RARE EVENTS FOR PRODUCT FRACTAL SETS

ANA CRISTINA MOREIRA FREITAS, JORGE MILHAZES FREITAS,
AND JORGE VALENTIM SOARES

Abstract. We analyse the existence of limiting laws of rare events for dynamically generated stochastic processes. We consider two-dimensional dynamical systems and observable functions maximised on Cantor dust sets and prove the existence of distributional limits for the partial maxima. We show how the Extremal Index, measuring the degree of clustering of rare events, is linked to the compatibility between the dynamics and the fractal structure of the maximal sets.

1. Introduction

The study of extreme events is of crucial importance in a multitude of scenarios where their occurrence has a serious disruption effect. This is the case of natural hazards such as earthquakes, storms, draughts, pandemics or human-made disasters such as industrial and transport accidents, oil spills, nuclear explosions, financial crashes, etc.

These phenomena correspond to very peculiar states of systems whose time evolution is, often, accurately described by mathematical models called dynamical systems. Namely, whenever the orbits of the system (the several successions of states through which the system goes during a certain realisation) hit small critical regions of the phase space corresponding to abnormal configurations, one observes extreme events. The critical regions are neighbourhoods of a critical set, which we will denote by $\mathcal{M}$ corresponding to configurations where appropriate observable functions achieve their maximum or minimum, *i.e.*, their extremes.

The recurrence properties of these critical sets $\mathcal{M}$ are intimately connected with the time distribution and respective impact of such abnormal observations, in particular, regarding the tendency to observe clusters or grouping of extremes. The study of rare events for dynamical systems has enjoyed enormous development in the last few years. The first analytical results considered that $\mathcal{M}$ was reduced to a single point in the phase space, which meant that clustering of extremes was directly associated to periodicity of $\mathcal{M}$ (see [15] and references therein). Recent works have considered $\mathcal{M}$ to be a finite set of points [3, 17], countably many points [4] or smooth submanifolds [7, 10]. In most of the literature, the maximal sets $\mathcal{M}$ have a fairly regular geometrical structure. The exceptions are the papers [13, 22], where $\mathcal{M}$ is taken as a Cantor set, which means it has a a more complex geometry reflected in its fractal nature.

Date: July 12, 2021.

2010 *Mathematics Subject Classification.* 37A50, 37B20, 60G70.

All authors were partially supported by by FCT project PTDC/MAT-PUR/28177/2017 and by CMUP, which is financed by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project with reference UIDB/00144/2020.

In this paper, we build up on [13] and study the existence of Extreme Value Laws for bi-dimensional fractal sets. We were motivated by the potential of application in climate dynamics. This area is a particular source of examples because the underlying mathematical models present chaotic behaviour, which is often revealed in the presence of critical regions with a complex multifractal structure. For example, in [11], the anomalies observed for the precipitation frequency data show a complex geometry consistent with an underlying fractal set. Moreover, because data is usually depicted in two-dimensional charts, then it is important to understand the connections between the fractal geometry of the maximal sets we consider here and the respective compatibility with the dynamics. We also mention the papers [9, 12, 20], where critical regions with multifractal properties appear in the study of turbulent datasets, greenhouse effect or the metastable states: warm and snow ball. Finally, we recommend the paper [6] for a nice discussion about the dimension of the underlying attractors and Extreme Value techniques.

In [13], we proved that the appearance of clustering of rare events was directly connected to the compatibility of the dynamical systems with the geometric fractal structure of the maximal set $\mathcal{M}$. Namely, we considered $\mathcal{M}$ to be a Cantor set and took uniformly expanding systems T such as those of the form $mx \mod 1$ and then observed that for compatible systems, as when $m = 3^k$, for some $k \in \mathbb{N}$, for which $\mathcal{M}$ is actually an invariant set, we obtained clustering of extremes. When the system was incompatible, *i.e.*, $m \neq 3^k$, for all $k \in \mathbb{N}$, for example, then the intersections $\mathcal{M} \cap T^{-j}(\mathcal{M})$, for $j \in \mathbb{N}$, although not empty, was not relevant in terms of its box dimension when compared with that of $\mathcal{M}$. This meant that, ultimately, we have no clustering of extreme values. Geometric tools such as fractal dimension and thickness proved to be very important in order to establish the results and the connections.

In here, we consider that $\mathcal{M}$ will be Cantor dust sets on the plane, which are obtained as the direct product of two Cantor sets. The dynamics will consist of the direct product of the uniformly expanding maps considered in [13]. Although these are very simple models, which is important in order to obtain the closed formulas and estimates we get here, they capture much of the possible behaviours in a rather transparent way. In fact, regarding the compatibility between the dynamics and the geometry, when compared with the one-dimensional situation studied in [13], we can have more cases here because we may have compatibility in both directions, only in one direction or in none of the directions. Yet, our simple models, allow us to obtain a rather global picture: compatibility in both directions yields clustering of extremal observations, while the mere existence of incompatibility in one of the directions is enough to guarantee that we have no clustering.

The paper is organised as follows. In Section 2, we introduce the framework regarding the study of rare events and provide conditions and results useful to prove the existence of Extreme Value Laws. In particular, we recall the notion of Extremal Index, which is a numerical indicator of the strength of clustering. In Section 3, we define the stochastic processes of interest, which are generated by observable functions, maximised at a Cantor dust subset of the plane, which are evaluated along the orbits of uniformly expanding dynamical systems consisting in product maps. The main results of the paper establishing the existence of limiting Extreme Value Laws for such stochastic processes are stated in the end of this section. In Section 4, we prove the existence of limiting laws in the presence of clustering created by the compatibility between the dynamics and the geometry of $\mathcal{M}$, while, in Section 5, we prove the results establishing the non existence of clustering of extremal observations.

2. Laws for rare events

Let $(\mathcal{X}, \mathcal{B}, T, \mu)$ be a discrete dynamical system, where $\mathcal{X}$ is a compact manifold, $\mathcal{B}$ is the corresponding Borel σ-algebra, $T : \mathcal{X} \to \mathcal{X}$ is a measurable map, and μ is the invariant measure associated with T. Start by considering an observable function $\varphi : \mathcal{X} \to \mathbb{R}^+ \cup \{\infty\}$ and assume that there exists $Z \in \mathbb{R}^+ \cup \{+\infty\}$ such that $Z = \max_{x \in \mathcal{X}} \varphi(x)$. Denote the set of global maximal points of φ by $\mathcal{M}$, *i.e*

$$\mathcal{M} = \{x \in \mathcal{X} : \ \varphi(x) = Z\}.$$

Define the stochastic process, $(X_n)_{n \in \mathbb{N}}$, as

$$X_n(x) = \varphi \circ T^n(x). \tag{2.1}$$

The process of partial maxima, $(M_n)_{n \in \mathbb{N}}$, associated with X_n is constructed in the following way:

$$M_n = \max\{X_0, \ldots, X_{n-1}\}. \tag{2.2}$$

The objective is to find a limiting distribution for the process $(M_n)_n$. To obtain such a law, we study the level sets $\{X_j > u\}$ which can be seen as exceedances of a given threshold u. The idea is to estimate the probability of not exceeding a high threshold u up to some time m, depending on u. This way, in the limit, we will be estimating the measure of the set $M_m \leq u$ as u approaches Z.

It is important to find the right dependence of m on u in order to find a non-degenerate distribution. Due to the types of observables used in this paper the measure of the level set $\{X_j > u\}$, seen as a function of u, is not smooth. Hence, we must use the relation first introduced in [14].

We consider sequences $(w_n)_{n \in \mathbb{N}}$ and $(u_n)_{n \in \mathbb{N}}$ satisfying

$$w_n \mu(X_0 > u_n) \to \tau \text{ as } n \to \infty \text{ for some } \tau \geq 0. \tag{2.3}$$

Our aim will be to find a non-degenerate distribution function H, whose support is $\mathbb{R}^+$, such that

$$\lim_{n \to \infty} \mu(M_{w_n} \leq u_n) = 1 - H(\tau). \tag{2.4}$$

This type of distributional limit has been called as cylinder Extreme Value Law. In the course of this article, we refer to this law as EVL.

Under the right normalisation, the limit $1 - H(\tau)$ can be represented by $\mathrm{e}^{-\theta \tau}$, where $\tau(y)$ must be one of the following types:

- $\tau_1(y) = e^{-y}$ for $y \in \mathbb{R}$. (Gumbel)
- $\tau_2(y) = y^{-\beta}$ for $y, \beta > 0$. (Frechet)
- $\tau_3(y) = (-y)^{\gamma}$ for $y \leq 0$ and $\gamma > 0$.(Weibull).

The parameter θ is called the *Extremal Index* and can be interpreted as a measure of the level of clustering of exceedances. Its value is closely linked to the recurrence behaviour of the set of global maxima of φ. It was proved in [2, 15, 16, 19] that, when $\mathcal{M}$ is reduced to a single periodic point, this periodicity would lead to clustering of exceedences which results in an EI smaller than 1. If, however, $\mathcal{M}$ was reduced to a single non-periodic point, then the non-recurrence properties of the point would lead to absence of clustering and to an EI equal to 1.

This idea was further extended in [4,13,17] for the cases where $\mathcal{M}$ is a larger set. It was shown that the nature of $T^{-j}(\mathcal{M})\cap\mathcal{M}$ determines the level of clustering appearing in the process $(X_n)_n$ and therefore the value of θ.

2.1. **Existence of Extreme Value Laws.** It is necessary to provide general conditions that guarantee the existence of distributional limit as stated in (2.4).

We start by considering a sequence of thresholds $(u_n)_{n\in\mathbb{N}}$ and a sequence $(w_n)_{n\in\mathbb{N}}$ as in (2.3). In addition, we define a sequence $(q_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} q_n=\infty \qquad \text{and} \qquad \lim_{n\to\infty}\frac{q_n}{w_n}=0. \tag{2.5}$$

For fixed $u\in\mathbb{R}$ and $q\in\mathbb{N}$, the event $U(u)$ is defined as

$$U(u):=\{X_0>u\}. \tag{2.6}$$

This event corresponds to the existence of an exceedance at time $t=0$.

Let T^{-i} denote the i-th preimage by the map T, then using $U(u)$, we construct the set

$$\mathcal{A}_q(U(u)):=U(u)\cap\bigcap_{i=1}^{q}T^{-i}(U(u)^c)=\{X_0>u,X_1\le u,\ldots,X_q\le u\}. \tag{2.7}$$

The event $\mathcal{A}_q(U(u))$ corresponds to the case where we have an exceedance at time $t=0$ that is not followed by another one up to time $t=q$. The occurrence of $T^{-i}(\mathcal{A}_q(U(u)))$ can be thought of as the expiration of a cluster of exceedances, whose last exceedance is observed precisely at time i. (See [1, Section 2.1], for more insight).

For all $s,\ell\in\mathbb{N}$ and any $B\in\mathcal{B}$, we set

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

For each $n\in\mathbb{N}$, set $U_n:=U(u_n)$ and $\mathcal{A}_{q_n,n}:=\mathcal{A}_{q_n}(U_n)$. With this notation, we can write

$$\mathscr{W}_{0,w_n}(U_n)=\{M_{w_n}\le u_n\}.$$

The existence of an Extreme Value Law is assured by two conditions.

Condition ($Д_{q_n}(u_n,w_n)$)**.** We say that $Д_{q_n}(u_n,w_n)$ holds for the stochastic process $(X_n)_{n\in\mathbb{N}}$ if for every $\ell,t,n\in\mathbb{N}$

$$\left|\mu\left(\mathcal{A}_{q_n,n}\cap\mathscr{W}_{t,\ell}\left(\mathcal{A}_{q_n,n}\right)\right)-\mu\left(\mathcal{A}_{q_n,n}\right)\mu\left(\mathscr{W}_{0,\ell}\left(\mathcal{A}_{q_n,n}\right)\right)\right|\le\gamma(n,t), \tag{2.8}$$

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

Consider the sequence $(t_n)_{n\in\mathbb{N}}$ used in condition $Д_{q_n}(u_n,w_n)$ and let $(k_n)_{n\in\mathbb{N}}$ be another sequence, such that

$$k_n\to\infty \quad \text{and} \quad k_nt_n=o(w_n). \tag{2.9}$$

Condition ($Д'_{q_n}(u_n, w_n)$)**.** We say that $Д'_{q_n}(u_n, w_n)$ holds for the stochastic process $(X_n)_{n\in\mathbb{N}}$ if there exists a sequence $(k_n)_{n\in\mathbb{N}}$ satisfying (2.9) such that

$$\lim_{n\to\infty} w_n \sum_{j=q_n+1}^{\lfloor w_n/k_n \rfloor -1} \mu\left(\mathcal{A}_{q_n,n} \cap T^{-j}\left(\mathcal{A}_{q_n,n}\right)\right) = 0. \tag{2.10}$$

Condition $Д_{q_n}(u_n, w_n)$ establishes an asymptotic independence between the occurrence of the event $\mathcal{A}_{q_n,n}$ and the absence of occurrences of such an event in the time interval $[t, t+\ell)$. In other words, if after an exceedance, we do not observe another exceedence for a run of q_n observations, then the non-occurrence of another closure of cluster of exceedances is asymptotically independent. On other hand, condition $Д'_{q_n}(u_n, w_n)$ guarantees that clusters of exceedances are well spaced along the time line, excluding the possibility of concentration of clusters.

Using O'Brien's formula, [23], we may define a finite time approximation of the EI, which we denote by θ_n, namely,

$$\theta_n := \frac{\mu\left(\mathcal{A}_{q_n,n}\right)}{\mu(U_n)}. \tag{2.11}$$

When the limit exists, we define:

$$\theta = \lim_{n\to\infty} \theta_n. \tag{2.12}$$

The existence of a limiting law for M_{w_n} is guaranteed by the following result.

Theorem 2.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a stochastic process constructed as in* (2.1)*. Consider the sequences $(u_n)_{n\in\mathbb{N}}$ and $(w_n)_{n\in\mathbb{N}}$ satisfying* (2.3) *for some $\tau \geq 0$. Assume that conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ hold for some $q_n \in \mathbb{N}_0$ satisfying* (2.5)*. Moreover, assume that the sequence $(\theta_n)_{n\in\mathbb{N}}$ defined in* (2.11) *converges to some $0 \leq \theta \leq 1$,* i.e.*, $\theta = \lim_{n\to\infty} \theta_n$. Then,*

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

The proof of this theorem follows from an easy adjustment of the proof of [21, Corollary 4.1.7]. The use of θ for the limit in (2.12) is justified by the previous theorem, which establishes that this limit (when it exists) can be identified as being the EI.

2.2. **Applications to systems defined in a two-dimensional space.**

The objective of this section is to achieve a set of sufficient conditions that guarantee that conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ hold for a class of systems defined in a two-dimensional space with some sort of decay of correlations against L^1.

Definition 2.2 (Decay of correlations)**.** Let $\mathcal{C}_1, \mathcal{C}_2$ denote Banach spaces of real valued measurable functions defined on $\mathcal{X}$. We denote the *correlation* of non-zero functions $\phi \in \mathcal{C}_1$ and $\psi \in \mathcal{C}_2$ with respect to a measure μ as

$$\mathrm{Cor}_\mu(\phi, \psi, n) := \frac{1}{\|\phi\|_{\mathcal{C}_1} \|\psi\|_{\mathcal{C}_2}} \left| \int \phi\, (\psi \circ T^n)\, \mathrm{d}\mu - \int \phi\, \mathrm{d}\mu \int \psi\, \mathrm{d}\mu \right|.$$

We say that the dynamical system $(\mathcal{X}, \mathcal{B}, T, \mu)$ has *decay of correlations*, with respect to the measure μ, for observables in $\mathcal{C}_1$ *against* observables in $\mathcal{C}_2$ if there exists a rate function

$\rho : \mathbb{N} \to \mathbb{R}$, with

$$\lim_{n\to\infty} \rho(n) = 0,$$

such that, for every $\phi \in \mathcal{C}_1$ and every $\psi \in \mathcal{C}_2$, we have

$$\mathrm{Cor}_\mu(\phi, \psi, n) \leq \rho(n).$$

In the remaining of this article we may use the notation ρ_n to represent $\rho(n)$.

We will work with systems that have decay of correlations of functions of Bounded Variation or quasi-Hölder functions, which we define below, against observables in $L^1(\mu)$. Hence, in our applications we will always have $\mathcal{C}_2 = L^1(\mu)$.

Definition 2.3. Given a potential $\psi : I \to \mathbb{R}^n$ on an interval I, the *variation* of ψ is defined as

$$\mathrm{Var}(\psi) := \sup \left\{ \sum_{i=0}^{n-1} |\psi(x_{i+1}) - \psi(x_i)| \right\},$$

where the supremum is taken over all finite ordered sequences $(x_i)_{i=0}^n \subset I$.

We use the norm $\|\psi\|_{BV} = \sup|\psi| + \mathrm{Var}(\psi)$, which makes the space of functions of Bounded Variation, $BV := \{\psi : I \to \mathbb{R} : \|\psi\|_{BV} < \infty\}$, into a Banach space.

Definition 2.4. Given an observable $\psi : I \to \mathbb{R}^n$ and a Borel set $Z \subseteq \mathbb{R}^n$, we define the *oscillation* of $\psi \in L^1(\mu)$ over Z as

$$\mathrm{osc}(\psi, Z) := \operatorname*{ess\,sup}_Z \psi - \operatorname*{ess\,inf}_Z \psi.$$

It is possible to verify that $x \mapsto \mathrm{osc}(\psi, B_\epsilon(x))$ is a measurable function (see [24, Proposition 3.1]). Consider real numbers $0 < \alpha \leq 1$ and $\epsilon_0 > 0$, the α - seminorm of ψ is defined as

$$|\psi|_\alpha = \sup_{0<\epsilon\leq\epsilon_0} \epsilon^{-\alpha} \int_{\mathbb{R}^n} \mathrm{osc}(\psi, B_\epsilon(x)) d\mu.$$

The space of functions with bounded α-seminorm is represented by

$$V_\alpha = \left\{ \psi \in L^1(\mu) : |\psi|_\alpha < \infty \right\}.$$

If we endow V_α with the norm

$$\|.\|_\alpha = \|.\|_{L^1(\mu)} + |.|_\alpha$$

then, it becomes a Banach space called the space of quasi-Hölder functions.

We will be considering two-dimensional dynamical systems that are constructed as the direct product of uni-dimensional maps. Consider the dynamical systems $(\mathcal{X}, \mathcal{B}, T_1, \mu)$ and $(\mathcal{X}, \mathcal{B}, T_2, \mu)$. From these maps, we define the product map $T : \mathcal{X}^2 \to \mathcal{X}^2$ by

$$T(x_1, x_2) = (T_1(x_1), T_2(x_2)), \tag{2.13}$$

whose invariant measure is $\mu \times \mu$.

Choosing an observable $\psi : \mathcal{X}^2 \to \mathbb{R}$, we will see that conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ hold for the stochastic process $(X_n)_n = (\psi \circ T^n)_n$ and follow from the decay of correlations mentioned above. Moreover, we will show that is possible to prove that condition $Д'_{q_n}(u_n, w_n)$ holds using only the decay of correlations of the maps T_1 and T_2. For that purpose, let φ_1

and φ_2 be two observables achieving a global maximum on the sets $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, and define the stochastic processes

$$X_n^1 = \varphi_1 \circ T_1^n(x) \qquad \text{and} \qquad X_n^2 = \varphi_2 \circ T_2^n(x), \qquad \text{for each } n \in \mathbb{N}$$

Let $(u_n)_{n\in\mathbb{N}}$ be a sequence of thresholds and consider the sets

$$U_n^{T_1} = \{x \in \mathcal{X} : \varphi_1(x) > u_n\} \qquad \text{and} \qquad U_n^{T_2} = \{x \in \mathcal{X} : \varphi_2(x) > u_n\}$$

associated with $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively.

Assume that the observable ψ achieves a global maximum on the set $\mathcal{M}_1 \times \mathcal{M}_2$, such that the set $U_n = \{x \in \mathcal{X}^2 : \psi(x) > u_n\}$ can be written as

$$U_n = U_n^{T_1} \times U_n^{T_2}. \tag{2.14}$$

Denoting the measure $\mu \times \mu$ by μ^2 and using the setting presented above and under the hypothesis of decay of correlations against L^1 of the maps involved, the next result gives sufficient conditions for $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ to hold.

Theorem 2.5. *Let T be a dynamical system defined as in* (2.13) *and consider an observable ψ, achieving a global maximum on a set $\mathcal{M}_1 \times \mathcal{M}_2$. Let $(X_n)_{n\in\mathbb{N}}$ be the stochastic process given by* (2.1) *and consider sequences $(u_n)_{n\in\mathbb{N}}$, $(w_n)_{n\in\mathbb{N}}$ and $(q_n)_{n\in\mathbb{N}}$ such that* (2.3), (2.5) *and* (2.14) *hold. Assume that T has decay of correlations of functions in $\mathcal{C}_1$ against observables in $L^1(\mu^2)$ and that T_1 and T_2 have decay of correlations of functions in $\mathcal{C}_2$ against observables in $L^1(\mu)$. If,*

(1) $\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{A}_{q_n,n}}\|_{\mathcal{C}_1} w_n \rho(t_n) = 0$ *or* $\lim_{n\to\infty} w_n \left(\|\mathbf{1}_{U_n}\|_{\mathcal{C}_1} \rho(t_n) + 2\mu^2(U_n \setminus \mathcal{A}_{q_n,n})\right) = 0$, *for some sequence $(t_n)_{n\in\mathbb{N}}$ such that $t_n = o(w_n)$*

(2) $\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{\mathcal{C}_2} \mu\left(U_n^{T_2}\right) \sum_{j=q_n}^{\infty} \rho_j^1 = 0$

(3) $\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{\mathcal{C}_2} \mu\left(U_n^{T_1}\right) \sum_{j=q_n}^{\infty} \rho_j^2 = 0$

(4) $\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{\mathcal{C}_2} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{\mathcal{C}_2} \sum_{j=q_n}^{\infty} \rho_j^1 \rho_j^2 = 0$

then conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ are satisfied. Furthermore, if the sequence $(\theta_n)_{n\in\mathbb{N}}$ defined in (2.11) *converges to some $0 \leq \theta \leq 1$ then*

$$\lim_{n\to\infty} \mu^2(M_{w_n} \leq u_n) = \mathrm{e}^{-\theta\tau}.$$

Proof. The dynamical system T has decay of correlations, with respect to the measure μ^2, for functions in $\mathcal{C}_1$ against observables in $L^1(\mu^2)$. Denote the correspondent rate function by ρ.

Similarly, the maps T_1 and T_2 have decay of correlations, with respect to the measure μ, for functions in $\mathcal{C}_2$ against observables in $L^1(\mu)$. Let ρ^1 and ρ^2 denote the respective rate functions.

By Theorem 2.1, we only need to check that the stochastic process $(X_n)_{n\in\mathbb{N}}$ satisfies conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$.

Consider $\phi = \mathbf{1}_{\mathcal{A}_{q_n,n}}$ and $\psi = \mathbf{1}_{\mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})}$ in Definition 2.2. Then, from the decay of correlations against L^1, it follows that there exists $C > 0$, such that, for any positive numbers ℓ and t, we have

$$\begin{aligned}
&|\mu^2(\mathcal{A}_{q_n,n} \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) - \mu^2(\mathcal{A}_{q_n,n})\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n}))| \\
&\qquad = \left| \int_{\mathcal{X}^2} \mathbf{1}_{\mathcal{A}_{q_n,n}} \cdot (\mathbf{1}_{\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n})} \circ T^t) d\mu^2 - \int_{\mathcal{X}^2} \mathbf{1}_{\mathcal{A}_{q_n,n}} d\mu^2 \int_{\mathcal{X}^2} \mathbf{1}_{\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n})} d\mu^2 \right| \\
&\qquad \leq C \|\mathbf{1}_{\mathcal{A}_{q_n,n}}\|_{\mathcal{C}_1} \rho(t).
\end{aligned}$$

If there exists a sequence $(t_n)_{n\in\mathbb{N}}$ such that $t_n = o(w_n)$ and $\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{A}_{q_n,n}}\|_{\mathcal{C}_1} w_n \rho_{t_n} = 0$, which is the content of hypothesis (1), then condition $Д_{q_n}(u_n, w_n)$ follows.

To verify the alternate version of hypothesis (1), consider $\phi = \mathbf{1}_{U_n \setminus \mathcal{A}_{q_n,n} \cup \mathcal{A}_{q_n,n}}$ and $\psi = \mathbf{1}_{\mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})}$ in Definition 2.2. Then, there exists a $C > 0$, such that, for any positive numbers ℓ and t,

$$\begin{aligned}
|\mu^2((U_n \setminus \mathcal{A}_{q_n,n} \cup \mathcal{A}_{q_n,n}) \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) - \mu^2((U_n \setminus \mathcal{A}_{q_n,n} \cup \mathcal{A}_{q_n,n}))\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n}))| \\
\leq C \|\mathbf{1}_{U_n}\|_{\mathcal{C}_1} \rho(t). \qquad (2.15)
\end{aligned}$$

Since $U_n \setminus \mathcal{A}_{q_n,n}$ and $\mathcal{A}_{q_n,n}$ are disjoint, we have that

$$\begin{aligned}
&|\mu^2((U_n \setminus \mathcal{A}_{q_n,n} \cup \mathcal{A}_{q_n,n}) \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) - \mu^2((U_n \setminus \mathcal{A}_{q_n,n} \cup \mathcal{A}_{q_n,n}))\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n}))| \\
=&|\mu^2((U_n \setminus \mathcal{A}_{q_n,n} \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) \cup (\mathcal{A}_{q_n,n} \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n}))) - \mu^2(U_n \setminus \mathcal{A}_{q_n,n} \cup \mathcal{A}_{q_n,n})\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n}))| \\
=&|\mu^2(U_n \setminus \mathcal{A}_{q_n,n} \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) - \mu^2(U_n \setminus \mathcal{A}_{q_n,n})\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n})) \\
&\qquad + \mu^2(\mathcal{A}_{q_n,n} \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) - \mu^2(\mathcal{A}_{q_n,n})\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n}))|.
\end{aligned}$$

Let $A := \mu^2(U_n \setminus \mathcal{A}_{q_n,n} \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) - \mu^2(U_n \setminus \mathcal{A}_{q_n,n})\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n}))$ and $B := \mu^2(\mathcal{A}_{q_n,n} \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) - \mu^2(\mathcal{A}_{q_n,n})\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n}))$, then using (2.15), we obtain that

$$|A + B| + |A| \leq C \|\mathbf{1}_{U_n}\|_{\mathcal{C}_1} \rho(t) + |A|.$$

Using the triangle inequality,

$$|A| \leq 2\mu^2(U_n \setminus \mathcal{A}_{q_n,n})$$

and

$$|B| \leq |A + B| + |A| \leq C \|\mathbf{1}_{U_n}\|_{\mathcal{C}_1} \rho(t) + 2\mu^2(U_n \setminus \mathcal{A}_{q_n,n}).$$

Using again the triangle inequality, we finally achieve that

$$|\mu^2(\mathcal{A}_{q_n,n} \cap \mathscr{W}_{t,\ell}(\mathcal{A}_{q_n,n})) - \mu^2(\mathcal{A}_{q_n,n})\mu^2(\mathscr{W}_{0,\ell}(\mathcal{A}_{q_n,n}))| \leq C \|\mathbf{1}_{U_n}\|_{\mathcal{C}_1} \rho(t) + 2\mu^2(U_n \setminus \mathcal{A}_{q_n,n}).$$

Therefore, condition $Д_{q_n}(u_n, w_n)$ follows if there exists a sequence $(t_n)_{n\in\mathbb{N}}$ such that $t_n = o(w_n)$ and

$$\lim_{n\to\infty} w_n \left(\|\mathbf{1}_{U_n}\|_{\mathcal{C}_1} \rho(t_n) + 2\mu^2(U_n \setminus \mathcal{A}_{q_n,n}) \right) = 0.$$

To prove condition $Д'_{q_n}(u_n, w_n)$, we start by noting that, due to (2.14) and since $\mathcal{A}_{q_n,n} \subseteq U_n$, we have that

$$\begin{aligned}
\mu^2 \left(\mathcal{A}_{q_n,n} \cap T^{-j}(\mathcal{A}_{q_n,n}) \right) &\leq \mu^2 \left((U_n^{T_1} \times U_n^{T_2}) \cap T^{-j}(U_n^{T_1} \times U_n^{T_2}) \right) \\
&= \mu^2 \left((U_n^{T_1} \times U_n^{T_2}) \cap (T_1^{-j}(U_n^{T_1}) \times T_2^{-j}(U_n^{T_2})) \right) \\
&= \mu \left(U_n^{T_1} \cap T_1^{-j}(U_n^{T_1}) \right) \mu \left(U_n^{T_2} \cap T_2^{-j}(U_n^{T_2}) \right).
\end{aligned}$$

The last inequality allows us to write

$$w_n \sum_{j=q_n+1}^{\lfloor w_n/k_n \rfloor} \mu^2\left(\mathcal{A}_{q_n,n} \cap T^{-j}(\mathcal{A}_{q_n,n})\right) \le w_n \sum_{j=q_n+1}^{\lfloor w_n/k_n \rfloor} \mu\left(U_n^{T_1} \cap T^{-j}(U_n^{T_1})\right) \mu\left(U_n^{T_2} \cap T^{-j}(U_n^{T_2})\right). \tag{2.16}$$

Take $\phi = \psi = \mathbf{1}_{U_n^{T_1}}$, in Definition 2.2, to obtain that

$$\mu\left(U_n^{T_1} \cap T_1^{-j}(U_n^{T_1})\right) = \int_{\mathcal{X}} \phi \cdot (\phi \circ T_1^j) d\mu \le \left(\mu(U_n^{T_1})\right)^2 + \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{\mathcal{C}_2} \mu\left(U_n^{T_1}\right) \rho^1(j). \tag{2.17}$$

Likewise, choosing $\phi = \psi = \mathbf{1}_{U_n^{T_2}}$, in Definition 2.2, we obtain that

$$\mu\left(U_n^{T_2} \cap T_2^{-j}(U_n^{T_2})\right) = \int_{\mathcal{X}} \phi \cdot (\phi \circ T_2^j) d\mu \le \left(\mu(U_n^{T_2})\right)^2 + \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{\mathcal{C}_2} \mu\left(U_n^{T_2}\right) \rho^2(j). \tag{2.18}$$

Take $(k_n)_{n\in\mathbb{N}}$ as in (2.9) and consider t_n as above. Recalling that $\lim_{n\to\infty} w_n \mu(U_n) = \tau$ and combining (2.16), (2.17) and (2.18), we can state that

$$\begin{aligned} & w_n \sum_{j=q_n+1}^{\lfloor w_n/k_n \rfloor} \left(\mu(U_n^{T_1})^2 + \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{\mathcal{C}_2} \mu(U_n^{T_1})\rho_j^1\right)\left(\mu(U_n^{T_2})^2 + \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{\mathcal{C}_2} \mu(U_n^{T_2})\rho_j^2\right) \\ & \le \frac{\tau^2}{k_n} + \tau \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{\mathcal{C}_2} \mu\left(U_n^{T_1}\right) \sum_{j=q_n}^{\infty} \rho_j^2 + \tau \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{\mathcal{C}_2} \mu\left(U_n^{T_2}\right) \sum_{j=q_n}^{\infty} \rho_j^1 + \tau \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{\mathcal{C}_2} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{\mathcal{C}_2} \sum_{j=q_n}^{\infty} \rho_j^1 \rho_j^2. \end{aligned}$$

Hence, condition $Д'_{q_n}(u_n, w_n)$ holds if we can verify the following conditions,

$$\begin{aligned} &\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{\mathcal{C}_2} \mu\left(U_n^{T_2}\right) \sum_{j=q_n}^{\infty} \rho_j^1 = 0 \\ &\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{\mathcal{C}_2} \mu\left(U_n^{T_1}\right) \sum_{j=q_n}^{\infty} \rho_j^2 = 0 \\ &\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{\mathcal{C}_2} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{\mathcal{C}_2} \sum_{j=q_n}^{\infty} \rho_j^1 \rho_j^2 = 0. \end{aligned} \tag{2.19}$$

□

Assuming that relation (2.14) holds, one can estimate the level of clustering associated with $(X_n)_n$ by means of the cluster level appearing in the processes $(X_n^1)_n$ and $(X_n^2)_n$.

For that purpose, let $\mathcal{A}_{q_n,n}^{T_1}$ and $\mathcal{A}_{q_n,n}^{T_2}$ represent the sets $\mathcal{A}_{q_n}(U_n^{T_1})$ and $\mathcal{A}_{q_n}(U_n^{T_2})$, respectively.

Let $(q_n)_{n\in\mathbb{N}}$ and $(q_n^*)_{n\in\mathbb{N}}$ be sequences and denote by θ_1 and θ_2 the following limits:

$$\theta_1 := \lim_{n\to\infty} \frac{\mu(\mathcal{A}_{q_n,n}^{T_1})}{\mu(U_n^{T_1})} \qquad \text{and} \qquad \theta_2 := \lim_{n\to\infty} \frac{\mu(\mathcal{A}_{q_n^*,n}^{T_2})}{\mu(U_n^{T_2})} \tag{2.20}$$

The product structure of the maximal set, $\mathcal{M}_1 \times \mathcal{M}_2$, allows for a decomposition of $\mathcal{A}_{q_n,n}$ using $\mathcal{A}_{q_n,n}^{T_1}$ and $\mathcal{A}_{q_n^*,n}^{T_2}$. Such fact, is the base of the following result.

Theorem 2.6. *Let T be a dynamical system defined as in* (2.13) *and consider an observable ψ, achieving a global maximum on a set $\mathcal{M}_1 \times \mathcal{M}_2$. Let $(X_n)_{n\in\mathbb{N}}$ be the stochastic process constructed as in* (2.1) *and consider sequences $(u_n)_{n\in\mathbb{N}}$, $(w_n)_{n\in\mathbb{N}}$, $(q_n)_{n\in\mathbb{N}}$ and $(q_n^*)_{n\in\mathbb{N}}$, such that* (2.3), (2.5) *and* (2.14) *hold.*

(1) assume that $q_n^* \geq q_n$ for all n larger than some $n^* \in \mathbb{N}$; assume further that conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ hold for $(X_n)_n$, the limit (2.12) *exists and the limits in* (2.20) *also exist, then*

$$\lim_{n\to\infty} \mu^2(M_{w_n} \leq u_n) = \mathrm{e}^{-\theta\tau}, \tag{2.21}$$

where

$$\theta > \theta_1 + \theta_2 - \theta_1\theta_2. \tag{2.22}$$

(2) if $\theta_1 = 1$ in (2.20) *and the conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ hold for $(X_n)_n$ or if $\theta_2 = 1$ in* (2.20) *and the conditions $Д_{q_n^*}(u_n, w_n)$ and $Д'_{q_n^*}(u_n, w_n)$ hold for $(X_n)_n$, then*

$$\lim_{n\to\infty} \mu^2(M_{w_n} \leq u_n) = \mathrm{e}^{-\tau}. \tag{2.23}$$

Proof. Since conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ hold, then Theorem 2.1 gives (2.23). Hence, in order to prove (1), we are left with the proof of the lower bound for θ stated in (2.22).

Let $x = (a, b)$ be a point in U_n and assume that $T_1^j(a) \in (U_n^{T_1})^c$ or $T_2^j(b) \in (U_n^{T_2})^c$, for all $j \leq q_n$. This implies that $x \in T^{-j}(U_n^c)$, for all $j \leq q_n$ and consequently $x \in \mathcal{A}_{q_n,n}$. Therefore,

$$\left(\mathcal{A}_{q_n,n}^{T_1} \times U_n^{T_2}\right) \cup \left(U_n^{T_1} \times \mathcal{A}_{q_n,n}^{T_2}\right) \subseteq \mathcal{A}_{q_n,n}. \tag{2.24}$$

Moreover, since for all $n > n^*$ we have that $q_n^* \geq q_n$, then we obtain that

$$\left(\mathcal{A}_{q_n,n}^{T_1} \times U_n^{T_2}\right) \cup \left(U_n^{T_1} \times \mathcal{A}_{q_n^*,n}^{T_2}\right) \subseteq \mathcal{A}_{q_n,n}. \tag{2.25}$$

But, the union of sets described above is not disjoint. The elements of the set $\mathcal{A}_{q_n,n}^{T_1} \times \mathcal{A}_{q_n^*,n}^{T_2}$ are being counted twice in (2.25). Hence, we can write, for all n sufficiently large,

$$\mu^2(\mathcal{A}_{q_n,n}) \geq \mu(\mathcal{A}_{q_n,n}^{T_1})\mu(U_n^{T_2}) + \mu(\mathcal{A}_{q_n^*,n}^{T_2})\mu(U_n^{T_1}) - \mu(\mathcal{A}_{q_n,n}^{T_1})\mu(\mathcal{A}_{q_n^*,n}^{T_2}).$$

Using O'Brien's formula, we obtain a lower bound for θ,

$$\begin{aligned}
\lim_{n\to\infty} \theta_n &= \lim_{n\to\infty} \frac{\mu^2(\mathcal{A}_{q_n,n})}{\mu^2(U_n)} \\
&\geq \lim_{n\to\infty} \frac{\mu(\mathcal{A}_{q_n,n}^{T_1})\mu(U_n^{T_2}) + \mu(\mathcal{A}_{q_n^*,n}^{T_2})\mu(U_n^{T_1}) - \mu(\mathcal{A}_{q_n,n}^{T_1})\mu(\mathcal{A}_{q_n^*,n}^{T_2})}{\mu(U_n^{T_1})\mu(U_n^{T_2})} \\
&\geq \theta_1 + \theta_2 - \theta_1\theta_2.
\end{aligned}$$

Now, we prove (2) and assume that θ_1 and both $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ hold for $(X_n)_n$. Note that, by (2.24), we have

$$\begin{aligned}\lim_{n\to\infty} \theta_n &\geq \lim_{n\to\infty} \frac{\mu(\mathcal{A}^{T_1}_{q_n,n})\mu(U_n^{T_2}) + \mu(\mathcal{A}^{T_2}_{q_n,n})\mu(U_n^{T_1}) - \mu(\mathcal{A}^{T_1}_{q_n,n})\mu(\mathcal{A}^{T_2}_{q_n,n})}{\mu(U_n^{T_1})\mu(U_n^{T_2})} \\ &= \theta_1 + \lim_{n\to\infty} \frac{\mu(\mathcal{A}^{T_2}_{q_n,n})}{\mu(U_n^{T_2})}\left(1 - \frac{\mu(\mathcal{A}^{T_1}_{q_n,n})}{\mu(U_n^{T_1})}\right).\end{aligned}$$

By assumption $\theta_1 = \lim_{n\to\infty} \frac{\mu(\mathcal{A}^{T_1}_{q_n,n})}{\mu(U_n^{T_1})} = 1$ and, since $0 \leq \frac{\mu(\mathcal{A}^{T_2}_{q_n,n})}{\mu(U_n^{T_2})} \leq 1$ by definition of the sets $U_n^{T_2}$ and $\mathcal{A}^{T_2}_{q_n,n}$, we conclude that

$$\lim_{n\to\infty} \theta_n \geq 1.$$

But by definition of θ_n it is clear that $\theta_n \leq 1$, for all $n \in \mathbb{N}$. It follows that the EI exists and $\theta = \lim_{n\to\infty} \theta_n = 1$.

If $\theta_2 = 1$ and $Д_{q_n^*}(u_n, w_n)$ and $Д'_{q_n^*}(u_n, w_n)$ hold for $(X_n)_n$, instead, then the same argument with the necessary adjustments would also lead to the conclusion that the EI exists and is equal to 1. □

3. Fractal Landscapes in two-dimensional Spaces

The starting point of this section is the ternary Cantor set that we denote by $\mathcal{C}$. Recall that, to construct this set we start with $\mathcal{C}_0 := [0, 1]$ and, by removing the middle third of this interval, we construct the first approximation of $\mathcal{C}$ designated by $\mathcal{C}_1$. From this point we start an iterative process, where the approximation $\mathcal{C}_n$ is constructed by removing the middle third of each connected component of $\mathcal{C}_{n-1}$. The result of this process, illustrated in Figure 1, is the set $\mathcal{C} = \cap_{n\geq 1}\mathcal{C}_n$.

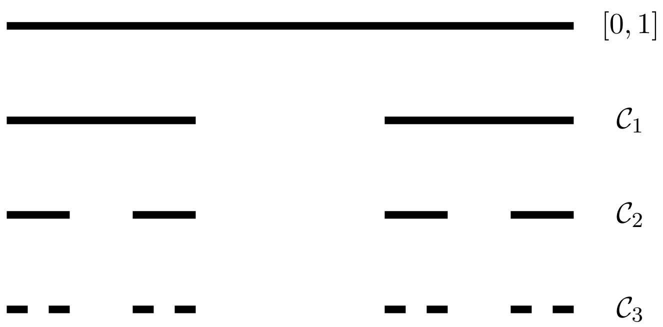

Figure 1. The construction of the ternary Cantor set.

For each $n \in \mathbb{N}$, let $B_n := \mathcal{C}_{n-1} \setminus \mathcal{C}_n$. We point out that, the sets B_n correspond to the gaps of the set $\mathcal{C}$ that are formed at the n-th approximation of its construction. The Cantor ladder function, φ is defined as

$$\varphi(x) = \begin{cases} n, & \text{if } x \in B_n,\ n = 1, 2, 3 \dots \\ \infty, & \text{otherwise.} \end{cases} \tag{3.1}$$

The function $\varphi(x)$ achieves its maximum value, ∞, if and only if x belongs to the Cantor set.

From the Cantor set $\mathcal{C}$, we can define a new fractal set given by $\mathfrak{C} := \mathcal{C} \times \mathcal{C}$. This set, usually called Cantor dust, is a self-similar set contained in the two-dimensional space $[0,1] \times [0,1]$. The Cantor dust can be seen as the final product of an algorithmic construction similar to the one presented for $\mathcal{C}$. One can define the n-th approximation of $\mathfrak{C}$, denoted by $\mathfrak{C}_n$, as the product $\mathcal{C}_n \times \mathcal{C}_n$. The set $\mathfrak{C}$ can then be described as $\cap_{n\geq 1}\mathfrak{C}_n$.

Using the Cantor ladder function we construct an observable, $\psi : [0,1]^2 \to \mathbb{R}$, whose maximal set $\mathcal{M}$ is exactly $\mathfrak{C}$ in the following way:

$$\psi(x,y) = \begin{cases} n, & \text{if } \min(\varphi(x), \varphi(y)) = n \\ \infty, & \text{otherwise.} \end{cases} \tag{3.2}$$

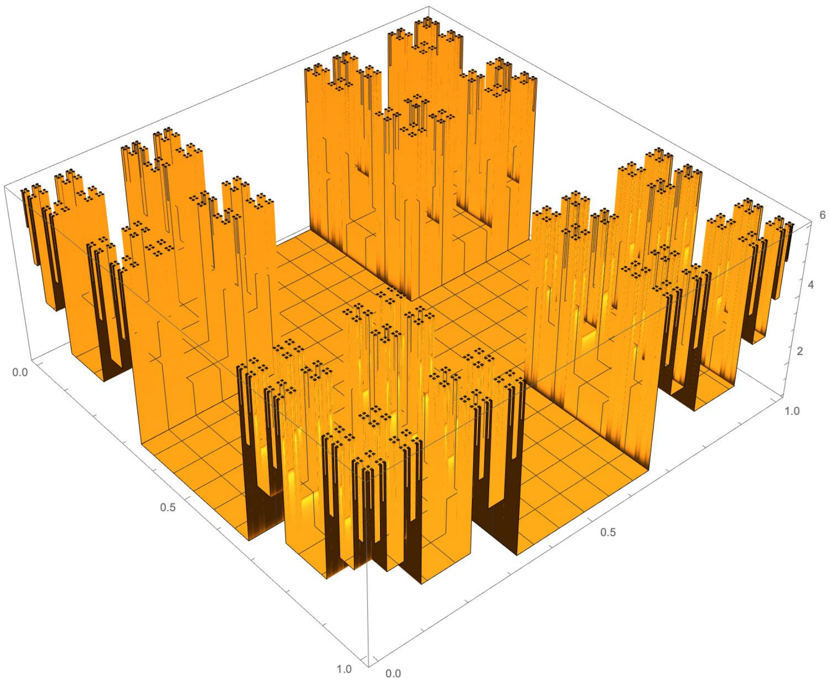

FIGURE 2. The observable ψ.

The two-dimensional dynamical systems that we will consider are given by,

$$\begin{aligned} T\colon [0,1]^2 &\longrightarrow [0,1]^2 \\ (x,y) &\mapsto (m_1 \cdot x \mod 1, m_2 \cdot y \mod 1), \end{aligned} \tag{3.3}$$

where $m_1, m_2 \in \mathbb{N}$.

From this point on, we will use Leb to denote the Lebesgue measure in $\mathbb{R}$. The systems in (3.3) preserve the product measure Leb $\times$ Leb, which we will denote by Leb^2 and belong to a larger class of maps defined by Saussol in [24]. Moreover, following the setting presented by Saussol, it is possible to prove that these systems have decay of correlations for quasi-Hölder observables against $L^1(\mathrm{Leb}^2)$ where the parameter α associated with the norm of the space of quasi-Hölder functions is equal to 1.

To be coherent with what was written in the last section, the natural choice for the maps T_1 and T_2 are the uniformly expanding maps $m_1 \cdot x \mod 1$ and $m_2 \cdot y \mod 1$, respectively. For these maps it was proved in [13] the existence of a limiting extreme value law for the stochastic process $(X_n^1)_n = (\varphi \circ T_1^n)_n$. Moreover, we were able to link the value of the EI to the compatibility between the maximal set of φ and the dynamics, *i.e*, how relevant is $T_1^{-j}(\mathcal{M}) \cap \mathcal{M}$ when compared with $\mathcal{M}$ itself.

Since for the observable φ, the maximal set is equal to $\mathcal{C}$ and using the fact that $\mathcal{C}$ has a thickness not less than 1, the relevance of the intersection $T_1^{-j}(\mathcal{M}) \cap \mathcal{M}$ was measured in terms of the comparison between the box dimension of $\mathcal{C}$ and $T_1^{-j}(\mathcal{C}) \cap \mathcal{C}$.

It was shown that if the box dimension of $T_1^{-j}(\mathcal{C}) \cap \mathcal{C}$ was lower than the box dimension of the Cantor set, which happens if $m_1 \neq 3^k$ for all $k \in \mathbb{N}$, this would lead to an EI equal to 1. However, when $T_1^{-j}(\mathcal{C}) = \mathcal{C}$, which happens if $m_1 = 3^k$ for some $k \in \mathbb{N}$, then the set $\mathcal{C}$ was playing the role of a periodic point. This creates a clustering effect due to the recurrence of the maximal set to itself and consequently the EI would be strictly smaller than 1 and equal to $1 - (2/3)^k$.

The main goal of this paper is to show that similar results hold when dealing with an observable whose maximal set is contained in a 2-dimensional space, as is the case of $\mathfrak{C}$. We will demonstrate that if $m_1 = 3^{k_1}$ and $m_2 = 3^{k_2}$, for some k_1, k_2 in $\mathbb{N}$, then there exists full compatibility between $\mathfrak{C}$ and $T^{-j}(\mathfrak{C})$.

Theorem 3.1. *Consider the stochastic process $(X_n)_{n\in\mathbb{N}}$ given as in* (2.1) *for the observable function ψ, defined in* (3.2)*, and the dynamical system T, defined in* (3.3)*, with $m_1 = 3^{k_1}$ and $m_2 = 3^{k_2}$, for some k_1, k_2 in $\mathbb{N}$ satisfying*

$$1 + \frac{min\{k_1, k_2\}}{max\{k_1, k_2\}} > \log_3(4). \tag{3.4}$$

Consider a sequence of thresholds $(u_n)_{n\in\mathbb{N}}$ such that $u_n = n$ and a sequence of times $(w_n)_{n\in\mathbb{N}}$, such that $w_n = \left\lfloor \tau\,(3/2)^{2n} \right\rfloor$.

Then, condition (2.3) *holds and*

$$\lim_{n\to\infty} Leb^2(M_{w_n} \leq n) = \mathrm{e}^{-\left(1-\frac{2^{k_1+k_2}}{3^{k_1+k_2}}\right)\tau}.$$

On other hand, we will show that when both m_1 and m_2 cannot be written as 3^k for any integer k, the compatibility between T and $\mathfrak{C}$ is broken.[1] The result is an insignificant clustering effect which, in the limit, will lead to an EI equal to 1.

Theorem 3.2. *Consider the stochastic process $(X_n)_{n\in\mathbb{N}}$ given as in* (2.1) *for the observable function ψ, defined in* (3.2)*, and the dynamical system T defined in* (3.3)*, where m_1 and m_2*

[1]We recall that, by compatibility between the maximal set and the dynamics, we mean how much of the maximal set is being preserved upon iteration of the set by the dynamics. High compatibility means that a significant portion of the maximal set is preserved by the dynamics, which leads to a smaller EI and the existence of clustering.

cannot be written as 3^k*, for any* $k \in \mathbb{N}$*. Set* $u_n = n$ *and let* $(w_n)_{n\in\mathbb{N}}$ *be a sequence of times, such that* $w_n = \left\lfloor \tau\,(3/2)^{2n} \right\rfloor$*.*

Then, condition (2.3) *holds and*

$$\lim_{n\to\infty} Leb^2(M_{w_n} \leq n) = \mathrm{e}^{-\tau}.$$

We will also be considering the case where the map T is such that $m_1 = 3^k$, for some $k \in \mathbb{N}$, and m_2 cannot be written in the form 3^j for any $j \in \mathbb{N}$. This case represents a middle ground between the two theorems presented above. On one hand, we should expect clustering to appear, due to compatibility between T_1 and the Cantor set $\mathcal{C}$, however, the incompatibility between T_2 and $\mathcal{C}$ is enough to guarantee the absence of clustering in the stochastic process $(X_n)_n$.

Theorem 3.3. *Consider* $(X_n)_{n\in\mathbb{N}}$ *to be the stochastic process given by* (2.1) *for the dynamical system* T *defined in* (3.3)*, where* $m_1 = 3^k$ *and* m_2 *cannot be written as* 3^j*, for any* $j \in \mathbb{N}$*. Moreover, assume that* k *and* m_2 *satisfy the inequality*

$$1 + k\frac{\log 3}{\log m_2} > \log_3 4. \tag{3.5}$$

Set $u_n = n$ *and let* $(w_n)_{n\in\mathbb{N}}$ *be a sequence of times, such that* $w_n = \left\lfloor \tau\,(3/2)^{2n} \right\rfloor$*.*

Then, condition (2.3) *holds and*

$$\lim_{n\to\infty} Leb^2(M_{w_n} \leq n) = \mathrm{e}^{-\tau}.$$

The strategy used to prove these statements follows the same lines taken in [13], where we proved a result that gives sufficient conditions for the existence of an EVL when $\mathcal{M}$ is a dynamically generated Cantor set Λ, *i.e.*, a Cantor set which is the attractor of a certain *Iterated Function System* (IFS). This IFS was then used to construct a map $\bar{T}$ such that, for all j, $\bar{T}^{-j}(\Lambda) = \Lambda$, which results in a closed formula for the value of the EI. We build up on these results to achieve similar conditions that guarantee the existence of an EVL with a non-trivial EI when $\mathcal{M}$ is the product of two dynamically generated Cantor sets.

To prove Theorem 3.2 and Theorem 3.3, we will rely on Theorem 2.6. This result, when used in conjunction with the previous mentioned results presented in [13], is sufficient to guarantee the existence of an EVL with θ equal to 1.

Remark 3.4. It will be clear in the proof of Theorems 3.1 and 3.3 that the technical conditions (3.4) and (3.5) assure the existence of sufficient expansion of the maps to prove conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ and establish the existence of EVL. However, this technical conditions are not important in the results regarding the value of the EI. The considerations made here regarding θ should hold in a much broader context.

4. Clustering and Fractal Landscapes

In this section, we prove a general result, Theorem 4.2, which gives the existence of an Extreme Value Law, when $\mathcal{M}$ is a product of dynamically generated Cantor sets and the thriving dynamics is compatible with that structure and, therefore, responsible for the appearance of clustering. Theorem 3.1 follows from Theorem 4.2, when $\mathcal{M} = \mathcal{C} \times \mathcal{C}$ is taken as the particular case corresponding to the Cantor dust set.

We will start by defining the concept of dynamically generated Cantor sets, which can be identified as the survivor sets of a certain dynamics. This identification allows to construct dynamical systems that are compatible with the fractal structure of these sets, leading to the appearance of an EI smaller than one.

Consider a finite family of C^1 diffeomorphisms on $[0,1]$, $\mathfrak{F} = \{f_1, f_2, \dots, f_s\}$, where

$$|f_i(x) - f_i(y)| \le \lambda_i |x - y|,$$

for some ratio $\lambda_i < 1$.

Assume further, that the intersection of the images of any two of these contractions is disjoint. When defined in this way, the family $\mathfrak{F}$ defines an Iterated Function System (IFS), that satisfies enough regularity conditions for the existence of a unique compact set, Λ, satisfying the equation

$$\Lambda = \cup_{i=1}^{s} f_i(\Lambda).$$

For more details on IFS we refer to [8, Chapter 9].

This set Λ can be called a dynamically generated Cantor set, since it can be identified as the survivor set of a dynamical system $G : \mathbb{R} \to \mathbb{R}$ defined as

$$G(x) = \begin{cases} f_i^{-1}(x), & \text{if } x \in f_i([0,1]) \\ 2, & \text{otherwise} \end{cases}.$$

Under this interpretation, Λ can be technically described as the set of points in $[0,1]$ whose orbit by $G(x)$ never leaves this interval, that is

$$\Lambda = \{x \in [0,1] \colon\ G^n(x) \in [0,1],\ \text{for all } n \in \mathbb{N}\}.$$

Let $\Lambda_0 = [0,1]$ and set for all $n \in \mathbb{N}$,

$$\Lambda_n = G^{-1}(\Lambda_{n-1}) = \{x \in [0,1] \colon G^l(x) \in [0,1],\ \text{for all } l = 1, \dots, n\}.$$

When defined this way, Λ_n represents the n-th approximation to Λ allowing us to write that $\Lambda = \cap_{n \ge 0} \Lambda_n$.

The setting above was the starting point that allowed us, in [13], to define a dynamics $\bar{T}$ that is fully compatible with the set Λ. Such dynamics is constructed using the dynamical system $G(x)$ in the following way.

Set $J_i = f_i([0,1])$ and let I denote a connected component of $[0,1] \setminus \cup_{i=1}^{s} J_i$ and consider $g_I(x)$ to be a linear function that maps I onto $[0,1]$. With this notation, we define $F : [0,1] \to [0,1]$ as

$$F(x) = \begin{cases} G(x), & \text{if } x \in \cup_{i=1}^{s} J_i \\ g_I(x), & \text{if } x \in I, \quad \text{where } I \text{ is a connected component of } [0,1] \setminus \bigcup_{i=1}^{s} J_i \end{cases}.$$

The function F is a piecewise uniformly expanding map and therefore admits an absolutely continuous invariant measure μ. Also, accordingly to [5, Corollary 8.3.1] this maps have decay of correlations of BV observables against $L^1(\mu)$.

The dynamics $\bar{T}$ is constructed by setting $\bar{T} = F^k$, for some $k \in \mathbb{N}$. In [13] it was proved a result that asserts the compatibility of $\bar{T}$ with Λ, which we transcribe here.

Lemma 4.1. *If $j \leq n/k$, then, $\bar{T}^{-j}(\Lambda_n) \cap \Lambda_n = \Lambda_{n+kj}$.*

Note that, in particular, Lemma 4.1 implies that $\bar{T}^{-j}(\Lambda) = \Lambda$. As pointed out before, it was this recurrence effect that resulted in an EI smaller than 1 for the case where $\mathcal{M} = \Lambda$ is defined to be the maximal set of the observable.

4.1. **Extreme laws and product of dynamically generated Cantor sets.** In what follows, we consider two dynamically generated sets, Λ_1 and Λ_2, with associated compatible maps denoted, respectively, by $T_1 = F_1^{k_1}$ and $T_2 = F_2^{k_2}$, for $k_1, k_2 \in \mathbb{N}$. The direct product of Λ_1 and Λ_2, represented by Σ, is a fractal set contained in $[0,1]^2$. The n-th approximation to Σ is defined as $\Sigma_n = \Lambda_{1,n} \times \Lambda_{2,n}$, where $\Lambda_{1,n}$ and $\Lambda_{2,n}$ represent the n-th approximation of Λ_1 and Λ_2, respectively.

Consider the map $T : [0,1]^2 \to [0,1]^2$, given by

$$T(x,y) = (T_1(x), T_2(y)). \tag{4.1}$$

Due to the product structure present in Σ, we can use Lemma 4.1 to establish the compatibility between T and Σ.

Let $k = \max\{k_1, k_2\}$, we claim that, if $j \leq n/k$, then

$$T^{-j}(\Sigma_n) \cap \Sigma_n = \Lambda_{1,n+k_1 j} \times \Lambda_{2,n+k_2 j}. \tag{4.2}$$

Using the properties of the direct product we write that

$$T^{-j}(\Sigma_n) \cap \Sigma_n = \left(T_1^{-j}(\Lambda_{1,n}) \cap \Lambda_{1,n}\right) \times \left(T_2^{-j}(\Lambda_{2,n}) \cap \Lambda_{2,n}\right). \tag{4.3}$$

Applying Lemma 4.1 to (4.3), we obtain that

$$T^{-j}(\Sigma_n) \cap \Sigma_n = \Lambda_{1,n+k_1 j} \times \Lambda_{2,n+k_2 j}$$

and the claim follows.

Again, these results imply that $T^{-j}(\Sigma) = \Sigma$ and clustering is to be expected when Σ is defined to be the maximal set of the considered observable.

To construct an observable whose maximal set is Σ, we use the function $\hat{\varphi}_\Lambda$ defined as,

$$\hat{\varphi}_\Lambda(x) = \begin{cases} n, & \text{if } x \in \Lambda_n \setminus \Lambda_{n+1}, \ n = 1, 2, 3 \dots \\ \infty, & \text{otherwise.} \end{cases}$$

The observable that we will be considering depends on the chosen sets Λ_1 and Λ_2 and is constructed as

$$\hat{\psi}(x,y) = \begin{cases} n, & \text{if } \min(\hat{\varphi}_{\Lambda_1}(x), \hat{\varphi}_{\Lambda_2}(y)) = n \\ \infty, & \text{otherwise.} \end{cases} \tag{4.4}$$

The maximal set of this observable is, precisely, the set Σ. Moreover, the functions φ and ψ defined in (3.1) and (3.2), respectively, are particular cases of $\hat{\varphi}$ and $\hat{\psi}$, when Λ_1 and Λ_2 coincide with the ternary Cantor set $\mathcal{C}$.

The map T is a multidimensional piecewise expanding map as defined by Saussol in [24]. Consequently, it has decay of correlations for quasi-Hölder observables against L^1, with respect to the invariant measure $\mu \times \mu$ denoted by μ^2. This allows us to prove the following Theorem that establishes sufficient conditions for the existence of an EVL, when the maximal set of the observable is the product of two dynamically generated Cantor sets. Furthermore, if the EI exists, this Theorem provides a closed formula for its calculation.

Theorem 4.2. *Consider $(X_n)_{n\in\mathbb{N}}$ to be the stochastic process constructed as in (2.1) for the dynamical system T in (4.1) and the observable $\hat{\psi}$ in (4.4). Consider a sequence of thresholds $(u_n)_{n\in\mathbb{N}}$ and a sequence $(w_n)_{n\in\mathbb{N}}$ such that,*

$$w_n = \left\lfloor \tau \left[\mu(\Lambda_{1,\lfloor u_n\rfloor})\mu(\Lambda_{2,\lfloor u_n\rfloor})\right]^{-1} \right\rfloor.$$

Let $k = max\{k_1, k_2\}$ and consider a sequence $(q_n)_{n\in\mathbb{N}}$ as in (2.5) satisfying $1 \le q_n \le u_n/k$.

Assume there exists $(t_n)_{n\in\mathbb{N}}$, where $t_n = o(w_n)$, such that the following conditions hold:

(1) $\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{A}_{q_n,n}}\|_\alpha w_n \rho(t_n) = 0$ *for some* $0 < \alpha \le 1$

(2) $\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{BV} \mu\left(U_n^{T_2}\right) \sum_{j=q_n}^{\infty} \rho_j^1 = 0$

(3) $\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{BV} \mu\left(U_n^{T_1}\right) \sum_{j=q_n}^{\infty} \rho_j^2 = 0$

(4) $\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{BV} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{BV} \sum_{j=q_n}^{\infty} \rho_j^1\rho_j^2 = 0.$

Furthermore, assume that there exists $0 \le \theta \le 1$ such that

$$\theta = \lim_{n\to\infty} \frac{\mu^2(\Sigma_{\lfloor u_n\rfloor} \setminus (\Lambda_{1,\lfloor u_n\rfloor+k_1} \times \Lambda_{2,\lfloor u_n\rfloor+k_2}))}{\mu^2(\Lambda_{1,\lfloor u_n\rfloor} \times \Lambda_{2,\lfloor u_n\rfloor})}.$$

Then,

$$\lim_{n\to\infty} \mu^2(M_{w_n} \le n) = \mathrm{e}^{-\theta\tau}.$$

Proof. Start by considering a sequence of thresholds $(u_n)_{n\in\mathbb{N}}$. Due to the definition of φ_Λ,

$$U_n^{T_1} = \{x \in [0,1] : \hat{\varphi}_{\Lambda_1}(x) > u_n\} = \Lambda_{1,\lfloor u_n\rfloor}$$

and

$$U_n^{T_2} = \{y \in [0,1] : \hat{\varphi}_{\Lambda_2}(y) > u_n\} = \Lambda_{2,\lfloor u_n\rfloor}.$$

By construction of $\hat{\psi}$, the point $(x, y) \in [0,1]^2$ satisfies the inequality, $\hat{\psi}(x, y) > u_n$, if and only if $\varphi_{\Lambda_1}(x) > u_n$ and $\varphi_{\Lambda_2}(y) > u_n$. Hence,

$$U_n = \left\{(x, y) \in [0,1]^2 : \hat{\psi}(x, y) > u_n\right\} = \Lambda_{1,\lfloor u_n\rfloor} \times \Lambda_{2,\lfloor u_n\rfloor} = U_n^{T_1} \times U_n^{T_2}.$$

Moreover, condition (2.3) is verified since,

$$w_n\mu^2(U_n) = \left\lfloor \tau(\mu(\Lambda_{1,\lfloor u_n\rfloor})\mu(\Lambda_{2,\lfloor u_n\rfloor}))^{-1} \right\rfloor \mu(\Lambda_{1,\lfloor u_n\rfloor})\mu(\Lambda_{2,\lfloor u_n\rfloor}) \xrightarrow[n\to\infty]{} \tau.$$

The maps T_1 and T_2 have decay of correlations for observables in BV against $L^1(\mu)$. Moreover, the dynamics T has decay of correlations for quasi-Hölder observables against $L^1(\mu \times \mu)$. Denote the rate functions of T_1, T_2 and T by ρ^1, ρ^2 and ρ, respectively. Together with Theorem 2.5, we obtain that conditions (1) through (4) guarantee that conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ hold.

Noting that, $\Lambda_{1,i+i^*} \times \Lambda_{2,j+j^*} \subseteq \Lambda_{1,i} \times \Lambda_{2,j}$, for all $i, i^*, j, j^* \in \mathbb{N}$ and using relation (4.2), we establish that

$$\mathcal{A}_{q_n,n} = \Sigma_{\lfloor u_n \rfloor} \setminus \left(\Lambda_{1,\lfloor u_n \rfloor + k_1} \times \Lambda_{2,\lfloor u_n \rfloor + k_2}\right).$$

for all q_n satisfying $1 \leq q_n \leq u_n/k$.

The value of the EI follows from O'Brien's formula. □

4.2. **Application to the Cantor dust.** We are now in conditions to use Theorem 4.2 to prove Theorem 3.1.

Proof of Theorem 3.1. The set $\mathcal{C}$ is dynamically generated by the IFS $f_1 = x/3$ and $f_2 = x/3 + 2/3$. Theorem 3.1 follows by taking $\Lambda_1 = \Lambda_2 = \mathcal{C}$ and applying Theorem 4.2.

The map considered is $T = (T_1(x), T_2(y))$ where $T_1(x) = 3^{k_1}x \mod 1$ and $T_2(x) = 3^{k_2}x \mod 1$, for $k_1, k_2 \in \mathbb{N}$ satisfying

$$1 + \frac{\min\{k_1, k_2\}}{\max\{k_1, k_2\}} > \log_3(4).$$

Observe that, the invariant measure associated with T is Leb^2 and the invariant measure associated with T_1 and T_2 is Leb. Set $u_n = n$, $w_n = \lfloor \tau(2/3)^{-2n} \rfloor$ and $q_n = \lfloor n/k \rfloor$, where $k = \max\{k_1, k_2\}$ and observe that,

$$U_n^{T_1} = U_n^{T_2} = \mathcal{C}_n.$$

Consequently,

$$\mu(U_n^{T_1}) = \mu(U_n^{T_2}) = \left(\frac{2}{3}\right)^n \qquad \text{and} \qquad \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{BV} = \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{BV} \leq 2^{n+1}.$$

The maps T_1 and T_2 have decay of correlations of BV observables against $L^1(\text{Leb})$, with rate functions $\rho_n^1 = (1/3)^{k_1 n}$ and $\rho_n^2 = (1/3)^{k_2 n}$. It is necessary to show that conditions (2) through (4) of Theorem 4.2 hold.

Making the necessary substitutions, there exist constants $C, C', C'' > 0$ such that,

$$\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{BV} \text{Leb}(U_n^{T_1}) \sum_{j=q_n}^{\infty} \rho_j^1 \leq C \lim_{n\to\infty} \frac{4^n}{3^{n(1+k_1/k)}} = 0$$

$$\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{BV} \text{Leb}(U_n^{T_2}) \sum_{j=q_n}^{\infty} \rho_j^2 \leq C' \lim_{n\to\infty} \frac{4^n}{3^{n(1+k_2/k)}} = 0$$

and

$$\lim_{n\to\infty} \left\|\mathbf{1}_{U_n^{T_1}}\right\|_{BV} \left\|\mathbf{1}_{U_n^{T_2}}\right\|_{BV} \sum_{j=q_n}^{\infty} \rho_j^1 \rho_j^2 \leq C'' \lim_{n\to\infty} \frac{4^n}{3^{n(k_1/k+k_2/k)}} = 0.$$

Therefore, considering the restrictions imposed on k_1 and k_2, conditions $(2) - (4)$ hold.

The next step is to prove that condition (1) of Theorem 4.2 holds. For that purpose, it is necessary to estimate $\|\mathbf{1}_{\mathcal{A}_{q_n,n}}\|_\alpha$. Let $C(A)$ denote the number of connected components of a set A and P denote the maximum perimeter of the connected components of $\mathcal{A}_{q_n,n}$. Then, for a given $0 < \alpha \leq 1$ and $\epsilon_0 > 0$,

$$|\mathbf{1}_{\mathcal{A}_{q_n,n}}|_\alpha \leq \sup_{0<\epsilon\leq\epsilon_0} \epsilon^{-\alpha} \left(\epsilon C(\mathcal{A}_{q_n,n}) P\right).$$

Hence, we achieve that

$$|\mathbf{1}_{\mathcal{A}_{q_n,n}}|_\alpha \leq PC(\mathcal{A}_{q_n,n})$$

and

$$\|\mathbf{1}_{\mathcal{A}_{q_n,n}}\|_\alpha \leq \mathrm{Leb}^2(\mathcal{A}_{q_n,n}) + PC(\mathcal{A}_{q_n,n}).$$

We now estimate the values of $\mathrm{Leb}^2(\mathcal{A}_{q_n,n})$, $C(\mathcal{A}_{q_n,n})$ and P. Set $\Sigma = \mathfrak{C}$ then, following the proof of Theorem 4.2, we can write,

$$\mathcal{A}_{q_n,n} = \mathfrak{C}_n \setminus (\mathcal{C}_{n+k_1} \times \mathcal{C}_{n+k_2})$$

and we obtain

$$\mathrm{Leb}^2(\mathcal{A}_{q_n,n}) = (2/3)^{2n}(1-(2/3)^{k_1+k_2}).$$

Each connected component of the set $\mathfrak{C}_n$ is a square with side length equal to $1/3^n$ and each connected component of $\mathcal{C}_{n+k_1} \times \mathcal{C}_{n+k_2}$ is a rectangle, where the side lengths are $1/3^{n+k_1}$ and $1/3^{n+k_2}$. This implies that, each candidate to connected component of $\mathcal{A}_{q_n,n}$ is a square, with side length $1/3^n$, with rectangular holes where the length of the sides of each hole are $1/3^{n+k_1}$ and $1/3^{n+k_2}$. Figure 3 aims to represent this reasoning when $k_1 = 1$ and $k_2 = 2$.

It is necessary to show that the candidates to connected components of $\mathcal{A}_{q_n,n}$ identified above are indeed connected. To make such verification, it is enough to note that, due to scaling properties of the ternary Cantor set, the pattern of the rectangular holes in each square is similar to the scheme of the connected components of the set $\mathcal{C}_{k_1} \times \mathcal{C}_{k_2}$. Since $\mathcal{C}_{k_1}$ and $\mathcal{C}_{k_2}$ always have a gap between each interval that belongs to the set, this is sufficient to show that

$$C(\mathcal{A}_{q_n,n}) = C(\mathfrak{C}_n) = 4^n.$$

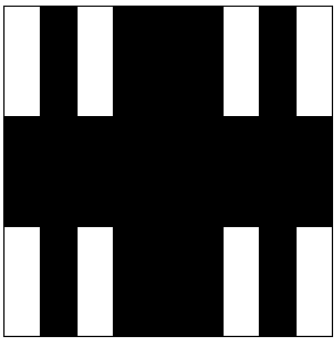

FIGURE 3. Representation of each connected component of the set $\mathcal{A}_{q_n,n}$ when $k_1 = 1$ and $k_2 = 2$. The white rectangular holes in the picture correspond to the connected components of the set $\mathcal{C}_{n+1} \times \mathcal{C}_{n+2}$ that we delete from each connected component of $\mathfrak{C}_n$. The remaining part of each connected component of $\mathfrak{C}_n$ forms a connected component of the set $\mathcal{A}_{q_n,n}$.

Moreover, the regularity of the connected components of $\mathcal{A}_{q_n,n}$ allows to estimate its maximum perimeter. Each connected component of $\mathcal{A}_{q_n,n}$ is a square with $2^{k_1+k_2}$ rectangular holes. Since each hole is a rectangle contained in $[0,1]\times[0,1]$ the maximum perimeter of each hole is 4. Hence, the maximum perimeter of each connected component is $4(2^{k_1+k_2}+1)$.

Making the necessary substitutions, we obtain that

$$\|\mathbf{1}_{\mathcal{A}_{q_n,n}}\|_\alpha \le \mathrm{Leb}^2(\mathcal{A}_{q_n,n}) + 4(2^{k_1+k_2}+1)C(\mathcal{A}_{q_n,n}). \tag{4.5}$$

The map T has decay of correlations for quasi-Hölder observables against $L^1(\mathrm{Leb}^2)$ with rate function $\rho_n = 1/3^n$. Let $t_n = n^2$, then $t_n = o(w_n)$ and there exists a constant $C''' > 0$ such that

$$\begin{aligned}\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{A}_{q_n,n}}\|_\alpha w_n \rho_{t_n} &\le \lim_{n\to\infty}\Big(\mathrm{Leb}^2(\mathcal{A}_{q_n,n}) + 4(2^{k_1+k_2}+1)C(\mathcal{A}_{q_n,n})\Big) w_n\rho_{t_n}\\ &\le \lim_{n\to\infty}(C'''(2/3)^{2n} + (2^{k_1+k_2}+1)4^{n+1})(\tau(2/3)^{-2n})\frac{1}{3^{n^2}}\\ &= 0\end{aligned}$$

and condition (1) of 4.2 holds.

To finish the proof, we use O'Brien's formula to establish that

$$\begin{aligned}\theta &= \lim_{n\to\infty}\frac{\mathrm{Leb}^2(\mathfrak{C}_n \setminus (\mathcal{C}_{n+k_1}\times\mathcal{C}_{n+k_2}))}{\mathrm{Leb}^2(\mathfrak{C}_n)}\\ &= \lim_{n\to\infty}\frac{(2/3)^{2n}(1-(2/3)^{k_1+k_2})}{(2/3)^{2n}}\\ &= 1-(2/3)^{k_1+k_2}.\end{aligned}$$

□

5. Absence of Clustering for maximal sets in two-dimensional spaces

The starting point of this section is the stochastic process $(X_n)_n$ constructed using the setting in Section 2.1 and the lower bound, obtained in Theorem 2.6, for the corresponding EI, θ. In there, we proved the existence of a link between the level of clustering appearing in $(X_n)_n$ and the clustering present in the processes $(X_n^1)_n$ and $(X_n^2)_n$. Assuming the existence of the limits in (2.20) then θ satisfies

$$\theta > \theta_1 + \theta_2 - \theta_1\theta_2. \tag{5.1}$$

The simplicity of such formula allows to draw very useful conclusions regarding the level of clustering associated with $(X_n)_n$ based on θ_1 and θ_2. In fact, as we can see in Figure 5, θ will be 1 if θ_1 or θ_2 is also 1.

This implies that the absence of clustering in either $(X_n^1)_n$ or $(X_n^2)_n$ will imply an absence of clustering in the process $(X_n)_n$. In fact, for $\theta_1, \theta_2 \in]0,1[$, we have

$$\theta_1 < \theta_1+\theta_2-\theta_1\theta_2 \qquad \text{and} \qquad \theta_2 < \theta_1+\theta_2-\theta_1\theta_2$$

implying that the level of clustering appearing in X_n is always smaller than the level of clustering appearing in $(X_n^1)_n$ or $(X_n^2)_n$. This smoothing effect of the clustering can be linked to the product structure present in $\mathcal{M}_1\times\mathcal{M}_2$. Due to this nature of the maximal set, if

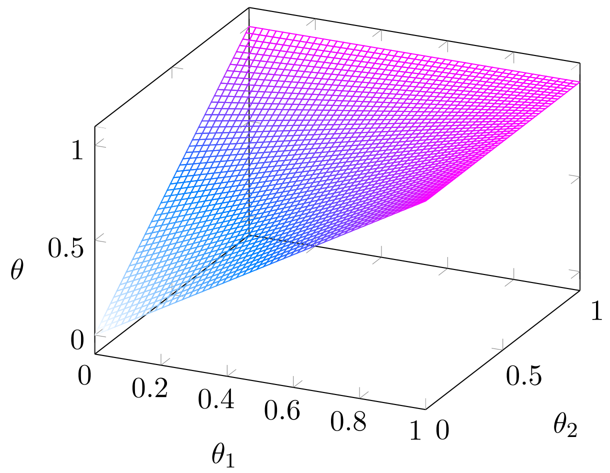

FIGURE 4. Graph of the lower bound for θ as a function of θ_1 and θ_2.

$T_1^{-j}(\mathcal{M}_1)\cap\mathcal{M}_1$ or $T_2^{-j}(\mathcal{M}_2)\cap\mathcal{M}_2$ is not relevant, this is enough to guarantee that $T^{-j}(\mathcal{M}_1\times\mathcal{M}_2)\cap\mathcal{M}_1\times\mathcal{M}_2$ is also not relevant when compared with $\mathcal{M}_1\times\mathcal{M}_2$ resulting in a low level of clustering appearing in $(X_n)_n$. This fact will be the key to prove Theorem 3.2 and Theorem 3.3.

For the remaining of this section, let $\bar{T}$ represent the dynamics $mx \mod 1$, for some $m\in\mathbb{N}$. Moreover, let $(\bar{X}_n)_n$ denote the stochastic process given by $\bar{X}_n=\varphi\circ\ \bar{T}^n$, where φ is the observable maximised on the ternary Cantor set $\mathcal{C}$ introduced in (3.1). We will build up on the following result from [13].

Theorem 5.1. *Let $(X_n)_{n\in\mathbb{N}}$ be the stochastic process given by $\bar{X}_n=\varphi\circ\ \bar{T}^n$. Assume that m is not a power of 3 and consider a sequence of thresholds $(u_n)_{n\in\mathbb{N}}$ such that $u_n=n$, a sequence of times $(w_n)_{n\in\mathbb{N}}$ such that $w_n=\lfloor\tau\,(3/2)^n\rfloor$ and a sequence $(q_n)_{n\in\mathbb{N}}$ such that $q_n=\left\lceil n\frac{\log 3}{\log m}\right\rceil$. Then, condition (2.3) holds and moreover*

$$\lim_{n\to\infty} Leb^2(M_{w_n}\le n)=\mathrm{e}^{-\tau}.$$

This result was proved using a link between the EI and the box dimension of the sets $\bar{T}^{-j}(\mathcal{C})\cap\mathcal{C}$. The low relevance of $\bar{T}^{-j}(\mathcal{C})\cap\mathcal{C}$ when compared with $\mathcal{C}$ translates to a smaller box dimension of $\bar{T}^{-j}(\mathcal{C})\cap\mathcal{C}$ when compared to the box dimension of $\mathcal{C}$. This difference allowed to compute an estimate for the value of $\mathrm{Leb}(\bar{T}^{-j}(\mathcal{C}_n)\cap\mathcal{C}_n)$ for n sufficiently large. In fact, if $3^{-n}\le m^{-j}$, then we estimated that

$$\mathrm{Leb}(\bar{T}^{-j}(\mathcal{C}_n)\cap\mathcal{C}_n)\le\frac{3e^{\gamma n\log 3}}{3^n},\tag{5.2}$$

where γ satisfies the inequality, $\gamma\log 3<\log 2$.

This estimate allowed to compute the measure of the set $\mathcal{C}_n\setminus\mathcal{A}_{q_n,n}$ which was the key to show that the EI is equal to 1. We will use estimate (5.2) in conjunction with Theorem 5.1 and Theorem 2.6 to prove Theorem 3.2.

Proof of Theorem 3.2. Set $T_1=m_1x \mod 1$ and $T_2=m_2y \mod 1$, where neither m_1 or m_2 can be written as 3^k, for any $k\in\mathbb{N}$. Consider $T=(T_1(x),T_2(y))$ and define the stochastic process $X_n = \psi \circ T$. We point out that, the invariant measures associated with T and with $T_1 = m_1 x \mod 1$ and $T_2 = m_2 y \mod 1$ are Leb^2 and Leb, respectively.

Let $u_n = n$ be the sequence of thresholds and set $w_n = \left\lfloor \tau\,(3/2)^{2n} \right\rfloor$. Due to the construction of the observable ψ, we have that

$$U_n^{T_1} = U_n^{T_2} = \mathcal{C}_n \qquad \text{and} \qquad U_n = \mathfrak{C}_n.$$

Checking that,

$$w_n \text{Leb}^2(U_n) = \left\lfloor \tau\,(3/2)^{2n} \right\rfloor \left(\frac{2}{3}\right)^{2n} \xrightarrow[n\to\infty]{} \tau,$$

we obtain that condition (2.3) is verified. Now, we assume that $m_2 \geq m_1$. The case $m_1 \geq m_2$ follows similarly. Let $m_2 \geq m_1$, set $q_n = \left\lceil n \frac{\log 3}{\log m_1} \right\rceil$ and note that $q_n = o(w_n)$. For such choice of q_n, Theorem 5.1 guarantees that

$$\theta_1 = 1.$$

The map T has decay of correlations for quasi-Hölder observables against $L^1(\text{Leb}^2)$, with rate function $\rho_n = 1/(m_1)^n$. Moreover, the maps T_1 and T_2 have decay of correlations for BV observables against $L^1(\text{Leb})$, with rate functions $\rho_n^1 = 1/m_1^n$ and $\rho_n^2 = 1/m_2^n$. Hence, to prove the validity of conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ it is only necessary to check hypothesis (1) through (4) of Theorem 2.5.

Since $\|\mathbf{1}_{\mathcal{C}_n}\|_{BV} \leq 2^{n+1}$, there exists a constant $C > 0$ such that,

$$\begin{aligned}
\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{C}_n}\|_{BV} \,\text{Leb}(\mathcal{C}_n) \sum_{j=q_n}^{\infty} \rho_j^1 &\leq C \lim_{n\to\infty} \frac{4^n}{3^n} \frac{1}{m_1^{q_n}} \\
&\leq C \lim_{n\to\infty} \frac{4^n}{3^{2n}} \\
&= 0.
\end{aligned}$$

Considering that $m_2 \geq m_1$, which means that $1/m_1 \geq 1/m_2$, there exist constants $C' > 0$ and $C'' > 0$ such that,

$$\begin{aligned}
\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{C}_n}\|_{BV} \sum_{j=q_n}^{\infty} \rho_j^1 \rho_j^2 &\leq C' \lim_{n\to\infty} \frac{4^n}{m_1^{q_n} m_2^{q_n}} \\
&\leq C' \lim_{n\to\infty} \frac{4^n}{m_1^{q_n} m_1^{q_n}} \\
&\leq C' \lim_{n\to\infty} \frac{4^n}{3^{2n}} \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{C}_n}\|_{BV}\,\mathrm{Leb}(\mathcal{C}_n)\sum_{j=q_n}^{\infty}\rho_j^2 &\le C''\lim_{n\to\infty}\frac{4^n}{3^n}\frac{1}{m_2^{q_n}}\\
&\le C''\lim_{n\to\infty}\frac{4^n}{3^n}\frac{1}{m_1^{q_n}}\\
&\le C''\lim_{n\to\infty}\frac{4^n}{3^{2n}}\\
&= 0.
\end{aligned}$$

To prove hypothesis (1), we will show that exists a sequence $(t_n)_{n\in\mathbb{N}}$, such that $t_n = o(w_n)$ and

$$\lim_{n\to\infty} w_n\left(\|\mathbf{1}_{\mathfrak{C}_n}\|_\alpha \rho(t_n) + 2\mathrm{Leb}^2(\mathfrak{C}_n\setminus\mathcal{A}_{q_n,n})\right) = 0.$$

Following the same reasoning as in the proof of Theorem 3.1, we have that

$$\|\mathbf{1}_{\mathfrak{C}_n}\|_\alpha \le \mathrm{Leb}^2(\mathfrak{C}_n) + PC(\mathfrak{C}_n),$$

where P denotes the maximum perimeter of the connected components of $\mathfrak{C}_n$ and $C(\mathfrak{C}_n)$ represents the maximum number of connected components of $\mathfrak{C}_n$.
Since, $\mathrm{Leb}^2(\mathfrak{C}_n) = (2/3)^{2n}$, $C(\mathfrak{C}_n) = 4^n$ and $P = 4/3^n$, we can write that,

$$\|\mathbf{1}_{\mathfrak{C}_n}\|_\alpha \le \left(\frac{2}{3}\right)^{2n} + \frac{4^{n+1}}{3^n}. \tag{5.3}$$

It is necessary to estimate $\mathrm{Leb}^2(\mathfrak{C}_n\setminus\mathcal{A}_{q_n,n})$. For that purpose, we point out that

$$\mathrm{Leb}^2(\mathfrak{C}_n\cap T^{-q}(\mathfrak{C}_n)) = \mathrm{Leb}(\mathcal{C}_n\cap T_1^{-q}(\mathcal{C}_n))\mathrm{Leb}(\mathcal{C}_n\cap T_2^{-q}(\mathcal{C}_n)) = \Big(\mathrm{Leb}(\mathcal{C}_n\cap T_1^{-q}(\mathcal{C}_n))\Big)^2$$

and consequently,

$$\mathrm{Leb}^2\left(\bigcup_{q=1}^{q_n}\mathfrak{C}_n\cap T^{-q}(\mathfrak{C}_n)\right) \le \sum_{q=1}^{q_n}\Big(\mathrm{Leb}(\mathcal{C}_n\cap T_1^{-q}(\mathcal{C}_n))\Big)^2.$$

Using the estimate (5.2) and noting that $3^{-n}\le m_1^{-q_n}$, we get that

$$\mathrm{Leb}^2\left(\bigcup_{q=1}^{q_n}\mathfrak{C}_n\cap T^{-q}(\mathfrak{C}_n)\right) \le q_n\frac{9e^{2\gamma n\log 3}}{3^{2n}}.$$

Considering that $\mathfrak{C}_n\setminus\mathcal{A}_{q_n,n}\subseteq\bigcup_{q=1}^{q_n}\mathfrak{C}_n\cap T^{-q}(\mathfrak{C}_n)$, we obtain

$$\mathrm{Leb}^2(\mathfrak{C}_n\setminus\mathcal{A}_{q_n,n}) \le q_n\frac{9e^{2\gamma n\log 3}}{3^{2n}}, \tag{5.4}$$

where γ satisfies the inequality $\gamma\log 3 < \log 2$.

Let $t_n = n^2$, then $t_n = o(w_n)$. Using (5.3) and (5.4), we get that, for all $m_1 > 1$,

$$\begin{aligned}
&\lim_{n\to\infty} w_n \left(\|\mathbf{1}_{\mathfrak{C}_n}\|_\alpha \rho(t_n) + 2\mathrm{Leb}^2(\mathfrak{C}_n \setminus \mathcal{A}_{q_n,n})\right) \\
\leq &\lim_{n\to\infty} \tau \left(\frac{3}{2}\right)^{2n} \left(\left(\frac{2}{3}\right)^{2n} + \frac{4^{n+1}}{3^n}\right) \frac{1}{m_1^{n^2}} + \tau \left(\frac{3}{2}\right)^{2n} q_n \frac{9e^{2\gamma n \log 3}}{3^{2n}} \\
\leq &\lim_{n\to\infty} \left(\tau + \tau \frac{4^{n+1}9^n}{3^n 4^n}\right) \frac{1}{m_1^{n^2}} + 9\tau q_n e^{2n(\gamma \log 3 - \log 2)} \\
= &0.
\end{aligned}$$

This implies that hypothesis (1) through (4) of Theorem 2.5 hold and conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ are satisfied. To conclude the proof it is only necessary to apply the second point of Theorem 2.6 and the result follows for $m_2 \geq m_1$.

For the case where $m_1 \geq m_2$ the proof of the result is analogous. It is only necessary to set $q_n^* = \left\lceil n \frac{\log 3}{\log m_2} \right\rceil$ and again use Theorem 2.5 and Theorem 5.1 to obtain that conditions $Д_{q_n^*}(u_n, w_n)$ and $Д'_{q_n^*}(u_n, w_n)$ hold and that $\theta_2 = 1$. From here, we apply again the second point of Theorem 2.6 and the result follows. $\square$

Theorem 3.2 is an example where there is no clustering associated with either $(X_n^1)_n$ or $(X_n^2)_n$. For that reason, the existence of clustering in the process $(X_n)_n$ was not expected.

A more interesting situation is the setting presented in Theorem 3.3. In this case, the map T is composed of two uni-dimensional dynamics with one of them preserving the structure of the Cantor set, $\mathcal{C}$. We already saw that this compatibility between the dynamics and the maximal set of the observable function leads to clustering. However, due to the product structure in $\mathfrak{C}$, the presence of a second dynamics that does not preserve the structure of $\mathcal{C}$ is enough to guarantee that $T^{-j}(\mathfrak{C}_n) \cap \mathfrak{C}$ is small when compared with $\mathfrak{C}$. This is then translated into a EI equal to 1.

Proof of Theorem 3.3. This proof follows the same structure as the proof of Theorem 3.2. Let $T_1(x) = 3^k x \mod 1$ and $T_2(y) = m_2 y \mod 1$. Assume that (3.5) holds and that m_2 cannot be written as 3^j for any $j \in \mathbb{N}$. Consider the dynamical system $T = (T_1(x), T_2(y))$ and the stochastic process defined by $X_n = \psi \circ T$. The invariant measures associated with these maps are again Leb and Leb^2.

The sequence of thresholds is $u_n = n$ and we set $w_n = \left\lfloor \tau\, (3/2)^{2n} \right\rfloor$. Again, we have that

$$U_n^{T_1} = U_n^{T_2} = \mathcal{C}_n \qquad \text{and} \qquad U_n = \mathfrak{C}_n,$$

which implies that condition 2.3 is satisfied.

Take $q_n = \left\lceil n \frac{\log 3}{\log m_2} \right\rceil$. Under this hypothesis Theorem 5.1 guarantees that $\theta_2 = 1$.

Put $r = \min\{3^k, m_2\}$, then T has decay of correlations for quasi-Hölder observables against $L^1(\mathrm{Leb}^2)$ with rate function $\rho_n = 1/r^n$. The maps $T_1 = 3^k x \mod 1$ and $T_2 = m_2 y \mod 1$ have decay of correlations for BV observables against $L^1(\mathrm{Leb})$ with rate functions $\rho_n^1 = (1/3^k)^n$ and $\rho_n^2 = (1/m_2)^n$.

We will check conditions (1) through (4) of Theorem 2.5 to prove $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$.

To prove hypothesis (1), we will show that there exists a sequence $(t_n)_{n\in\mathbb{N}}$, such that $t_n = o(w_n)$ and

$$\lim_{n\to\infty} w_n \left(\|\mathbf{1}_{\mathfrak{C}_n}\|_\alpha \rho(t_n) + 2\mathrm{Leb}^2(\mathfrak{C}_n \setminus \mathcal{A}_{q_n,n})\right) = 0.$$

Similarly to the proof of Theorem 3.2, we can write that

$$\|\mathbf{1}_{\mathfrak{C}_n}\|_\alpha \leq \mathrm{Leb}^2(\mathfrak{C}_n) + PC(\mathfrak{C}_n),$$

where P denotes the maximum perimeter of the connected components of $\mathfrak{C}_n$ and $C(\mathfrak{C}_n)$ represents the maximum number of connected components of $\mathfrak{C}_n$.

Again, as in the proof of Theorem 3.2, $\mathrm{Leb}^2(\mathfrak{C}_n) = (2/3)^{2n}$, $C(\mathfrak{C}_n) = 4^n$ and $P = 4/3^n$. Therefore,

$$\|\mathbf{1}_{\mathfrak{C}_n}\|_\alpha \leq \left(\frac{2}{3}\right)^{2n} + \frac{4^{n+1}}{3^n}. \tag{5.5}$$

We need an estimate for $\mathrm{Leb}^2(\mathfrak{C}_n \setminus \mathcal{A}_{q_n,n})$. We can write that

$$\mathrm{Leb}^2(\mathfrak{C}_n \cap T^{-q}(\mathfrak{C}_n)) = \mathrm{Leb}(\mathcal{C}_n \cap T_1^{-q}(\mathcal{C}_n))\mathrm{Leb}(\mathcal{C}_n \cap T_2^{-q}(\mathcal{C}_n)) \leq \mathrm{Leb}(\mathcal{C}_n \cap T_1^{-q}(\mathcal{C}_n))\mathrm{Leb}(\mathcal{C}_n)$$

and consequently,

$$\mathrm{Leb}^2\left(\bigcup_{q=1}^{q_n} \mathfrak{C}_n \cap T^{-q}(\mathfrak{C}_n)\right) \leq \sum_{q=1}^{q_n} \mathrm{Leb}(\mathcal{C}_n \cap T_1^{-q}(\mathcal{C}_n))\mathrm{Leb}(\mathcal{C}_n).$$

Using the estimate (5.2) and noting that $3^{-n} \leq m_2^{-q_n}$, we get that

$$\mathrm{Leb}^2\left(\bigcup_{q=1}^{q_n} \mathfrak{C}_n \cap T^{-q}(\mathfrak{C}_n)\right) \leq q_n \left(\frac{3e^{\gamma n \log 3}}{3^n}\right)\left(\frac{2}{3}\right)^n.$$

Considering that $\mathfrak{C}_n \setminus \mathcal{A}_{q_n,n} \subseteq \bigcup_{q=1}^{q_n} \mathfrak{C}_n \cap T^{-q}(\mathfrak{C}_n)$, we obtain that

$$\mathrm{Leb}^2(\mathfrak{C}_n \setminus \mathcal{A}_{q_n,n}) \leq q_n \left(\frac{3e^{\gamma n \log 3}}{3^n}\right)\left(\frac{2}{3}\right)^n, \tag{5.6}$$

where γ satisfies the inequality $\gamma \log 3 < \log 2$. Set $t_n = n^2$, then $t_n = o(w_n)$ and using (5.5) and (5.6), we obtain

$$\begin{aligned}
&\lim_{n\to\infty} w_n \left(\|\mathbf{1}_{\mathfrak{C}_n}\|_\alpha \rho(t_n) + 2\mathrm{Leb}^2(\mathfrak{C}_n \setminus \mathcal{A}_{q_n,n})\right)\\
\leq &\lim_{n\to\infty} \tau \left(\frac{3}{2}\right)^{2n} \left(\left(\frac{2}{3}\right)^{2n} + \frac{4^{n+1}}{3^n}\right)\frac{1}{r^{n^2}} + \tau\left(\frac{3}{2}\right)^{2n} q_n \left(\frac{3e^{\gamma n \log 3}}{3^n}\right)\left(\frac{2}{3}\right)^n\\
\leq &\lim_{n\to\infty} \left(\tau + \tau\frac{4^{n+1}9^n}{3^n 4^n}\right)\frac{1}{r^{n^2}} + 3\tau q_n e^{n(\gamma\log 3 - \log 2)}\\
= &0.
\end{aligned}$$

Since inequality 3.5 holds, there exists a constant $C' > 0$ such that

$$\begin{aligned}\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{C}_n}\|_{BV} \operatorname{Leb}(\mathcal{C}_n) \sum_{j=q_n}^{\infty} \rho_j^1 &\le C' \lim_{n\to\infty} 2^n (2/3)^n \left(\frac{1}{3^k}\right)^{n \log_{m_2}(3)} \\ &\le C' \lim_{n\to\infty} \frac{4^n}{3^{n(1+k\log_{m_2}(3))}} \\ &= 0.\end{aligned}$$

Similarly, there exists a constant $C'' > 0$ such that

$$\begin{aligned}\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{C}_n}\|_{BV} \operatorname{Leb}(\mathcal{C}_n) \sum_{j=q_n}^{\infty} \rho_j^2 &\le C'' \lim_{n\to\infty} 2^n (2/3)^n \left(\frac{1}{m_2}\right)^{n \log_{m_2}(3)} \\ &\le C'' \lim_{n\to\infty} \frac{4^n}{9^n} \\ &= 0.\end{aligned}$$

To finish and again assuming that (3.5) holds, there exists another constant $C''' > 0$ such that

$$\begin{aligned}\lim_{n\to\infty} \|\mathbf{1}_{\mathcal{C}_n}\|_{BV}^2 \sum_{j=q_n}^{\infty} \rho_j^1 \rho_j^2 &\le C''' \lim_{n\to\infty} 4^n \frac{1}{m_2}^{n \log_{m_2}(3)} \left(\frac{1}{3^k}\right)^{n \log_{m_2}(3)} \\ &\le C''' \lim_{n\to\infty} \frac{4^n}{3^{n(1+k\log_{m_2}(3))}} \\ &= 0.\end{aligned}$$

This computation shows that conditions (1) through (4) of Theorem 2.5 hold and conditions $Д_{q_n}(u_n, w_n)$ and $Д'_{q_n}(u_n, w_n)$ are satisfied. To conclude the proof it is only necessary to apply the second point of Theorem 2.6 and the result follows. $\square$

6. Numerical simulation study

In this section we present a small numerical simulation study to illustrate the theoretical results presented in Section 3 for finite time.

We consider the estimator of the EI introduced by Hsing in [18]. Namely, we consider:

$$\hat{\theta}_n(u, q) = \frac{\sum_{i=0}^{n-1} \mathbf{1}_{T^{-i}(\mathcal{A}_q(u))}}{\sum_{i=0}^{n-1} \mathbf{1}_{T^{-i}(U(u))}}, \tag{6.1}$$

where $U(u)$ and $\mathcal{A}_q(u)$ are the sets defined in (2.6), (2.7) and u and q are tuning parameters which determine the estimate's quality. In principle, high values of u should be considered so that the tail behaviour is captured by the quantities in $\hat{\theta}_n(u, q)$, but if u is too high there may not be enough information to estimate the EI accurately. We know that if $Д'_{q*}(u_n, w_n)$ holds for some fixed $q* \in \mathbb{N}$ then $Д'_q(u_n, w_n)$ holds for all $q > q*$. So, the parameter q should not affect too much the estimator's quality. We test several values of u and a few for q and then we analyse the data in order to identify regions of stability of the estimator.

We consider the observable $\psi : [0, 1]^2 \to \mathbb{R}$ defined in (3.2), which is maximised in the Cantor dust set $\mathfrak{C} = \mathcal{C} \times \mathcal{C}$, where $\mathcal{C}$ is the standard ternary Cantor set. The dynamics is ruled by the two-dimensional product map T given in (3.3).

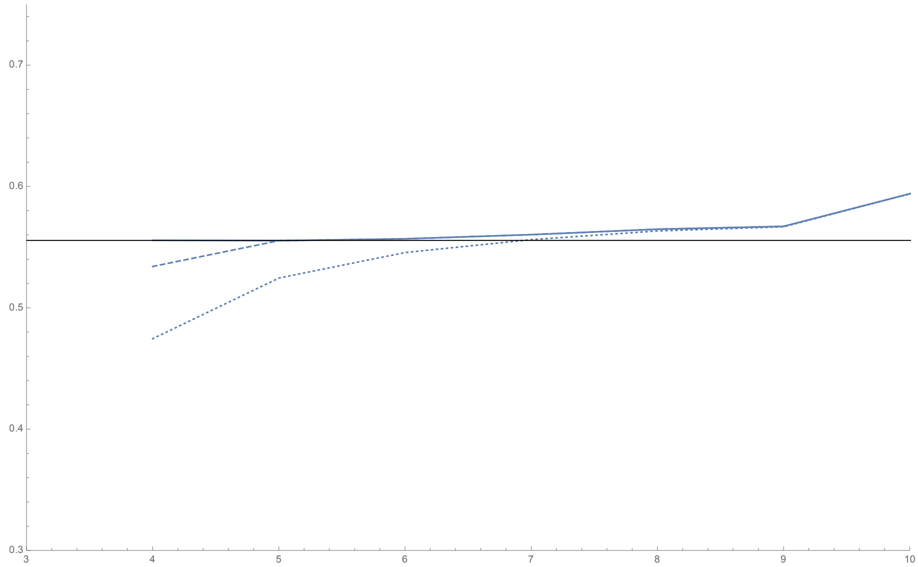

FIGURE 5. On the y-axis, mean values of $\hat{\theta}_n(u, q)$ for each u of the x-axis, with $n = 50.000$ and $\ell = 500$. The full line corresponds to $q = 1$, the dashed line to $q = 5$ and the dotted line to $q = 10$. The black horizontal line represents the exact value of the EI given by Theorem 3.1. In this case, $T(x, y) = (3x \mod 1, 3y \mod 1)$.

The numerical simulations consisted in randomly generating ℓ uniformly distributed points on $[0, 1]$ (recall that Lebesgue measure is invariant for the linear maps considered in the definition of T) and, for each one, compute the first n iterates of the respective orbit and evaluate the observable function ψ, along each orbit. Then, for each the ℓ time series obtained as described above, we compute $\hat{\theta}_n(u, q)$, for several values of u and q, which are adequately chosen for the range of u values.

We observe an excellent agreement between the theoretical value of θ and the observed estimates of $\hat{\theta}_n(u, q)$, in the regions of stability which correspond to the values of u in $[6, 9]$, in the case $m_1 = m_2 = 3$.

In the case $m_1 = m_2 = 5$, there is also an excellent agreement between the theoretical value $\theta = 1$ and the observed estimates of $\hat{\theta}_n(u, q)$, in the regions of stability which correspond to higher values of u, namely, for $u \in [8, 11]$.

In the case $m_1 = 5, m_2 = 3$, where we have a competition between the compatibility observed in the vertical direction versus incompatibility in the horizontal direction, we also observe that, as predicted in Theorem 3.3, the incompatibility prevails and the EI is equal to 1. We note that the agreement improves considerably when we increase the number iterations, n, which allows to have more information on the tails.

REFERENCES

[1] M. Abadi, A. C. M. Freitas, and J. M. Freitas. Dynamical counterexamples regarding the extremal index and the mean of the limiting cluster size distribution. *J. Lond. Math. Soc. (2)*, 102(2):670–694, 2020.

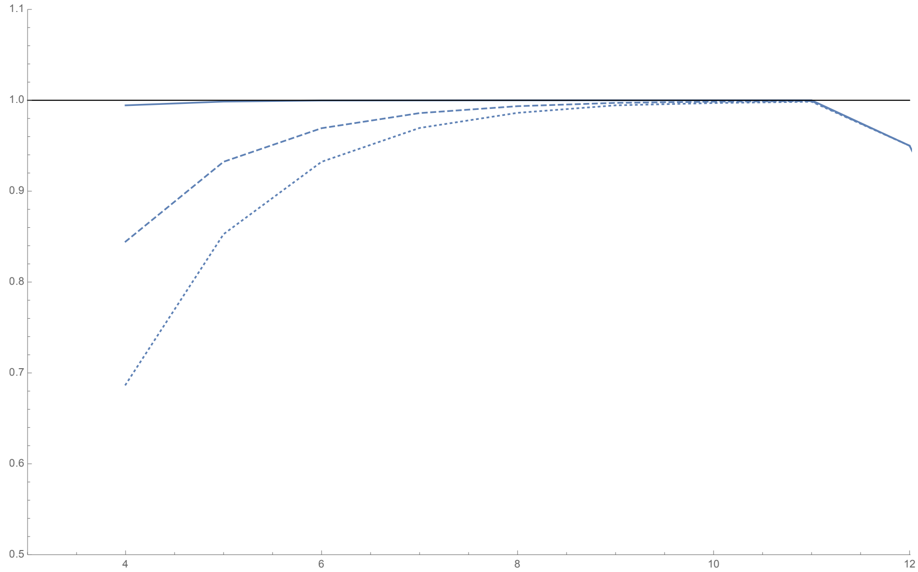

FIGURE 6. On the y-axis, mean values of $\hat{\theta}_n(u,q)$ for each $5 \leq u \leq 20$ of the x-axis, with $n = 50.000$ and $\ell = 500$. The full line corresponds to $q = 1$, the dashed line to $q = 5$ and the dotted line to $q = 10$. The black horizontal line represents the exact value of the EI given by Theorem 3.2. In this case, $T(x,y) = (5x \mod 1, 5y \mod 1)$.

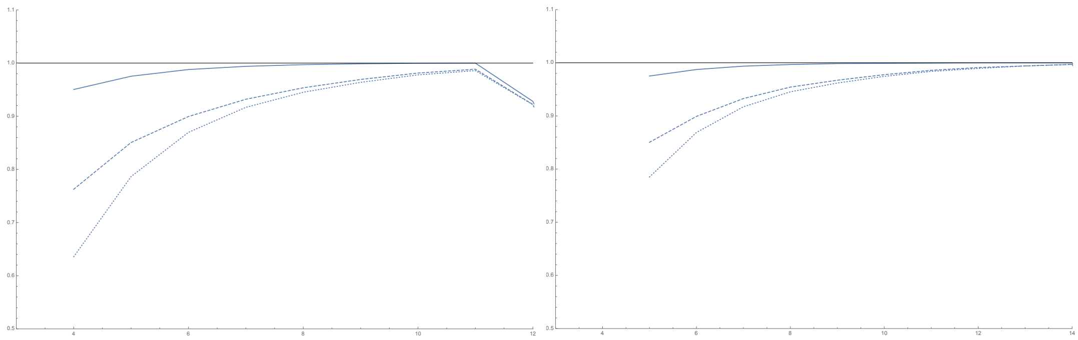

FIGURE 7. On the y-axis, mean values of $\hat{\theta}_n(u,q)$ for each $5 \leq u \leq 20$ of the x-axis. On the left we have $n = 50.000$ and $\ell = 500$, while on the right we have $n = 400.000$ and $\ell = 100$. The full line corresponds to $q = 1$, the dashed line to $q = 5$ and the dotted line to $q = 10$. The black horizontal line represents the exact value of the EI given by Theorem 3.3. In this case, $T(x,y) = (5x \mod 1, 3y \mod 1)$.

[2] H. Aytaç, J. M. Freitas, and S. Vaienti. Laws of rare events for deterministic and random dynamical systems. *Trans. Amer. Math. Soc.*, 367(11):8229–8278, 2015.

[3] D. Azevedo, A. C. M. Freitas, J. M. Freitas, and F. B. Rodrigues. Clustering of extreme events created by multiple correlated maxima. *Phys. D*, 315:33–48, 2016.

[4] D. Azevedo, A. C. M. Freitas, J. M. Freitas, and F. B. Rodrigues. Extreme Value Laws for Dynamical Systems with Countable Extremal Sets. *J. Stat. Phys.*, 167(5):1244–1261, 2017.

[5] A. Boyarsky and P. Góra. *Laws of chaos*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1997. Invariant measures and dynamical systems in one dimension.
[6] T. Caby, D. Faranda, G. Mantica, S. Vaienti, and P. Yiou. Generalized dimensions, large deviations and the distribution of rare events. *Phys. D*, 400:132143, 15, 2019.
[7] M. Carney, M. Holland, and M. Nicol. Extremes and extremal indices for level set observables on hyperbolic systems. Preprint arXiv:1909.04748, 2019.
[8] K. Falconer. *Fractal geometry*. John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2003. Mathematical foundations and applications.
[9] D. Faranda, M. C. Alvarez-Castro, G. Messori, D. Rodrigues, and P. Yiou. The hammam effect or how a warm ocean enhances large scale atmospheric predictability. *Nature Communications*, 10(1):1316, 2019.
[10] D. Faranda, H. Ghoudi, P. Guiraud, and S. Vaienti. Extreme value theory for synchronization of coupled map lattices. *Nonlinearity*, 31(7):3326–3358, 2018.
[11] D. Faranda, G. Messori, M. C. Alvarez-Castro, and P. Yiou. Dynamical properties and extremes of northern hemisphere climate fields over the past 60 years. *Nonlinear Processes in Geophysics*, 24(4):713–725, 2017.
[12] D. Faranda, Y. Sato, B. Saint-Michel, C. Wiertel, V. Padilla, B. Dubrulle, and F. Daviaud. Stochastic chaos in a turbulent swirling flow. *Phys. Rev. Lett.*, 119:014502, Jul 2017.
[13] A. C. M. Freitas, J. M. Freitas, F. B. Rodrigues, and J. V. Soares. Rare events for Cantor target sets. *Comm. Math. Phys.*, 378(1):75–115, 2020.
[14] 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):108–126, 2011.
[15] A. C. M. Freitas, J. M. Freitas, and M. Todd. The extremal index, hitting time statistics and periodicity. *Adv. Math.*, 231(5):2626–2665, 2012.
[16] A. C. M. Freitas, J. M. Freitas, M. Todd, and S. Vaienti. Rare events for the Manneville-Pomeau map. *Stochastic Process. Appl.*, 126(11):3463–3479, 2016.
[17] M. Holland, M. Nicol, and A. Török. Extreme value theory for non-uniformly expanding dynamical systems. *Trans. Amer. Math. Soc.*, 364(2):661–688, 2012.
[18] T. Hsing. Extremal index estimation for a weakly dependent stationary sequence. *Ann. Statist.*, 21(4):2043–2071, 1993.
[19] G. Keller. Rare events, exponential hitting times and extremal indices via spectral perturbation. *Dyn. Syst.*, 27(1):11–27, 2012.
[20] V. Lucarini and T. Bódai. Transitions across melancholia states in a climate model: Reconciling the deterministic and stochastic points of view. *Phys. Rev. Lett.*, 122:158701, Apr 2019.
[21] V. Lucarini, D. Faranda, A. C. M. Freitas, J. M. Freitas, M. Holland, T. Kuna, M. Nicol, and S. Vaienti. *Extremes and Recurrence in Dynamical Systems*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, Hoboken, NJ, 2016.
[22] G. Mantica and L. Perotti. Extreme value laws for fractal intensity functions in dynamical systems: Minkowski analysis. *J. Phys. A*, 49(37):374001, 21, 2016.
[23] G. L. O'Brien. Extreme values for stationary and Markov sequences. *Ann. Probab.*, 15(1):281–291, 1987.
[24] B. Saussol. Absolutely continuous invariant measures for multidimensional expanding maps. *Israel Journal of Mathematics*, 116(1):223–248, Dec 2000.

Ana Cristina Moreira Freitas, Centro de Matemática & Faculdade de Economia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal

Email address: amoreira@fep.up.pt

URL: http://www.fep.up.pt/docentes/amoreira/

Jorge Milhazes Freitas, Centro de Matemática & Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal

Email address: jmfreita@fc.up.pt

URL: http://www.fc.up.pt/pessoas/jmfreita/

Jorge Valentim Soares, Centro de Matemática da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal

Email address: jsoares235@gmail.com