta_n - (\beta + \varepsilon_n)^{-p}C\varepsilon_n$. Recall that for all $i \geq n^\gamma$, we have that $(\beta - \varepsilon_n) \leq \beta_i \cdot \cdots \cdot \beta_{i+p} \leq (\beta + \varepsilon_n)$.

Since we are composing β transformations, then for all $i \geq n^\gamma$, we have $\mathrm{dist}(T_{i+p} \circ \cdots \circ T_i(\xi_u), T_{i+p} \circ \cdots \circ T_i(\zeta)) \geq \delta_n + (\beta - \varepsilon_n)^p C\varepsilon_n$. Using the triangle inequality it follows that

$$\mathrm{dist}\big(T_{i+p} \circ \cdots \circ T_i(\xi_u), \zeta\big) \geq \delta_n.$$

Similarly, $\mathrm{dist}(T_{i+p} \circ \cdots \circ T_i(\xi_l), T_{i+p} \circ \cdots \circ T_i(\zeta)) \leq \delta_n - C\varepsilon_n$ and

$$\mathrm{dist}\big(T_{i+p} \circ \cdots \circ T_i(\xi_l), \zeta\big) \leq \delta_n.$$

If we assume that both ξ_u and ξ_l are on the right hand side with respect to ζ and ξ_u^* and ξ_l^* are the corresponding points on the left hand side of ζ, then

$$\big(\zeta - \delta_n, \xi_u^*\big] \cup [\xi_u, \zeta + \delta_n) \subset D_{n,i}^{(p)} \subset \big(\zeta - \delta_n, \xi_l^*\big] \cup [\xi_l, \zeta + \delta_n).$$

Hence,

$$\delta_n - (\beta - \varepsilon_n)^{-p}\delta_n - C\varepsilon_n \leq \frac{1}{2} m\big(D_{n,i}^{(p)}\big) \leq \delta_n - (\beta + \varepsilon_n)^{-p}\delta_n + (\beta + \varepsilon_n)^{-p}C\varepsilon_n.$$

Since $\varepsilon_n = o(n^{-1}) = o(\delta_n)$ then we easily get that

$$\lim_{n\to\infty} \frac{m(D_{n,i}^{(p)})}{m(U_n)} = 1 - \beta^{-p}.$$

Now, observe that by (4.3), $m(A_{n,i}^{(p)}) = m(\mathcal{T}_i^{-1}(D_{n,i}^{(p)})) = \mu(D_{n,i}^{(p)}) + o(n^{-1})$ and $m(X_i > u_n) = \mu(U_n) + o(n^{-1})$. Hence, we have that

$$\lim_{n\to\infty} \frac{m(A_{n,i}^{(p)})}{m(X_i > u_n)} = \lim_{n\to\infty} \frac{\mu(D_{n,i}^{(p)})}{\mu(U_n)}.$$

The density $\frac{d\mu}{dm}$, which can be found in [29, Theorem 2], is sufficiently regular so that, as in [15, Section 7.3], one can see that

$$\lim_{n\to\infty} \frac{\mu(D_{n,i}^{(p)})}{\mu(U_n)} = \lim_{n\to\infty} \frac{m(D_{n,i}^{(p)})}{m(U_n)}.$$

It follows that

$$\lim_{n\to\infty} \frac{m(A_{n,i}^{(p)})}{m(X_i > u_n)} = 1 - \beta^{-p}.$$

Since, as we have seen in (4.6), we can write that $\ell_{n,i} = \frac{n}{k_n}(1 + o(1))$, then the previous equation can easily be used to prove that condition (2.8) holds, with $\theta = 1 - \beta^{-p}$.

In the case $\zeta = 0 \sim 1$, the argument follows similarly but this time we have to take into account the fact that the density is discontinuous at $0 \sim 1$. By [29] we have that

$$\frac{d\mu}{dm}(x) = \frac{1}{M(\beta)} \sum_{x<T^n(1)} \frac{1}{\beta^n},$$

where $M(\beta) := \int_0^1 \sum_{x<T^n(1)} \frac{1}{\beta^n}\, dm$. In this case, we have $\theta = \frac{d\mu}{dm}(0)(1-\beta^{-1}) + \frac{d\mu}{dm}(1)$ if 1 is not periodic and $\theta = \frac{d\mu}{dm}(0)(1-\beta^{-1}) + \frac{d\mu}{dm}(1)(1-\beta^{-p})$ if 1 is periodic of period p.

4.5. *An example with an EI equal to* 1 *at periodic points*

In the previous subsections, we used (4.1), which imposes a fast accumulation rate of β_n to β, to show that the EI equals the EI observed for the unperturbed dynamics. If this condition fails then the EI for the sequential dynamics does not need to be the same as the one of the original system.

Let $\beta = 5/2$ and $T = T_\beta = 5/2x \bmod 1$. Let $\zeta = 2/3$. Note that $T(2/3) = 2/3$. Consider a sequence $\beta_j = 5/2 + \varepsilon_j$, with $\varepsilon_j = j^{-\alpha}$, where $\alpha < 1$. Note that $1/n = o(\varepsilon_n)$.

Observe that $T_j(2/3) = 2/3 + O(\varepsilon_j)$. Also note that, since we are choosing, deliberately, $\varepsilon_j > 0$ for all j, then the orbit of ζ is being pulled to the right everytime we iterate. Moreover, by letting j be sufficiently large we can keep it inside a small neighbourhood of 2/3 at least up to a certain fixed number of iterates $J \in \mathbb{N}$.

For $\delta > 0$, we have that $T_j(2/3 - \delta) = 2/3 + O(\delta) + O(\varepsilon_j)$. So if we take $\delta = \delta_n$ such that $B_{\delta_n}(\zeta) = U_n$ then $\delta_n = O(1/n)$ and we see that if j and n are sufficiently large then $T_j(2/3 - \delta_n) > 2/3 + \delta_n$. Hence, by continuity, for some fixed $J \in \mathbb{N}$, we can show that for j and n sufficiently large then for all $i = 1, \dots, J$ we have $T_{j+i} \circ \cdots \circ T_j(U_n) \cap U_n = \emptyset$. This means that we would be able to show that $Д'_0(u_n)$ would hold with $A^{(q)}_{n,i} = U_n$ (meaning that $q = 0$).

The conclusion then is that at $\zeta = 2/3$, although for the unperturbed system T shows an EI equal to $1 - 2/5 = 3/5$, for the sequential systems chosen as above the EI is equal to 1.

Remark 4.5. *Note that condition* (4.1) *was used to prove* (2.2) *so*, *in this case*, *we may need to use different* $u_{n,i}$ *for each* i *but*, *since the invariant measure of each* T_i *is equivalent to Lebesgue measure*, *the corresponding* $\delta_{n,i}$ *still satisfies* $\delta_{n,i} = O(1/n)$ *for all* $i \in \mathbb{N}$.

5. Random fibered dynamical systems

We now provide a second example of non-stationary dynamical systems, this time arising from suitable random perturbations.

We consider a probability space $(\Omega, \mathcal{G}, P)$ with an invertible P-preserving transformation $\vartheta : \Omega \to \Omega$; then we let $(\boldsymbol{\Xi}, \mathcal{F})$ another measurable space and Ξ a measurable (with respect to the product $\mathcal{G} \times \mathcal{F}$) subset of $\boldsymbol{\Xi} \times \Omega$ with the fibers $\Xi^\omega = \{\xi \in \boldsymbol{\Xi} : (\xi, \omega) \in \Xi\} \in \mathcal{F}$. We define the (skew) map $s : \Xi \to \Xi$ by $s(\xi, \omega) = (f_\omega \xi, \vartheta\omega)$, with $f_\omega : \Xi^\omega \to \Xi^{\vartheta\omega}$ being measurable fiber maps with the composition rule

$$f^n_\omega : \Xi^\omega \to \Xi^{\vartheta^n\omega}, \qquad f^n_\omega = f_{\vartheta^{n-1}\omega} \circ \cdots \circ f_\omega.$$

We also put

$$f^j_{\vartheta^l\omega} : \Xi^{\vartheta^l\omega} \to \Xi^{\vartheta^{l+j}\omega}; \qquad f^j_{\vartheta^l\omega} = f_{\vartheta^{l+j-1}\omega} \circ \cdots \circ f_{\vartheta^l\omega}.$$

Moreover we set

$$f^{-1}_{\vartheta^j\omega} : \Xi^{\vartheta^{j+1}\omega} \to \Xi^{\vartheta^j\omega} \quad \text{and} \quad \left(f^k_\omega\right)^{-1} := f^{-1}_\omega \circ \cdots \circ f^{-1}_{\vartheta^{k-1}\omega}.$$

This will allow us to introduce the σ-algebras $\mathcal{T}^\omega_k := (f^k_\omega)^{-1}\mathcal{T}_0^{\vartheta^k\omega}$ where $\mathcal{T}_0^{\vartheta^k\omega}$ is the restriction of the σ-algebra $\mathcal{F}$ to $\Xi^\omega \subset \boldsymbol{\Xi}$.

It is well known that a measure μ disintegrated with respect to the measure P will be s-invariant if the conditional measures μ^ω will verify the quasi-invariant relation

$$(f_\omega)_*\mu^\omega = \mu^{\vartheta\omega}. \tag{5.1}$$

An interesting case is whenever all the fibers Ξ^ω coincide with the metric space X. In this case we can also define a marginal measure μ on X in the following way: given $A \subset X$, define

$$\mu(A) = \tilde{\mu}(\Omega \times A) = \int_\Omega \mu^\omega(A)\, dP(\omega).$$

Also in this case, the stochastic process is defined by

$$X_i = \varphi \circ f_\omega^i, \tag{5.2}$$

where $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ is as in (3.9). This stochastic process $X_0, X_1, \ldots$ is not necessarily stationary and, by (5.1), the distribution function of X_i is given by

$$F_i(u) = \mu^{\vartheta^i\omega}\big(\{x \in X : \varphi(x) \le u\}\big).$$

In this setting, we will consider that the boundary levels $u_{n,0}, u_{n,1}, \ldots$ are such that $u_n = u_{n,0} = u_{n,1} = \cdots$, where u_n is determined by the marginal measure μ so that

$$u_n = \inf\left\{u \in \mathbb{R} : \mu\big(\{x \in X : \varphi(x) \le u\}\big) \ge 1 - \frac{\tau}{n}\right\}.$$

Then as a result of the theory developed in Section 1.2, we can write a quenched distributional limit for the partial maxima of the process $X_0, X_1, \ldots$. Namely, as a consequence of Theorem 2.4 we have

Corollary 5.1. *Let $X_0, X_1, \ldots$ be a stationary stochastic process defined as above, based on the action of the fiber maps f_ω^n. Assume that for P-a.e. $\omega \in \Omega$ conditions (2.1) and (2.2) hold for some $\tau > 0$. Assume that there exists $q \in \mathbb{N}_0$, defined as in (2.4), and (2.8) holds for P-a.e. $\omega \in \Omega$. Assume moreover that conditions $\mathcal{Д}_q(u_{n,i})$ and $\mathcal{Д}'_q(u_{n,i})$ are satisfied for P-a.e. $\omega \in \Omega$. Then*

$$\lim_{n\to\infty} \boldsymbol{\mu}^\omega\big(\max\{X_0, \ldots, X_{n-1}\} \le u_n\big) = \mathrm{e}^{-\theta\tau}, \quad \text{for } P\text{-a.e. } \omega \in \Omega.$$

To illustrate an application of the theory developed here and in particular of Corollary 5.1, we look into random subshifts.

5.1. *Random subshifts*

We consider the random subshifts studied in [31] and [32], in the setting of Hitting Times. Here we will keep using an Extreme Values approach and the statements can be seen as a translation of the corresponding results in [31,32], in light of the connection between HTS and EVL proved in [12,13].

Since the target sets, in this example, are dynamically defined cylinders, we need to produce some adjustments to the definition of the observable and to the time scale, as in [13, Section 5] (where the notion of cylinder EVL was introduced), in order to properly use an EVL approach. We return to this issue below. Meanwhile, we introduce the notions using mostly the notation of [32].

Let (Ω, ϑ, P) be an invertible ergodic measure preserving system, set $X = \mathbb{N}^{\mathbb{N}_0}$ and let $\sigma : X \to X$ denote the shift. Let $A = \{A(\omega) = (a_{ij}(\omega)) : \omega \in \Omega\}$ be a random transition matrix, i.e., for any $\omega \in \Omega$, $A(\omega)$ is in an $\mathbb{N} \times \mathbb{N}$-matrix with entries in $\{0, 1\}$, with at least one non-zero entry in each row and each column and such that $\omega \to a_{ij}(\omega)$ is measurable for any $i \in \mathbb{N}$ and $j \in \mathbb{N}$. For any $\omega \in \Omega$ define

$$X_\omega = \big\{x = (x_0, x_1, \ldots) : x_i \in \mathbb{N} \text{ and } a_{x_i x_{i+1}}\big(\vartheta^i\omega\big) = 1 \text{ for all } i \in \mathbb{N}\big\}$$

and

$$\mathcal{E} = \{(\omega, x) : \omega \in \Omega, x \in X_\omega\} \subset \Omega \times X.$$

We consider the random dynamical system coded by the skew-product $S : \mathcal{E} \to \mathcal{E}$ given by $S(\omega, x) = (\vartheta\omega, \sigma x)$. While we allow infinite alphabets here, we nevertheless call S a *random subshift of finite type* (SFT). Assume that ν is an S-invariant probability measure with marginal P on Ω. Then we let $(\mu^\omega)_\omega$ denote its decomposition on X_ω, that is, $d\nu(\omega, x) = d\mu_\omega(x)\,dP(\omega)$. The measures μ^ω are called the *sample measures*. Note $\mu^\omega(A) = 0$ if $A \cap X_\omega = \emptyset$. As before, we denote by $\mu = \int \mu^\omega\,dP$ the marginal of ν on X.

For any $y \in X$ we denote by $C_n(y) = \{z \in X : y_i = z_i \text{ for all } 0 \le i \le n-1\}$ the n-cylinder that contains y. Let $\mathcal{F}_0^n$ be the σ-algebra in X, generated by all the n-cylinders.

We assume the following: there are constants $h_0 > 0$, $c_0 > 0$ and a summable function ψ such that for all m, n, $\kappa \in \mathbb{N}$, $A \in \mathcal{F}_0^n$ and $B \in \mathcal{F}_0^m$:

(1) the marginal measure μ satisfies

$$\left|\mu\left(A \cap \sigma^{-\kappa-n} B\right) - \mu(A)\mu(B)\right| \le \psi(\kappa);$$

(2) for P-almost every $\omega \in \Omega$, if $y \in X_\omega$ and $n \ge 1$ then $c_0^{-1}\mathrm{e}^{-h_0 n} \le \mu(c_n(y))$;
(3) for P-almost every $\omega \in \Omega$,

$$\left|\mu^\omega\left(A \cap \sigma^{-\kappa-n} B\right) - \mu^\omega(A)\mu^{\vartheta^{n+\kappa}\omega}(B)\right| \le \psi(\kappa)\mu^\omega(A)\mu^{\vartheta^{n+\kappa}\omega}(B);$$

(4) the sample measure satisfies

$$\operatorname*{essup}_{\omega\in\Omega} \sup_{x\in X} \mu^\omega\left(C_1(x)\right) < 1.$$

The following lemma has been proved in [32].

Lemma 5.2. *For a random SFT such that assumptions* (5.1) *and* (5.1) *hold*, *there exist* c_1, $c_2 > 0$ *and* $h_1 > 0$ *such that for any* $y \in X$, $n \ge 1$ *and* $m \ge 1$, *for almost* P*-almost every* $\omega \in \Omega$,

$$\mu^\omega\left(C_n(y)\right) \le c_1 \mathrm{e}^{-h_1 n}$$

and

$$\sum_{k=m}^{n} \mu^\omega\left(C_n(y) \cap \sigma^{-k} C_n(y)\right) \le c_2 \mathrm{e}^{-h_1 m} \mu^\omega\left(C_n(y)\right).$$

Since the target sets are cylinders, in order to state the result using an EVL approach, as mentioned earlier, we need to make some adjustments to the definition of the observable function and to the time scale. Hence, proceeding as in [13, Section 5], the stochastic process is defined by $X_i = \varphi \circ \sigma^i$, where $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ instead of being given by (3.9) is given by

$$\varphi(x) = g\left(\mu\left(C_{n(x)}(\zeta)\right)\right),$$

where $n(x) := \max\{j \in \mathbb{N} : x \in C_j(\zeta)\}$ and g is as in Section 3.2. As in [13, (5.5)] we let the sequence $(u_n)_{n\in\mathbb{N}}$ be such that $\{x \in X : \varphi(x) > u_n\} = C_n(\zeta)$. Moreover, for the time scale we use the sequence $(w_n)_{n\in\mathbb{N}}$ given by [13, (5.6)]:

$$w_n = \left[\tau\mu\left(\{x \in X : \varphi(x) > u_n\}\right)\right],$$

for some $\tau \ge 0$.

Now, we can apply Corollary 5.1 to obtain the following result, which is a translation to the EVL setting of [32, Theorem 2.2].

Theorem 5.3. *Assume (5.1)–(5.1) hold and there exists a constant* $q > 2\frac{h_0}{h_1}$ *such that* ψ *satisfies* $\psi(\kappa)\kappa^q \to$ *as* $\kappa \to +\infty$. *Let* $\zeta \in X$. *Then for* P*-almost every* ω, *either*

(a) ζ *is a periodic point of period* p *and if the limit* $\theta := \lim_{n\to\infty} \frac{\mu(C_n(\zeta)\setminus C_{n+p}(\zeta))}{\mu(C_n(\zeta))}$ *exists, then for all* $\tau \geq 0$ *we have*

$$\lim_{n\to\infty} \mu^{\omega}(M_{w_n} \leq u_n) = \mathrm{e}^{-\theta\tau};$$

or

(b) *for all* $\tau \geq 0$ *we have*

$$\lim_{n\to\infty} \mu^{\omega}(M_{w_n} \leq u_n) = \mathrm{e}^{-\tau}.$$

In order to use Corollary 5.1 to prove Theorem 5.3, one needs to check that conditions (2.2), $Д_q(u_{n,i})$, $Д'_q(u_{n,i})$ and (2.8) hold for P-a.e. $\omega \in \Omega$.

Note that because of the adjustments required to the cylinder setting, for condition (2.2), one needs to check that for P-a.e. $\omega \in \Omega$ we have

$$\lim_{n\to\infty} \sum_{i=0}^{w_n} \mu^{\vartheta^i(\omega)}\big(C_n(\zeta)\big) = \tau,$$

which follows immediately from [32, Lemma 4.5]. In the same way, conditions $Д_q(u_{n,i})$, $Д'_q(u_{n,i})$ follow from [32, Lemma 4.8] and [32, Lemma 4.9] respectively and condition (2.8) from the discussion in [32, Section 5].

6. Concluding remarks

The sequential systems considered in this paper were built on uniformly expanding maps, for which the transfer operators admits a spectral gap and the correlations decay exponentially. In a different direction, a class of sequential systems given by composition of non-uniformly expanding maps of Pomeau–Manneville type was studied in [1], by perturbing the slope at the indifferent fixed point 0. Polynomial decay of correlations was proved for particular classes of centred observables, which could also be interpreted as the decay of the iterates of the transfer operator on functions of zero (Lebesgue) average, and this fact is better known as *loss of memory*. In the successor paper [27], a (non-stationary) central limit theorem was shown for sums of centred observables and with respect to the Lebesgue measure. In the forthcoming paper [16] we will continue the statistical analysis of these indifferent transformations by proving the existence of extreme value distributions under suitable normalisation for the threshold of the exceedances.

Acknowledgements

ACMF was partially supported by FCT (Portugal) grant SFRH/BPD/66174/2009. JMF was partially supported by FCT grant SFRH/BPD/66040/2009. All these grants are financially supported by the program POPH/FSE. ACMF and JMF were partially supported by FCT project FAPESP/19805/2014 and by CMUP (UID/MAT/00144/2013), which is funded by FCT with national (MEC) and European structural funds through the programs FEDER, under the partnership agreement PT2020. SV was supported by the ANR-Project *Perturbations* and by the project *Atracción de Capital Humano Avanzado del Extranjero* MEC 80130047, CONICYT, at the CIMFAV, University of Valparaiso. All authors were partially supported by FCT projects PTDC/MAT/120346/2010 and PTDC/MAT-CAL/3884/2014, which are funded by national and European structural funds through the programs FEDER and COMPETE. SV is grateful to N. Haydn, M. Nicol and A. Török for several and fruitful discussions on sequential systems, especially in the framework of indifferent maps. JMF is grateful to M. Todd for fruitful discussions and careful reading of a preliminary version of this paper. All authors acknowledge the Isaac Newton Institute for Mathematical Sciences, where this work was initiated during the program Mathematics for the Fluid Earth.

References

[1] R. Aimino, H. Hu, M. Nicol, A. Török and S. Vaienti. Polynomial loss of memory for maps of the interval with a neutral fixed point. *Discrete Contin. Dyn. Syst.* **35** (3) (2015) 793–806. doi:10.3934/dcds.2015.35.793. MR3277171

[2] J. F. Alves, J. M. Freitas, S. Luzzatto and S. Vaienti. From rates of mixing to recurrence times via large deviations. *Adv. Math.* **228** (2) (2011) 1203–1236. Available at arXiv:1007.3771. doi:10.1016/j.aim.2011.06.014. MR2822221

[3] L. Arnold. *Random Dynamical Systems*. *Springer Monographs in Mathematics*. Springer-Verlag, Berlin, 1998. doi:10.1007/978-3-662-12878-7. MR1374107

[4] H. Aytaç, J. M. Freitas and S. Vaienti. Laws of rare events for deterministic and random dynamical systems. *Trans. Amer. Math. Soc.* **367** (11) (2015) 8229–8278. doi:10.1090/S0002-9947-2014-06300-9. MR3391915

[5] W. Bahsoun and S. Vaienti. Escape rates formulae and metastability for randomly perturbed maps. *Nonlinearity* **26** (5) (2013) 1415–1438. doi:10.1088/0951-7715/26/5/1415. MR3056132

[6] D. Berend and V. Bergelson. Ergodic and mixing sequences of transformations. *Ergodic Theory Dynam. Systems* **4** (3) (1984) 353–366. doi:10.1017/S0143385700002509. MR0776873

[7] M. R. Chernick, T. Hsing and W. P. McCormick. Calculating the extremal index for a class of stationary sequences. *Adv. in Appl. Probab.* **23** (4) (1991) 835–850. doi:10.2307/1427679. MR1133731

[8] P. Collet. Statistics of closest return for some non-uniformly hyperbolic systems. *Ergodic Theory Dynam. Systems* **21** (2) (2001) 401–420. doi:10.1017/S0143385701001201. MR1827111

[9] J.-P. Conze and A. Raugi. Limit theorems for sequential expanding dynamical systems on $[0,1]$. In *Ergodic Theory and Related Fields* 89–121. *Contemp. Math.* **430**. Amer. Math. Soc., Providence, RI, 2007. doi:10.1090/conm/430/08253. MR2331327

[10] M. Falk, J. Hüsler and R.-D. Reiss. *Laws of Small Numbers: Extremes and Rare Events*, extended edition. Birkhäuser/Springer Basel AG, Basel, 2011. doi:10.1007/978-3-0348-0009-9. MR2732365

[11] A. C. M. Freitas and J. M. Freitas. On the link between dependence and independence in extreme value theory for dynamical systems. *Statist. Probab. Lett.* **78** (9) (2008) 1088–1093. doi:10.1016/j.spl.2007.11.002. MR2422964

[12] A. C. M. Freitas, J. M. Freitas and M. Todd. Hitting time statistics and extreme value theory. *Probab. Theory Related Fields* **147** (3–4) (2010) 675–710. Available at arXiv:0804.2887. doi:10.1007/s00440-009-0221-y. MR2639719

[13] A. C. M. Freitas, J. M. Freitas and M. Todd. Extreme value laws in dynamical systems for non-smooth observations. *J. Stat. Phys.* **142** (1) (2011) 108–126. Available at arXiv:1006.3276. doi:10.1007/s10955-010-0096-4. MR2749711

[14] A. C. M. Freitas, J. M. Freitas and M. Todd. The extremal index, hitting time statistics and periodicity. *Adv. Math.* **231** (5) (2012) 2626–2665. Available at arXiv:1008.1350. doi:10.1016/j.aim.2012.07.029. MR2970462

[15] 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. Available at arXiv:1401.4206. doi:10.1016/j.spa.2014.11.011. MR3310360

[16] A. C. M. Freitas, J. M. Freitas and S. Vaienti Extreme value laws for non stationary, sequential intermittent systems. Preprint, 2016. Available at arXiv:1605.06287.

[17] C. González-Tokman, B. R. Hunt and P. Wright. Approximating invariant densities of metastable systems. *Ergodic Theory Dynam. Systems* **31** (5) (2011) 1345–1361. doi:10.1017/S0143385710000337. MR2832249

[18] N. Haydn, M. Nicol, A. Török and S. Vaienti Almost sure invariance principle for sequential and non-stationary dynamical systems. Preprint. Available at arXiv:1406.4266.

[19] N. Haydn, M. Nicol, S. Vaienti and L. Zhang. Central limit theorems for the shrinking target problem. *J. Stat. Phys.* **153** (5) (2013) 864–887. doi:10.1007/s10955-013-0860-3. MR3124980

[20] H. Hu and S. Vaienti. Absolutely continuous invariant measures for non-uniformly expanding maps. *Ergodic Theory Dynam. Systems* **29** (4) (2009) 1185–1215. doi:10.1017/S0143385708000576. MR2529645

[21] J. Hüsler. Asymptotic approximation of crossing probabilities of random sequences. *Z. Wahrsch. Verw. Gebiete* **63** (2) (1983) 257–270. doi:10.1007/BF00538965. MR0701529

[22] J. Hüsler. Extreme values of nonstationary random sequences. *J. Appl. Probab.* **23** (4) (1986) 937–950. MR0867190

[23] Y. Kifer. *Ergodic Theory of Random Transformations*. *Progress in Probability and Statistics* **10**. Birkhäuser Boston, Inc., Boston, MA, 1986. doi:10.1007/978-1-4684-9175-3. MR0884892

[24] Y. Kifer. *Random Perturbations of Dynamical Systems*. *Progress in Probability and Statistics* **16**. Birkhäuser Boston, Inc., Boston, MA, 1988. doi:10.1007/978-1-4615-8181-9. MR1015933

[25] Y. Kifer. Limit theorems for random transformations and processes in random environments. *Trans. Amer. Math. Soc.* **350** (4) (1998) 1481–1518. doi:10.1090/S0002-9947-98-02068-6. MR1451607

[26] Y. Kifer and A. Rapaport. Poisson and compound Poisson approximations in conventional and nonconventional setups. *Probab. Theory Related Fields* **160** (3–4) (2014) 797–831. doi:10.1007/s00440-013-0541-9. MR3278921

[27] M. Nicol, A. Török and S. Vaienti Central limit theorems for sequential and random intermittent dynamical systems. Preprint, 2015. Available at arXiv:1510.03214.

[28] X.-F. Niu. Extreme value theory for a class of nonstationary time series with applications. *Ann. Appl. Probab.* **7** (2) (1997) 508–522. doi:10.1214/aoap/1034625342. MR1442324

[29] W. Parry. On the β-expansions of real numbers. *Acta Math. Acad. Sci. Hung.* **11** (1960) 401–416. MR0142719

[30] J. Rousseau. Hitting time statistics for observations of dynamical systems. *Nonlinearity* **27** (9) (2014) 2377–2392. doi:10.1088/0951-7715/27/9/2377. MR3266858

[31] J. Rousseau, B. Saussol and P. Varandas. Exponential law for random subshifts of finite type. *Stochastic Process. Appl.* **124** (10) (2014) 3260–3276. doi:10.1016/j.spa.2014.04.016. MR3231619

[32] J. Rousseau and M. Todd. Hitting times and periodicity in random dynamics. *J. Stat. Phys.* **161** (1) (2015) 131–150. doi:10.1007/s10955-015-1325-7. MR3392511

[33] B. Saussol. Absolutely continuous invariant measures for multidimensional expanding maps. *Israel J. Math.* **116** (2000) 223–248. doi:10.1007/BF02773219. MR1759406