Discrete and Continuous Dynamical Systems

doi:10.3934/dcds.2025030

EXTREME VALUE THEORY WITH SPECTRAL TECHNIQUES: APPLICATION TO A SIMPLE ATTRACTOR

Jason Atnip[1], Nicolai Haydn[2], and Sandro Vaienti*[3]

[1]School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia

[2]Mathematics Department, USC, Los Angeles, 90089-2532, USA

[3]Université de Toulon, Aix Marseille Université, CNRS, CPT, UMR 7332, Marseille, France

Abstract. We give a brief account of application of extreme value theory in dynamical systems by using perturbation techniques associated to the transfer operator. We apply it to the baker's map and we get a precise formula for the extremal index. We also show that the statistics of the number of visits in small sets is compound Poisson distributed.

Contents

2020 *Mathematics Subject Classification.* Primary: 37E30, 37C05, 37A50; Secondary: 60G70.
Key words and phrases. Baker attractor, anisotropic Banach spaces, extreme value theory.

1. **Introduction.** Extreme value theory (EVT) has been widely studied in recent years in application to dynamical systems both deterministic and random. A review of the recent results with an exhaustive bibliography is given in our collective work [26]. As we will see, there is a close connection between EVT and the statistics of recurrence, and both could be worked out simultaneously by using perturbations theories of the transfer operator. This powerful approach is limited to systems with quasi-compact transfer operators and exponential decay of correlations; nevertheless it can be applied to situations where more standard techniques meet obstructions and difficulties, in particular to:
- non-stationary and random dynamical systems,
- observable with non-trivial extremal sets,
- higher-dimensional systems.

Another big advantage of this technique is the possibility of defining in a precise and universal way the extremal index (EI). We defer to our recent paper [7] for a critical discussion of this issue with several explicit computations of the EI in new situations. The germ of the perturbative technique of the transfer operator applied to EVT is in the fundamental paper [24] by G. Keller and C. Liverani; the explicit connection with recurrence and extreme value theory was done by G. Keller in [23], which also contains a list of suggestions for further investigations. We successively applied this method to i.i.d. random transformations in [5, 7], to randomly quenched dynamical systems in [2], to coupled maps on finite lattices in [13], and to open systems with targets and holes in [18].

The object of this note is to illustrate this technique by presenting a new application to a bi-dimensional invertible system. We will see that the perturbative technique could be applied in this case as well provided one could find good functional spaces where the transfer operator exhibits quasi-compactness.

We will find a few limitations to a complete application of the theory and to its generalization to a wider class of maps in higher dimensions, see Remarks 5.3 and 5.4.

When the first version of this paper circulated, the spectral technique discussed above did not allow us to get another property related to limiting return and hitting times distribution in small sets (sometimes also called *target* or *holes*), namely the statistics of the number of visits, which usually takes the form of a compound Poisson distribution. This has been recently achieved in [3], and it could be easily applied to the system under investigation in this paper. We will briefly quote this technique in Section 7. As for the EVT, such a technique suffers the limitation imposed by the shape of the target sets, and for the choice of the parameters, see remark 5.4. The former will be particularly important for us when we decide to use the rectangular target set, see Section 6.1. These cases could be worked out with another technique developed by two of us, see [21], which allows for the recovery of compound Poisson distributions for invertible maps in a higher dimension and for arbitrarily small sets. By using this approach, we will also be able to construct an example for the *fat* baker map with a compound Poisson distribution which is neither the standard Poisson nor the Pòlya-Aeppli, which are the most common compound distributions.

We will finally discuss the extension to the compound Poisson point process on the real line.

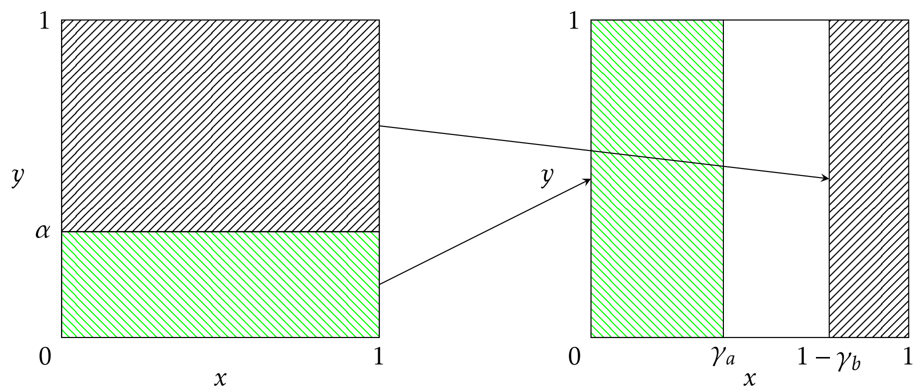

Figure 1. Action of the baker's map on the unit square. The lower part of the square is mapped to the left part and the upper part is mapped to the right part.

2. A simple example: The generalized baker's map.

2.1. The map.

We now treat an example for which there are not apparently established results for the extreme value distributions. This example, the generalized baker's map, from now on simply abbreviated as baker's map, is a prototype for uniformly hyperbolic transformations in more than one dimension, two in our case, and in order to study it with the transfer operator, we will introduce suitable anisotropic Banach spaces. Our original goal was to directly investigate larger classes of uniformly hyperbolic maps, including Anosov ones, but, as we said above, the generalizations do not seem straightforward; we will explain the reason later on. With the usual probabilistic approaches, extreme value distributions have been obtained for the linear automorphisms of the torus in [8].

We will refer to the baker's transformation studied in Section 2.1 in [10], but we will write it in a particular case in order to make the exposition more accessible. The baker's transformation $T(x_n, y_n)$ is defined on the unit square $X = [0,1]^2 \subset \mathbb{R}^2$ into itself by

$$x_{n+1} = \begin{cases} \gamma_a x_n & \text{if } y_n < \alpha \\ (1-\gamma_b) + \gamma_b x_n & \text{if } y_n > \alpha \end{cases} \tag{1}$$

$$y_{n+1} = \begin{cases} \frac{1}{\alpha} y_n & \text{if } y_n < \alpha \\ \frac{1}{v}(y_n - \alpha) & \text{if } y_n > \alpha, \end{cases} \tag{2}$$

with $v = 1-\alpha$, $\gamma_a + \gamma_b \leq 1$; see Fig. 1. To simplify some of the following formulae, we will take $\alpha = v = 0.5$ and $\gamma_a = \gamma_b < 0.5$. This last value must be strictly less than 1/2 since Lemma 5.2 requires the stable dimension d_s strictly less than one, which corresponds to a fractal invariant set (*thin baker's map*). This condition will be relaxed in example 7.5 (*fat baker's map*), but using an approach different of the spectral one leading to Lemma 5.2.

The map T is discontinuous at the horizontal line $\Gamma : \{y = \alpha\}$. The singularity curves for $T^l, l > 1$ are given by $T^{-l}\Gamma$, and they are constructed in this way: Take the preimages $T_Y^{-l}(\alpha)$ of $y=\alpha$ on the y-axis according to the map

$$T_Y(y)=\begin{cases}\frac{1}{\alpha}y, y<\alpha\\ \frac{1}{v}y-\frac{\alpha}{v}, y\ge\alpha.\end{cases}\tag{3}$$

Then, $T^{-l}\Gamma=\{y=T_Y^{-l}(\alpha)\}$. Any other horizontal line will be a stable manifold of T. The invariant non-wandering set Λ will be at the end an attractor foliated by vertical lines which are all unstable manifolds. We denote by $\mathcal{W}^s(\mathcal{W}^u)$ the set of full horizontal (vertical) stable (unstable) manifolds of length 1 just constructed. We point out that a stable horizontal manifold W_s will originate two disjoint full stable manifolds when iterated backward by T^{-1}, not for the presence of singularity, but because the map T^{-1} will only be defined on the two images of $T(X)$ as illustrated in Fig. 1.

2.2. **The functional space.** In order to obtain useful spectral information from the transfer operator $\mathcal{L}$, its action is restricted to a Banach space $\mathcal{B}$. We now give the construction of the norms on $\mathcal{B}$ and an associated "weak" space $\mathcal{B}_w$ in the case of the baker's map, following partly the exposition in [10]. In this case, the norms will be constructed directly on the horizontal stable manifolds instead of admissible leaves, which are smooth curves in approximately the stable direction, see [11]. As we anticipated above, we follow [10], but we slightly change the definition of the stable norms by adapting ourselves to that originally introduced in [11]. Let us explain why. First of all, we will consider the collection Σ of all the intervals W of length less or equal to 1 that are contained in the same stable manifold $W_s\in\mathcal{W}^s$. We will take such a value equal to γ_a for reasons which will be clear in the next considerations. Instead, in [11], Σ was the set of full horizontal line segments of length 1 in X. The reason for our choice is that we will introduce small target sets B_n and the preimages of such sets will cut the W_s. The smaller pieces generated in this way will enter the three norms given below, and therefore it will be useful to count such pieces in Σ.

We now to consider the set C^ρ of continuous complex-valued functions over X with Hölder exponent $0\le\rho\le 1$. When we set C^1, we mean C^ρ with $\rho=1$, which is simply Lipschitz. Given a stable leaf W and a Hölder function φ, we define the norm along W as

$$|\varphi|_{C^\rho(W)}=|\varphi|_{C^0(W)}+H^\rho(\varphi),\ H^\rho(\varphi)=\sup_{\substack{x,y\in W\\ x\neq y}}\frac{|\varphi(x)-\varphi(y)|}{|x-y|^\rho}.$$

Another norm will be considered later on, namely

$$|\varphi|_{W,\rho}:=|W|^\rho\cdot|\varphi|_{C^\rho(W)},\tag{4}$$

where $|W|$ denotes the length of W.

We now closely follow Section 2.2 in [10] and put

$$|\varphi|_{C^\rho(\mathcal{W}^s)}=\sup_{W\in\mathcal{W}^s}|\varphi|_{C^\rho(W)}.\tag{5}$$

We call $C^1(\mathcal{W}^s)$ the set of functions that are Lipschitz along stable manifolds, i.e., for which the quantity (5) is finite. For $\rho<1$, we set $C^\rho(\mathcal{W}^s)$, as the completion of $C^1(\mathcal{W}^s)$ in the $|\cdot|_{C^\rho(\mathcal{W}^s)}$ norm. Analogously, $C^\rho(W^s)$ denotes the completion of $C^1(W^s)$ in the $|\cdot|_{C^\rho(W^s)}$ norm. One can show that $|\varphi|_{C^\rho(\mathcal{W}^s)}$ is a Banach space for $0\le\rho\le 1$; the Banach space $|\varphi|_{C^\rho(\mathcal{W}^u)}$ is defined similarly. By the very structure of the map, it follows that whenever $\varphi\in C^\rho(\mathcal{W}^s)$, then $\varphi\circ T\in C^\rho(\mathcal{W}^s)$. This allows one to define the transfer operator $\mathcal{L}$ associated with T on the dual space $(C^\rho(\mathcal{W}^s))^*$ as[1]

$$(\mathcal{L}h)(\varphi) = h(\varphi \circ T),\ \forall \varphi \in C^\rho(\mathcal{W}^s),\ h \in (C^\rho(\mathcal{W}^s)^*.$$

If we denote by m_L the Lebesgue measure over X and we take $h \in C^1(\mathcal{W}^u)$, then we identify h with the measure hdm_L so that $h \in (C^\rho(\mathcal{W}^s))^*$ and $\mathcal{L}h$ is now identified with the measure having density

$$\mathcal{L}h(x) = \left(\frac{h}{|\det DT|}\right) \circ T^{-1}(x) = \frac{h \circ T^{-1}(x)}{\alpha^{-1}\gamma_a}, \tag{6}$$

where the last equality on the r.h.s. uses the particular choices for the parameters defining the map T. When $h \in C^1(\mathcal{W}^u)$, we therefore set

$$h(\varphi) = \int_X h\varphi\, dm_L, \quad \text{for } \varphi \in C^1(\mathcal{W}^s). \tag{7}$$

We are now ready to construct the Banach spaces.

For $h \in C^1(\mathcal{W}^u)$, we define the *weak norm* of h by

$$|h|_w = \sup_{W\in\Sigma} \sup_{\substack{\varphi\in C^1(W) \\ |\varphi|_{C^1(W)}\leq 1}} \left| \int_W h\varphi\, dm \right|$$

where dm is the unnormalized Lebesgue measure along W.

We now take[2] $(\kappa, \beta) \in (0,1)$ with $0 < \beta \leq \min(\kappa, 1-\kappa)$.

The strong stable norm is defined as

$$\|h\|_s = \sup_{W\in\Sigma} \sup_{\substack{\varphi\in C^\kappa(W) \\ |\varphi|_{W,\kappa}\leq 1}} \left| \int_W h\varphi\, dm \right|. \tag{8}$$

We then need to define the strong unstable norm, which allows us to compare expectations along different stable manifolds. If W_1 is a subset of the stable manifold W_s we could parameterize it as (s_{W_1}, t) where s_{W_1} is the common ordinate of the points in W_1 and $t \in [a_1, b_1] \subset [0,1]$. If W_2 is a subset of another stable manifold, parametrized as (s_{W_2}, t) with $t \in [a_2, b_2]$, we pose

$$d(W_1, W_2) = |s_{W_1} - s_{W_2}| + |[a_1, b_1]\Delta[a_2, b_2]|,$$

where Δ means the symmetric difference, and for test functions $\varphi_i \in C^1(W_i), i = 1,2:$

$$d_\kappa(\varphi_1, \varphi_2) = \sup_{t\in[a_1,b_1]\cap[a_2,b_2]} |\varphi_1(s_{W_1}, t) - \varphi_2(s_{W_2}, t)|_{C^\kappa(W_i)}.$$

The strong unstable norm of h is defined as

$$\|h\|_u = \sup_{\epsilon\leq 1} \sup_{\substack{W_1,W_2\in\mathcal{W}_s \\ d(W_1,W_2)\leq\epsilon}} \sup_{\substack{\varphi_i\in C^1(W_i) \\ |\varphi_i|_{C^1(W)}\leq 1 \\ d_\kappa(\varphi_1,\varphi_2)\leq\epsilon}} \frac{1}{\epsilon^\beta} \left| \int_{W_1} h\varphi_1\, dm - \int_{W_2} h\varphi_2 dm \right|, \tag{9}$$

[1]Notice that although the map T is discontinuous, the fact that $\varphi \in C^\rho(\mathcal{W}^s)$ implies $\varphi \circ T \in C^\rho(\mathcal{W}^s)$, allows us to define the transfer operator on the space $(C^\rho(\mathcal{W}^s))^*$ without the need for two scales of space as in [11].

[2]The bound $\beta \leq \min(\kappa, 1-\kappa)$ is needed in the proof of Lemma 3.1 in [10]. Notice that such a lemma only requires $\beta \leq 1-\kappa$. The additional constraint $\beta \leq \kappa$ comes from the fact that in the proof of Lemma 3.1 in [10], in particular in the estimate of the strong unstable norm, there are not *unmatched* pieces since all the stable leaves have full length; see also footnote 3.

Finally we can define the *strong norm* of h by

$$\|h\| = \|h\|_s + b\|h\|_u,$$

where b is a small constant to be fixed later on.

We set $\mathcal{B}$ to be the completion of $C^1(\mathcal{W}^u)$ with respect to the norm $\|\cdot\|$, and, similarly, we define $\mathcal{B}_w$ to be the completion of $C^1(\mathcal{W}^u)$ with respect to the norm $|\cdot|_w$.

We now list a few important results whose proof can be found in [10] and which we will use frequently in the next sections.

- (Lemma 2.4, [10]) For any $\beta' \in (\beta, 1)$, we have the following sequence of continuous embeddings:

  $$C^1(X) \hookrightarrow C^{\beta'}(\mathcal{W}^u) \hookrightarrow \mathcal{B} \hookrightarrow \mathcal{B}_w \hookrightarrow (C^1(\mathcal{W}^s))^*.$$

  Moreover, the embedding $\mathcal{B} \hookrightarrow \mathcal{B}_w$ is relatively compact.
- For $h \in \mathcal{B}$ and $\varphi \in C^1(\mathcal{W}^s)$, we have

  $$|h(\varphi)| \le |h|_w |\varphi|_{C^1(\mathcal{W}^s)}. \tag{10}$$

  Moreover,

  $$|h|_w \le \|h\|_s. \tag{11}$$
- (Lemma 4.1, [10]). The transfer operator $\mathcal{L}$ is a bounded, linear operator on both $\mathcal{B}$ and $\mathcal{B}_w$.
- (Lemma 3.1, [10]). If $g \in C^1(X)$ and $h \in \mathcal{B}$, then[3]

  $$\|gh\| \le (5b+1)|g|_{C^1(X)}\|h\|. \tag{12}$$
- (Theorem 2.5, [10]). $\mathcal{L}$ is quasi-compact as an operator on $\mathcal{B}$. Its spectral radius is 1 and its essential spectral radius is bounded by $\max\{\lambda_a^\kappa, \alpha^\beta\} < 1$.

  Then:

  (a) $\mathcal{L}$ has 1 as a simple eigenvalue and all other eigenvalues have modulus less than 1.

  (b) There is a unique solution $\mu \in \mathcal{B}$ of $\mathcal{L}\mu = \mu$ with $\mu(1) = 1$, and such a solution is the Sinai-Bowen-Ruelle, SRB-measure. Its conditional measures on unstable leaves are equal to arclength.

  (c) There exist $\mathfrak{a} < 1$ and $\mathfrak{C} > 0$ such that for any $h \in \mathcal{B}$ with $h(1) = 1$, we have

  $$\|\mathcal{L}^n h - \mu\| \le \mathfrak{C}\mathfrak{a}^n \|h\|, \ \forall n \ge 0.$$

3. **The spectral approach for EVT.**

3.1. **Formulation of the problem.** We now take a ball $B(z,r)$ of center $z \in X$ and radius r and denote with $B(z,r)^c$ its complement, where $d(\cdot,\cdot)$ is the Euclidean metric.

Let us consider for $x \in X$ the observable[4]

$$\Xi(x) = -\log d(x,z) \tag{13}$$

and the function

$$M_n(x) := \max\{\Xi(x), \cdots, \Xi(T^{n-1}x)\}. \tag{14}$$

[3] In [10], the constant on the r.h.s.of (12) is simply 3. We should modify it since the presence of unmatched pieces adds two more contributions of the factor $|g|_{C1(X)}\|h\|_s$ in the computation of the strong unstable norm. Finally, the factor b comes from the very definition of the Banach norm $\|\cdot\|$.

[4] See section 8 for a discussion about the choice of the observable.

For $u \in \mathbb{R}_+$, we are interested in the distribution of $M_n \leq u$, where M_n is now seen as a random variable on the probability space (X, μ). Notice that the event $\{M_n \leq u\}$ is equivalent to the set $\{\Xi \leq u, \ldots, \Xi \circ T^{n-1} \leq u\}$, which in turn coincides with the set

$$E_n := B(z, e^{-u})^c \cap T^{-1} B(z, e^{-u})^c \cap \cdots \cap T^{-(n-1)} B(z, e^{-u})^c.$$

We are therefore following points which will enter the ball $B(z, e^{-u})$ for the first time after at least n steps (see e.g. Eq. (67) in Section 8), and $u \mapsto \mu(E_n)$ is the distribution function of the maximum of the observable $\Xi \circ T^j, j = 0, \ldots, n-1$. It is well known from elementary probability that the distribution of the maximum of a sequence of i.i.d. random variables is degenerate. One way to overcome this is to make the *boundary level* u depend upon the time n in such a way the sequence u_n grows to infinity and gives, hopefully, a non-degenerate limit for $\mu(M_n \leq u_n)$.

From now on we set $B_n = B(z, e^{-u_n})$ and B_n^c the complement of B_n; the dependence upon the "center" z will be discussed in Remark 5.3.
We easily have

$$\mu(M_n \leq u_n) = \int \mathbf{1}_{B_n^c}(x) \mathbf{1}_{B_n^c}(Tx) \cdots \mathbf{1}_{B_n^c}(T^{n-1}x)\, d\mu. \tag{15}$$

By introducing the perturbed operator, for $h \in \mathcal{B}$,

$$\mathcal{L}_n h := \mathcal{L}(\mathbf{1}_{B_n^c} h), \tag{16}$$

and we can write (15) as

$$\mu(M_n \leq u_n) = \mathcal{L}_n^n \mu(1). \tag{17}$$

Notice that

$$\mathcal{L}_n^k h = \mathcal{L}(h \mathbf{1}_{B_n^c} \mathbf{1}_{B_n^c} \circ T \ldots \mathbf{1}_{B_n^c} \circ T^{k-1}).$$

3.2. **Target sets and the space $\mathcal{B}$.** We explicitly used the above two facts, which require justification.

3.2.1. $\mathbf{1}_{B_n^c}$ *is in the Banach space* $\mathcal{B}$. Of course, the same proof should hold for functions of the form $\mathbf{1}_{B_n^c} \mathbf{1}_{B_n^c} \circ T \ldots \mathbf{1}_{B_n^c} \circ T^{k-1}$. The geometric shape of the sets $B_n^c \cap T^{-1} B_n^c \cap \cdots \cap T^{-(k-1)} B_n^c$ plays an important role in the proof. Those sets are equivalently given by $\left(B_n \cup T^{-1} B_n \cup \cdots \cup T^{-(k-1)} B_n\right)^c$ and we call $B^{(k)}$ one of them. Suppose we could find a sequence $\{h_l\}_{l \in \mathbb{N}}$ in $C^1(X)$ which is Cauchy in $\mathcal{B}$ and such that for any $\varphi \in C^1(\mathcal{W}^s)$ we have

$$\int h_l \varphi \, dm \to \int \mathbf{1}_{B^{(k)}} \varphi \, dm, \; l \to \infty. \tag{18}$$

Then $\mathbf{1}_{B^{(k)}}$ is in $\mathcal{B}$ since the latter is continuously embedded in the dual space $(C^1(\mathcal{W}^s))^*$, see Lemma 2.4 in [10] quoted in section 2. We now construct such a sequence. Take a bump function ϕ with support in the unit ball of $\mathbb{R}^2$, normalized to 1 and put $\phi_\delta(x) := \frac{1}{\delta^2} \phi(\frac{x}{\delta})$, where δ designs the δ-neighborhood $B_\delta^{(k)} := \{x \in \mathbb{R}^2; \mathrm{dist}(x, B^{(k)}) \leq \delta\}$ of $B^{(k)}$. Then define the convolution product

$$h_\delta := \mathbf{1}_{B_\delta^{(k)}} * \phi_\delta. \tag{19}$$

Then h_δ is equal to 1 on $B^{(k)}$ and equal to zero outside the 2δ-neighborhood $B_{2\delta}^{(k)}$. Moreover it is straightforward to get (18). It remains to prove that $\{h_\delta\}_{\delta > 0}$

is Cauchy, for $\delta \to 0$. Call $U_\delta = B^{(k)}_{2\delta} \setminus B^{(k)}$. To control the strong stable norm, we observe that, if $\delta_2 < \delta_1$

$$\left|\int_W (h_{\delta_1} - h_{\delta_2})\varphi dm\right| = \left|\int_{W\cap U_{\delta_1}} (h_{\delta_1} - h_{\delta_2})\varphi dm\right| \le 2|W|^{-\kappa}|W\cap U_{\delta_1}| \le 2|W\cap U_{\delta_1}|^{1-\kappa} \tag{20}$$

There are now two cases:

(i) suppose first that $|W| \le \delta_1$; then (20)$\le \delta_1^{1-\kappa}$.

(ii) suppose now that $|W| > \delta_1$. As we will write in footnote 5, each W could meet at most $k-1$ sets of the form $T^{-j}B_n, j = 1, \dots, k$. These sets are ellipses with the major axis along the stable manifolds. Therefore each W could meet at most $(k-1)$ δ_1–neighborhoods of the preimages $T^{-j}B_n, j = 1, \dots, k$. It is a simple exercise to show that the maximum intersection of W with one of the previous δ_1–neighborhoods is bounded by a constant $\tilde{C}$ depending only on the size of X times $(\delta_1)^{1/2}$. Then (20)$\le 2(k-1)[\tilde{C}\delta_1^{1/2}]^{1-\kappa}$. In conclusion

$$\|h_{\delta_1} - h_{\delta_2}\|_s \le 2(k-1)[\tilde{C}\delta_1^{1/2}]^{1-\kappa}.$$

We now compute the strong unstable norm. We proceed in two different manners. First of all we could simply bound the difference

$$\frac{1}{\epsilon^\beta}\left|\int_{W_1} (h_{\delta_1} - h_{\delta_2})\varphi_1 dm - \int_{W_2} (h_{\delta_1} - h_{\delta_2})\varphi_2 dm\right| \tag{21}$$

by $\frac{1}{\epsilon^\beta}4(k-1)\tilde{C}\delta_1^{1/2}$, since the term $h_{\delta_1} - h_{\delta_2}$ is different from zero only on the intersections of the manifolds W_1, W_2 with a δ_1-neighborhood. We now pass to a finer estimate of (21) and we will use the same trick to control the strong unstable norm in the Lasota-Yorke inequality, see section 4.3. We will see that giving two stable manifolds W_1, W_2 at a distance at most ϵ, there will be two *matched* subsets of those manifolds whose points have the same y-ordinate, and the x-components belong to the same interval. The complement of the matched piece on each manifold has length less or equal to ϵ (*unmatched pieces*). The contribution given by those two unmatched pieces is $2(k-1)\epsilon^{1-\beta}$. We now parametrize the two matched pieces, where all the points of W_1 (resp. W_2), have the same ordinate s_1 (resp. s_2) and the abscissa t varies in the interval $I_{1,2}$. Then we can write the difference of the two integrals in (21) as

$$\left|\int_{I_{1,2}} (h_{\delta_1} - h_{\delta_2})(s_1,t)\varphi_1(s_1,t)dt - \int_{I_{1,2}} (h_{\delta_1} - h_{\delta_2})(s_2,t)\varphi_2(s_2,t)dt\right|. \tag{22}$$

Notice that from now on it only matters the intersection of W_1, W_2 with U_{δ_1}, since outside it the quantity (22) is zero. We begin to control the piece $D_1 := \int_{I_{1,2}} h_{\delta_1}(s_1,t)\varphi_1(s_1,t)dt - \int_{I_{1,2}} h_{\delta_1}(s_2,t)\varphi_2(s_2,t)dt$, the other one involving h_{δ_2}, call it D_2, behaves in the same way. We split it as

$$\begin{aligned} D_1 = &\int_{I_{1,2}} h_{\delta_1}(s_1,t)\varphi_1(s_1,t)dt - \int_{I_{1,2}} h_{\delta_1}(s_1,t)\tilde{\varphi}_1(s_1,t)dt \\ &+ \int_{I_{1,2}} h_{\delta_1}(s_1,t)\tilde{\varphi}_1(s_1,t)dt - \int_{I_{1,2}} h_{\delta_1}(s_2,t)\varphi_2(s_2,t)dt, \end{aligned}$$

where we put

$$\tilde{\varphi}_1(s_1,t) = \varphi_2(s_2,t),\ t \in I_{1,2}.$$

The absolute value of the first difference in D_1 is simply bounded by ϵ (remember $d_\kappa(\varphi_1,\varphi_2)\le\epsilon$), times $2(k-1)\tilde{C}\delta_1^{1/2}$. The absolute value of the second piece in D_1 is bounded by $2(k-1)\tilde{C}\delta_1^{1/2}$ times $H(h_{\delta_1})\epsilon$, where $H(h_{\delta_1})\le\frac{C_\phi}{\delta_1^3}$ is the Lipschitz constant of h_{δ_1} and C_ϕ depends only on ϕ. By dividing for ϵ^β we have that $|D_1|+|D_2|$ is bounded by

$$2(k-1)\frac{\hat{C}}{\delta_1^{5/2}}\epsilon^{1-\beta}, \tag{23}$$

where $\hat{C}=\max(\tilde{C},C_\phi)$. In conclusion we have (the term coming from the unmatched pieces being incorporated in (23))

$$\|h_{\delta_1}-h_{\delta_1}\|_u\le\min\{4(k-1)\frac{\hat{C}}{\delta_1^{5/2}}\epsilon^{1-\beta},\frac{1}{\epsilon^\beta}4(k-1)\tilde{C}\delta_1^{1/2}\}.$$

By interpolating we finally get

$$\|h_{\delta_1}-h_{\delta_1}\|_u\le 4(k-1)\hat{C}\epsilon^{1-(q+\beta)}\delta_1^{3q-5/2},$$

where $q\in(0,1)$ must be chosen such that $\beta+q<1$ and $q>5/6$.

3.2.2. $\mathbf{1}_{B^{(k)}}h\in\mathcal{B}$. Take again a set like $B^{(k)}$ such that $\mathbf{1}_{B^{(k)}}\in\mathcal{B}$; what we need is that $\mathbf{1}_{B^{(k)}}h\in\mathcal{B}$, where $h\in\mathcal{B}$. First of all we have to define the object $\mathbf{1}_{B^{(k)}}h\in\mathcal{B}$. Take a sequence $\{h_l\}_{l\ge1}\in C^1(\mathcal{W}^u)$ converging to h in the $\mathcal{B}$–norm. Whenever $\mathbf{1}_{B^{(k)}}h_l\in\mathcal{B}$, we set

$$\mathbf{1}_{B^{(k)}}h=\lim_{l\to\infty}\mathbf{1}_{B^{(k)}}h_l, \tag{24}$$

provided the limit exists. So, first of all we have to show that for any l, $\mathbf{1}_{B^{(k)}}h_l\in\mathcal{B}$. This is proved exactly in the same manner as in the previous item, where $h_l=1$.

We are left by showing that $\mathbf{1}_{B^{(k)}}h_l$ is Cauchy. To get it we begin to prove a preliminary result, namely in the computation of the strong stable and unstable norm of $\mathbf{1}_{B^{(k)}}f$, where $\mathbf{1}_{B^{(k)}}\in\mathcal{B}$ and $f\in C^1(\mathcal{W}^u)$, such norms can be computed by using directly the (non smooth) product $\mathbf{1}_{B^{(k)}}f$. By (24), if we call g_l a sequence converging to $\mathbf{1}_{B^{(k)}}$, we put

$$\mathbf{1}_{B^{(k)}}f=\lim_{l\to\infty}g_lf,$$

but now we are sure the limit exists since the sequence g_lf is Cauchy by (12):

$$\|(g_l-g_k)f\|\le(5b+1)\,\|g_l-g_k\|\,|f|_{C^1(X)}.$$

Therefore we have that

$$\|g_lf\|\to\|\mathbf{1}_{B^{(k)}}f\|,$$

and the norm on the l.h.s. is the norm ”before” completion. So we have

$$\|\mathbf{1}_{B^{(k)}}f\|=\lim_{l\to\infty}(\|g_lf\|_s+b\|g_lf\|_u).$$

The result follows by replacing the strong stable and unstable norms on the right-hand side respectively with the expressions (8) and (9) and by passing the limit inside the integrals by dominated convergence. The same argument shows also that for $h\in C^1(\mathcal{W}^u)$,

$$\mathbf{1}_{B^{(k)}}h(\varphi)=\mathbf{1}_{B^{(k)}}(h\varphi)=\int_X\mathbf{1}_{B^{(k)}}h_l\varphi dm. \tag{25}$$

We now return to prove that $\mathbf{1}_{B^{(k)}}h_l$ is Cauchy by computing the norm of the generic element $\mathbf{1}_{B^{(k)}}h$, $h \in C^1(\mathcal{W}^u)$, directly along the stable manifolds and showing that it is bounded by a constant depending only on $B^{(k)}$ times $\|h\|$.[5] If we take a stable manifold W of length at most γ_a, the intersection $W \cap B^{(k)}$ is given by a finite number $\#(W,k)$ of smaller stable intervals $W_i, 1 \le i \le \#(W,k)$. The latter are generated by removing from W the intersections with the preimages of B_n up to order k (see the beginning of section 3.2.1). Therefores, $\#(W,k) \le k, \forall W$, as explained in footnote 7. By using the arguments in **A2** or in **A3** in the next sections, it is straightforward to check that $\|\mathbf{1}_{B^{(k)}}h\|_s \le k\|h\|_s$ and $|\mathbf{1}_{B^{(k)}}h|_w \le k|h|_w$. It remains to compute the strong unstable norm, and this reduces to bounding the difference for a smooth h: $\frac{1}{\epsilon^\beta}\left|\int_{W_1\cap B^{(k)}} h\varphi_1 dm - \int_{W_2\cap B^{(k)}} h\varphi_2 dm\right|$, where $W_1 W_2, \varphi_1, \varphi_2$ verify the constraints given in (9). We recall that the preimages of B_n of order $l \le k$ are ellipsis whose axis along the vertical unstable direction has length at most α^l. We now split the computation in two parts. Suppose first that $\epsilon \ge 0.5\alpha^k$. Then, a rough estimate gives

$$\frac{1}{\epsilon^\beta}\left|\int_{W_1\cap B^{(k)}} h\varphi_1 dm - \int_{W_2\cap B^{(k)}} h\varphi_2 dm\right| \le 4\frac{1}{\alpha^{k\beta}}\|h\|_s. \tag{26}$$

Take now $\epsilon < 0.5\alpha^k$. By following the strategy used in dealing with the strong unstable norm in Sections 3.2.1 and 4.3, we split the difference above over unmatched and matched pieces. Suppose now that $\#(W_1,k) > \#(W_2,k)$. Then, there could be at most $\#(W_1,k)$ matched pieces given a final contribution bounded by $k\|h\|_u$. We now count the unmatched pieces. First, there are the two intervals of length $\le \epsilon$ at the extremities of W_1, W_2. Notice that there could be at most $\#(W_2,k) - 1$ preimages of B_n which cut both W_1 and W_2. Therefore, there will be at most $2[\#(W_2,k) - 1]$ unmatched pieces generated at the intersection with the boundaries of such preimages and the length of each of those unmatched pieces is bounded by a constant $C_1(B_n,k)$ depending solely on the radius of the ball B_n and on k times $\sqrt{\epsilon}$, as we argued in Section 4.3. But now there will also be $[\#(W_1,k)] - [\#(W_2,k)]$ unmatched pieces given by preimages of B_n which cut W_1 but not W_2. The length of those pieces will be again bounded by a constant $C_2(B_n,k)$ times $\sqrt{\epsilon}$. Hence, we get for $\epsilon < 0.5\alpha^k$:[6]

$$\left|\int_{W_1\cap B^{(k)}} h\varphi_1 dm - \int_{W_2\cap B^{(k)}} h\varphi_2 dm\right| \le \epsilon^\beta k\|h\|_u + \left[2\epsilon^\kappa + [2C_1(B_n,k) + C_2(B_n,k)]k\epsilon^{\kappa/2}\right]\|h\|_s. \tag{27}$$

In conclusion, for $2\beta < \kappa$ we get

$$\|\mathbf{1}_{B^{(k)}}h\|_u \le \max\{4\frac{1}{\alpha^{k\beta}}, 2 + [1 + 2C_1(B_n,k) + C_2(B_n,k)]k\}\|h\|_u,$$

which immediately implies the Cauchly property for the sequence $\mathbf{1}_{B^{(k)}}h_l$.

We finally consider (25) when h is a Borel measure $\tilde{\mu} \in \mathcal{B}$. Suppose also that the $\tilde{\mu}$–measure of the boundary of $B^{(k)}$ is zero. This happens for instance if $\tilde{\mu}$ is the SRB-measure, which is our case. Then, if $h_l \to \tilde{\mu}$, for a test function φ, we have

$$\mathbf{1}_{B^{(k)}}\tilde{\mu}(\varphi) = \lim_{l\to\infty} \mathbf{1}_{B^{(k)}}h_l(\varphi) = \lim_{l\to\infty} \int \mathbf{1}_{B^{(k)}}h_l\varphi dm = \int \mathbf{1}_{B^{(k)}}\varphi d\tilde{\mu} = \tilde{\mu}(\mathbf{1}_{B^{(k)}}\varphi),$$

[5]Refer to Lemma 4.3 in [9] for a similar computation.

[6]We incorporate the exponent κ directly into the constants $C_1(B_n,k), C_2(B_n,k)$.

where the third equality follows from the Portmanteau theorem. This fact will be explicitly used in equation (49) below.

3.3. **The perturbative approach.** The quasi-compacity of the operator $\mathcal{L}$ stated in (Theorem 2.5, [10]) and quoted in Section 2 implies that[7]

$$\mathcal{L} = \mu \otimes Z + Q, \tag{28}$$

where again $\mu = \mathcal{L}\mu$ is the SRB measure in $\mathcal{B}$ normalized in such a way that $\mu(1) = 1$ and spans the one-dimensional eigenspace corresponding to the eigenvalue 1; Z is the generator of the one-dimensional eigenspace of $\mathcal{L}^*$ in the dual space $\mathcal{B}^*$ corresponding to the eigenvalue 1, and is normalized in such a way that $Z(\mu) = 1$; and Q is a linear operator on $\mathcal{B}$ with spectral radius $sp(Q)$ strictly less than one. We now introduce the assumptions which allow us to apply the perturbative technique of Keller and Liverani [24]. They are split in two blocks: **A0**, **A2**, and **A3** are needed to get the quasi-compact decomposition (31), which extends to the perturbed operators $\mathcal{L}_n$ the same decomposition for $\mathcal{L}$ required by **A1**. The assumptions **A4** and **A5** together with (31) are finally needed to apply the perturbative technique in [24] we referred to at the beginning of this section.

- **A0** $\mathcal{B}$ is continuously embedded into $\mathcal{B}_w$.
- **A1** The unperturbed operator $\mathcal{L}$ is quasi-compact in the sense expressed by (28).
- **A2** There are constants $0 < \rho < 1, D_1, D_2, D_3 > 0$, and $\rho < M$, such that $\forall n$ sufficiently large, $\forall h \in \mathcal{B}$, and $\forall k \in \mathbb{N}$, we have

$$|\mathcal{L}_n^k h|_w \leq D_1 M^k |h|_w, \tag{29}$$

$$\|\mathcal{L}_n^k h\| \leq D_2 \rho^k \|h\| + D_3 M^k |h|_w. \tag{30}$$

  This will be proved below.
- **A3** We can bound the weak norm of $(\mathcal{L} - \mathcal{L}_n)h$, with $h \in \mathcal{B}$, in terms of the norm of h as

$$|(\mathcal{L} - \mathcal{L}_n)h|_w \leq \chi_n \|h\|$$

  where χ_n is a sequence converging to zero. We give immediately the proof of this fact since it is achieved by a simple adaptation of the computation of the strong stable norm in the proof of item **A2** below. Looking at the notations and at the steps of such a demonstration, we have to control the term $\int_W (\mathcal{L} - \mathcal{L}_n)h\varphi dm = \int_W \mathcal{L}(\mathbf{1}_{B_n} h)\varphi dm = \sum_{i=1,2} \int_{W_i \cap B_n} h(y)\varphi(Ty)\alpha\, dm(y) \leq \|h\|_s |B_n|^\kappa$. Then, $\chi_n = |B_n|^\kappa$.

  Thanks to assumptions **A2** (*uniform Lasota-Yorke inequalities*) and **A3** (*closeness of the operators in the triple norm*), we can apply the spectral theory in [25][8] and get that the decomposition (28) holds for n large enough, namely

$$\lambda_n^{-1} \mathcal{L}_n = \mu_n \otimes Z_n + Q_n, \tag{31}$$

$$\mathcal{L}_n \mu_n = \lambda_n \mu_n, \tag{32}$$

$$Z_n \mathcal{L}_n = \lambda_n Z_n, \tag{33}$$

$$Q_n(\mu_n) = 0, \ \ Z_n Q_n = 0, \tag{34}$$

[7] If φ is a test function, equation (28) means that $(\mathcal{L}h)(\varphi) = Z(h)\mu(\varphi) + Q(h)(\varphi)$.

[8] This spectral theory also requires that if z is in the spectrum of $\mathcal{L}_n$ and $|z| > s$, then z is not in the residual spectrum of $\mathcal{L}_n$. This last fact is guaranteed by **A0**, which implies that the spectral radius of $\mathcal{L}_n$ is bounded by s.

where $\lambda_n \in \mathbb{C}$, $\mu_n \in \mathcal{B}$, $Z_n \in \mathcal{B}^*$, $Q_n \in \mathcal{B}$, and $\sup_n sp(Q_n) \leq sp(Q)$. We observe that the previous assumptions (31)–(34) imply that $Z_n(\mu_n) = 1, \forall n$. Moreover, μ_n can be normalized in such a way that $\mu_n(1) = 1$ and $Z(\mu_n) = 1$; see [24].

We now state assumption **A4**, leaving **A5** to Section 6.1.

- **A4** If we define
$$\Delta_n = Z(\mathcal{L} - \mathcal{L}_n)(\mu), \tag{35}$$
and for $h \in \mathcal{B}$
$$\eta_n := \sup_{||h|| \leq 1} |Z(\mathcal{L}(h\mathbf{1}_{B_n}))|, \tag{36}$$
we must assume that
$$\lim_{n\to\infty} \eta_n = 0, \tag{37}$$
$$\eta_n ||\mathcal{L}(\mathbf{1}_{B_n}\mu)|| \leq \text{const } \Delta_n. \tag{38}$$

Notice that **A0** and **A1** are the content of the aforementioned Lemma 2.4 and Theorem 2.5 in [10]; it remains to prove **A2** and **A4**. The proofs, especially that of **A2**, are quite long and we will defer them to the following sections.

4. Proof of A2.

We start by noticing that the proof we present is also valid for the unperturbed operator, and this will be explicitly used in the following. The proof is basically the same as the proof of Proposition 4.2 in [10], with the difference that we allow subsets of the stable manifolds of length less than γ_a. By the density of $C^1(\mathcal{W}^u)$ in both $\mathcal{B}$ and $\mathcal{B}_w$, it will be enough to take that h is such a smaller space. We have to control integrals of the type $\int_W \mathcal{L}_n h\varphi\, dm$, where $W \in \Sigma$ and $\varphi \in C^1(W)$ (resp. $C^\kappa(W)$), according to the estimate of the weak (resp. strong) norm.

4.1. Weak norm.

Let us start with the weak norm and consider, for instance, $\mathcal{L}_n^3$. We have
$$\int_W \mathcal{L}_n^3 h\varphi dm = \int_W \frac{(h\mathbf{1}_{B_n^c}\mathbf{1}_{B_n^c}\circ T\mathbf{1}_{B_n^c}\circ T^2)(T^{-3}x)\varphi(x)}{\alpha^{-3}\gamma_a^3} dm(x). \tag{39}$$

We successively perform three changes of variable along the backward images of W each with Jacobian γ_a, which will cancel the factor γ_a^3 in the denominator in (39). But, we must now understand how those backward images are produced.

Since $|W| \leq \gamma_a$, its inverse image will give rise to at most two pieces A_1, A_2 of respective lengths a_1, a_2 such that $a_1 + a_2 \leq \frac{|W|}{\gamma_a}$. But now, $T^{-2}W$ is equal to $T^{-2}(A_1 \cup A_2)$ and $T^{-1}(A_1)$ (resp. $T^{-1}(A_2)$) will produce the pieces B_1, B_2 (resp. C_1, C_2). If we denote by $b_{1,2}, c_{1,2}$ the length of those pieces, we have $b_1 + b_2 \leq \frac{a_1}{\gamma_a}, c_1 + c_2 \leq \frac{a_2}{\gamma_a}$.

Our last step consists of iterating backward $B_{1,2}, C_{1,2}$. Each of them will be expanded by a factor γ_a, so we get
$$(39) = \sum_{*\in\{1,2\}} \alpha^3 \int_{T^{-1}B_*} (h\mathbf{1}_{B_n^c}\mathbf{1}_{B_n^c}\circ T\mathbf{1}_{B_n^c}\circ T^2)(x)\varphi(T^3x)dm(x) +$$
$$\sum_{*\in\{1,2\}} \alpha^3 \int_{T^{-1}C_*} (h\mathbf{1}_{B_n^c}\mathbf{1}_{B_n^c}\circ T\mathbf{1}_{B_n^c}\circ T^2)(x)\varphi(T^3x)dm(x),$$
where the measure m is again the unnormalized Lebesgue measure. We now cut the eight intervals $T^{-1}B_*, T^{-1}C_*$ into pieces of length $|W|$. For instance, $T^{-1}B_1$ will give $\frac{b_1}{\gamma_a|W|}$ pieces of length $|W|$ plus two pieces of length less than $|W|$. But, in the

present case, we have to twice add 3 to those pieces for the presence of [9] B_n^c . Then, we split the previous eight integrals over those smaller pieces, which we denote with $\tilde{W}_j^{(n)}$, where $1 \le j \le M_3$, and $M_3 = \frac{1}{\gamma_a |W|}(b_1 + b_2 + c_1 + c_2) + 2^3(1+3)$ is just an upper bound of the number of the smaller pieces after three backward iterations. Then, we can write

$$(39) \le \alpha^3 \sum_{j=1}^{M_3} \int_{\tilde{W}_j^{(n)}} h(x)\varphi(T^3 x) dm(x)$$

Using the preceding bounds on the couples $b_{1,2}, c_{1,2}$ and $a_{L,R}$, we see that $M_3 \le \frac{1}{|W|\gamma_a^3}|W| + 2^3(1+3)$. By iterating backward k times, the cardinality becomes

$$M_k \le \frac{1}{\gamma_a^k} + 2^k(1+k).$$

In order to compute the weak norm of $\mathcal{L}_n^3$, we must take a test function φ verifying $|\varphi|_{C^1(W)} \le 1$. If we now take two points $y_1, y_2 \in \tilde{W}_j^{(n)}$, we have

$$\frac{|\varphi(T^3x) - \varphi(T^3y)|}{|x-y|} = \frac{|\varphi(T^3x) - \varphi(T^3y)|}{|T^3x - T^3y|}\frac{|T^3x - T^3y|}{|x-y|} \le H^1(\varphi)\gamma_a^3 \le H^1(\varphi)$$

where $H^1(\varphi)$ is the Hölder exponent of φ (on W). Therefore,

$$|\varphi \circ T^3|_{C^1(\tilde{W}_j^{(n)})} = |\varphi \circ T^3|_{C^0(\tilde{W}_j^{(n)})} + H^1(\varphi \circ T^3) \le 1.$$

By multiplying and dividing the integral in (39) by $|\varphi \circ T^3|_{C^1(\tilde{W}_j^{(n)})}$, we finally get, for any $k \ge 1$ and remembering that $\alpha = 1/2$,

$$|\mathcal{L}_n^k h|_w \le \left[\left(\frac{\alpha}{\gamma_a}\right)^k + 1 + k\right]|h|_w. \tag{40}$$

Remark 4.1. *The kind of partitioning we consider above, namely by cutting the preimages into pieces of length $|W|$, was not really necessary to estimate the weak norm, but it is particularly adapted to control the strong stable norm, see below. For this reason, we anticipated it here. We will see how one could have proceeded more directly in estimating the strong unstable norm; in this case one gets a weaker bound on the cardinality of the preimages, nevertheless this will not significantly improve the final result.*

4.2. **Strong stable norm.** To compute the strong stable norm, we closely follow the same calculations of Section 4.1 in [10] and we write, still for the third iterate of the perturbed operator and using the notations above,

$$\int_W \mathcal{L}_n^3 h\varphi dm = \alpha^3 \left\{ \sum_{j}^{M_3} \int_{\tilde{W}_j^{(n)}} h(y)[\varphi(T^3y) - \overline{\varphi_{j,n}}]dm(y) + \int_{\tilde{W}_j^{(n)}} h(y)\overline{\varphi_{j,n}} dm(y) \right\}, \tag{41}$$

[9] If we consider higher iterates of $\mathcal{L}$, for instance of order k, we should control terms like $W \cap B_n^c \cap T^{-1}B_n^c \cap \cdots \cap T^{-(k-1)}B_n^c$, where W is a piece of stable manifold. Notice that each preimage $T^{-l}B_n, l = 1, \ldots, k-1$, is contained in 2^l disjoint horizontal rectangles. Therefore, W could meet at most $k-1$ of such rectangles of different generation and hence at most $k-1$ preimages of B_n. This implies that the complement in W of such intersection is at most composed by k connected intervals

where

$$\overline{\varphi_{j,n}} = \frac{1}{|\tilde{W}_j^{(n)}|}\int_{\tilde{W}_j^{(n)}} \varphi(T^3 y)dm(y).$$

Since $|\overline{\varphi_{j,n}}| \leq \sup_W |\varphi|$, we immediately have that the rightmost term in (41) is bounded by the right-hand side of (40). Instead, the first piece on the right-hand side is bounded by

$$\alpha^3 \sum_j^{M_3} \|h\|_s |\varphi \circ T^3 - \overline{\varphi}_{j,n}|_{\tilde{W}_j^{(n)},\kappa} = \sum_j^{M_3} \alpha^3 \|h\|_s |\tilde{W}_j^{(n)}|^\kappa |\varphi \circ T^3 - \overline{\varphi}_{j,n}|_{C^\kappa(\tilde{W}_j^{(n)})} \tag{42}$$

But, $|\varphi \circ T^3 - \overline{\varphi}_{j,n}|_{C^\kappa(W_j^{(n)})} = |\varphi \circ T^3 - \overline{\varphi}_{j,n}|_{C^0(W_j^{(n)})} + \sup_{x\neq y} \frac{|\varphi(T^3x)-\varphi(T^3y)|}{|x-y|^\kappa}$. We now treat the last term on the right-hand side, the first giving the same result after having noticed that $|\varphi \circ T^3 - \overline{\varphi}_{j,n}| = |\varphi(T^3x) - \varphi(T^3x^*)|$, x^* being some point in $W_j^{(n)}$ by the mean value theorem, and having multiplied and divided it by $\frac{|T^3x-T^3y|^\kappa}{|x-y|^\kappa}$. We have

$$\frac{|\varphi(T^3x)-\varphi(T^3y)|}{|x-y|^\kappa} = \frac{|\varphi(T^3x)-\varphi(T^3y)|}{|T^3x-T^3y|^\kappa}\frac{|T^3x-T^3y|^\kappa}{|x-y|^\kappa} \leq$$
$$H^\kappa(\varphi)\gamma_a^{3\kappa} \leq \gamma_a^{3\kappa}|\varphi|_{C^\kappa(W)} = \gamma_a^{3\kappa}|W|^{-\kappa}|\varphi|_{W,\kappa} \leq \gamma_a^{3\kappa}|W|^{-\kappa},$$

where $H^\kappa(\varphi)$ is the Hölder exponent of φ (on W). The sum (42) is therefore bounded by

$$2\alpha^3\gamma_a^{3\kappa}|W|^{-\kappa}\sum_j^{M_3}\|h\|_s|W_j^{(n)}|^\kappa. \tag{43}$$

By construction, all the intervals $|W_j^{(n)}| \leq |W|$[10]. By using the bound on the cardinality of such intervals given by M_k, we finally get

$$\|\mathcal{L}_n^k h\|_s \leq \left[\left(\frac{\alpha}{\gamma_a}\right)^k + 1 + k\right]|h|_w + \|h\|_s\left[2(\alpha\gamma_a^{\kappa-1})^k + 2\gamma_a^{k\kappa}(1+k)\right].$$

4.3. **Strong unstable norm.** In order to treat the strong unstable norm, we follow Section 4.3 in [11] adapted to our case, which is considerably easier. Therefore, we take two stable manifolds $W_{1,2}$ at distance at most ϵ, and φ_i on $W_i, i = 1,2$ with $|\varphi_i|_{C^1(W_i)} \leq 1$. Call $U_1 \subset W_1$ and $U_2 \subset W_2$ the connected intervals parametrized respectively by (s_{W_1}, t) and (s_{W_2}, t), with t belonging to the same interval. We call these two pieces *matched*. We refer to $V_{1,2}$ as the two *unmatched* pieces in $W_{1,2}$. Notice that the length of these two pieces is less than ϵ. Now, by $U_{1,k}^{(j)}, U_{2,k}^{(j)}, j = 1,\dots 2^k$, we define the two preimages of order k of U_1 and U_2, respectively, with the same history, which means that if $s_{U_{1,k}^{(j)}}, s_{U_{2,k}^{(j)}}$ are the common ordinates of the points in respectively $U_{1,k}^{(j)}$ and $U_{2,k}^{(j)}$, then $s_{U_{1,k}^{(j)}}$ and $s_{U_{2,k}^{(j)}}$ belong to the same inverse branch of the map T_Y^k given in (3). Due to the linearity of the map, the sets $U_{1,k}^{(j)}$ and $U_{2,k}^{(j)}$ will again be matched, and $d(U_{1,k}^{(j)}, U_{2,k}^{(j)}) = |s_{U_{1,k}^{(j)}} - s_{U_{2,k}^{(j)}}| \leq \alpha^k d(U_1, U_2) \leq \alpha^k\epsilon$. Since $U_{1,k}^{(j)}$ and $U_{2,k}^{(j)}$ could contain each at most k preimages of the ball B_n, we could have

[10] It is at this point where the partitioning we argued in Remark 4.1 becomes useful.

at most $2k$ matched intervals inside $U_{1,k}^{(j)}$ and $U_{2,k}^{(j)}$. Call $U_{1,k}^{(j,l)}$ and $U_{2,k}^{(j,l)}, l = 1, \dots, 2k$, those smaller matched pieces.

The points of $U_{1,k}^{(j,l)}$ (resp. of $U_{2,k}^{(j,l)}$), will be parametrized as $(s_{1,k}^{(j,l)}, t), t \in U_{1,k}^{(j,l)}$ (resp. $(s_{2,t}^{(j,l)}, t \in U_{2,k}^{(j,l)})$)[11]. We have to control pieces of the type

$$\frac{1}{\epsilon^\beta}\left(\int_{U_{1,k}^{(j,l)}} h(s_{1,k}, t)\varphi_1(T^k(s_{1,k}, t))dt - \int_{U_{2,k}^{(j,l)}} h(s_{2,k}, t)\varphi_2(T^k(s_{2,k}, t))dt\right). \tag{44}$$

We now introduce the auxiliary function

$$\overline{\varphi}_2(s_{W_2}, t) = \varphi_1(s_{W_1}, t), t \in U_1{}^{12}.$$

Then, we bound (44) as

$$\frac{1}{\epsilon^\beta}\left|\int_{U_{1,k}^{(j,l)}} h(s_{1,k}^{(j,l)}, t)\varphi_1(T^k(s_{1,k}^{(j,l)}, t))dt - \int_{U_{2,k}^{(j,l)}} h(s_{2,k}^{(j,l)}, t)\overline{\varphi}_2(T^k(s_{2,k}^{(j,l)}, t))dt\right| +$$
$$\frac{1}{\epsilon^\beta}\left|\int_{U_{1,k}^{(j,l)}} h(s_{2,k}^{(j,l)}, t)[\overline{\varphi}_2(T^k(s_{2,k}^{(j,l)}, t))dt - \varphi_2(T^k(s_{2,k}^{(j,l)}, t))]dt\right| = (I) + (II)$$

We begin by treating the first piece (I): Notice that $T^k(s_{2,k}^{(j,l)}, t)$ is a point of the form $(s_{W_2}, t^*), t^* \in U_2$, and therefore $\overline{\varphi}_2(T^k(s_{2,k}^{(j,l)}, t)) = \overline{\varphi}_2(s_{W_2}, t^*) = \varphi_1(s_{W_1}, t^*) = \varphi_1(T^k(s_{1,k}^{(j,l)}, t))$ since the points $(s_{1,k}^{(j,l)}, t)$ and $(s_{2,k}^{(j,l)}, t)$ are aligned on the same vertical line. Notice now that $|\varphi_1 \circ T^k|_{C^1(U_{1,k})} \le 1$, (we did a similar computation for the strong stable norm), and moreover $d_q(\varphi_1 \circ T^k, \varphi_2 \circ T^k) = 0$. We also have $d(U_{1,k}^{(j,l)}, U_{2,k}^{(j,l)}) \le \alpha^k \epsilon < \epsilon$, which finally implies $(I) \le \alpha^{k\beta} \|h\|_u$. We now pass to estimate (II) using the strong stable norm as

$$(II) \le \frac{1}{\epsilon^\beta}\|h\|_s |U_{1,k}^{(j,l)}|^\kappa |\varphi_2 \circ T^k - \varphi_1 \circ T^k|_{C^\kappa(U_{1,k})}.$$

We now have, using estimates as above,

$$|\varphi_2 \circ T^k - \varphi_1 \circ T^k|_{C^\kappa(U_{1,k})} = |\varphi_2 \circ T^k - \varphi_1 \circ T^k|_{C^0(U_{1,k})} +$$
$$\sup_{y_1, y_2 \in U_{1,k}^{(j,l)}, x \ne y} \frac{|\varphi_2 \circ T^k(y_1) - \varphi_1 \circ T^k(y_1) - \varphi_2 \circ T^k(y_2) + \varphi_1 \circ T^k(y_2)|}{|y_1 - y_2|^\kappa} \le$$
$$|\varphi_2 - \varphi_1|_{C^0(U_1)} + \gamma_a^{k\kappa} H^\kappa(\varphi_1 - \varphi_2) = d_\kappa(\varphi_1, \varphi_2) \le \epsilon.,$$

where H^κ is computed on U_1. Therefore,

$$(II) \le \epsilon^{1-\beta}\gamma_a^{-k\kappa}\|h\|_s.$$

For the unmatched pieces, we have to take into account those generated by the 2^k preimages of $V_{1,2}$, but also the unmatched pieces in the $U_{m,k}^{(j)}, m = 1, 2, j = 1, \dots, 2^k$.

Let us consider first the pieces generated by the 2^k preimages of $V_{1,2}$: Their total number is at most $2k2^k$. If we call one of them $V_{(k)}$ and suppose it belongs

[11] With abuse of notation, $U_{1,k}^{(j,l)}$ and $U_{2,k}^{(j,l)}$ denote the segments of stable manifolds, where $s_{1,t}^{(j,l)}$ (resp. $s_{2,t}^{(j,l)}$) is the common ordinate of the points in $U_{1,k}^{(j,l)}$ (resp. in $U_{2,k}^{(j,l)}$)

[12] Same convention for U_1 as in the previous footnote.

to the backward images of W_1, we must estimate the strong stable norm of the quantity $\frac{1}{\epsilon^\beta}\left|\int_{V_{(k)}} h(y)\varphi(T^k y)dm(y)\right|$. We multiply it by $|V_{(k)}|^\kappa|\varphi\circ T^k|_{C^\kappa(V_k)}$. But, $|\varphi\circ T^k|_{C^\kappa(V_{(k)})}\le|\varphi|_{C^0(W_1)}+H^\kappa(\varphi)\gamma_a^{k\kappa}\le 1$, and $|V_{(k)}|^\kappa\le\epsilon^\kappa\gamma_a^{-k\kappa}$. We now consider the unmatched pieces in $U^{(j)}_{m,k}, m=1,2$. These are generated by the intersections of the preimages of the ball B_n with the preimages of $W_{1,2}$. These intersections could affect one or both of the stable segments $U^{(j)}_{m,k}$. As soon as an intersection occurs, it could generate at most three unmatched pieces (the intersection itself and two short segments on both sides of the intersection). Therefore, we could have at most $6k$ unmatched pieces for the couple $U^{(j)}_{m,k}$. About their size, we use the convexity argument given in Section 6.3 in [11]. If an intersection occurs with one or both the $U^{(j)}_{m,k}$, it also happens between the ball B_n and some backward iterate of order $l\le k$ of the couple $W_{1,2}$. In this case, the intersection will be of order $\sqrt{\epsilon}$, namely $C_{B_n}\sqrt{\epsilon}$ (our B_n is a real ball), where the constant $C_{B_n}<1$ depends on the radius of B_n, and therefore the backward intersections with $U^{(j)}_{m,k}$ will be of order $(\sqrt{\epsilon}\gamma_a^{-1})^k$. Putting these contributions together and asking for

$$\kappa>2\beta,$$

we have, since $\alpha^k=2^{-k}$,

$$\|\mathcal{L}_n^k h\|_u\le 2k\alpha^{k\beta}\|h\|_u+12k\gamma_a^{-k\kappa}\|h\|_s$$

In conclusion, for $k\ge 1$ we get

$$\|\mathcal{L}_n^k h\|=\|\mathcal{L}_n^k h\|_s+b\|\mathcal{L}_n^k h\|_u\le \tag{45}$$

$$\left[\left(\frac{\alpha}{\gamma_a}\right)^k+1+k\right]|h|_w+\left[2(\alpha\gamma_a^{\kappa-1})^k+2\gamma_a^{k\kappa}(1+k)\right]\|h\|_s+b\left(2k\alpha^{k\beta}\|h\|_u+12k\gamma_a^{-k\kappa}\|h\|_s\right). \tag{46}$$

We now put $g_k:=\left[\left(\frac{\alpha}{\gamma_a}\right)^k+1+k\right]$ and $u_\kappa:=\alpha\gamma_a^{\kappa-1}$. Then, we say that $u_\kappa<1$, which needs

$$\kappa>1-\frac{\log\alpha}{\log\gamma_a}.$$

Then, we can rewrite (46) as

$$\|\mathcal{L}_n^k h\|\le g_k|h|_w+[2(2+k)u_\kappa^k+12bk\gamma_a^{-k\kappa}]\|h\|_s+2bk\alpha^{k\beta}\|h\|_u. \tag{47}$$

Then, we choose b such that[13]

$$b<\frac{(2+k)u_\kappa^k}{k\gamma_a^{-k\kappa}},$$

which allows us to rewrite (47) as

$$\|\mathcal{L}_n^k h\|\le g_k|h|_w+4(2+k)u_\kappa^k\|h\|_s+2bk\alpha^{k\beta}\|h\|_u.$$

Then, we pose

$$\sigma=\max\left(u_\kappa,\alpha^\beta\right)<1,$$

which gives

$$\|\mathcal{L}_n^k h\|\le g_k|h|_w+r_k\sigma^k\|h\|,$$

[13] We will see in a moment that this choice will be done for a particular k.

where we set $r_k = 4(2+k)$.

We now fix a value of k, say k_0, such that

$$\rho := (r_{k_0}\sigma^{k_0})^{\frac{1}{k_0}} < 1$$

and we replace k with k_0 in the bound above for b. With these positions and by using blocks of length k_0, it is immediate to rewrite (46) for any $k > 0$ as

$$\|\mathcal{L}_n^k h\| \le \rho^k \|h\| + M^k |h|_w,$$

where $M := g_{k_0}^{\frac{1}{k}0}(1 - r_{k_0}\sigma^{k_0})^{-1}$, and this proves (30).

Remark 4.2. *We summarize the bounds we imposed on the relevant quantities we used up to now: We have, since $\alpha = 1/2$:*

- $0 < \beta < 1 - \kappa$ *and* $2\beta < \kappa$. *This first requires* $\beta < 1/3$.
- $\beta + q < 1$, *with* $q \in (0,1)$, *and* $q > 5/6$, *which implies* $\beta < 1/6$.
- $\kappa > 1 - \frac{\log \alpha}{\log \gamma_a} = 1 + \frac{\log 2}{\log \gamma_a}$.
- *Finally, we will see below that* $\kappa > \frac{\alpha \log \alpha^{-1} + (1-\alpha)\log(1-\alpha)^{-1}}{\log \gamma_a^{-1}} = -\frac{\log 2}{\log \gamma_a}$.

  This is verified by several couples of the parameters κ, γ_a. *For instance, for any* $1/2 < \kappa < 1$, *we could take* $\gamma_a = 1/4$. *Alternatively, by choosing* $\kappa = 3/4$, *we could take* $2^{-4} < \gamma_a < 2^{-4/3}$.

Warning: From now on we will consider the baker's map (1) with the parameters $\alpha = v = 0.5$ and γ_a satisfying the constraints given in the previous remark 4.2.

5. Proof of A4.

We now aim to justify **A4.** We remind that Z is the unique solution of the eigenvalue equation $\mathcal{L}^* Z = Z$, where $\mathcal{L}^*$ is the dual of the transfer operator. By setting

$$Z(h) := h(1),\ h \in \mathcal{B}, \tag{48}$$

we have for $h \in \mathcal{B}$,

$$\mathcal{L}^* Z(h) = Z(\mathcal{L}h) = (\mathcal{L}h)(1) = h(1 \circ T) = h(1) = Z(h).$$

Coming back to Δ_n, we see immediately that

$$\Delta_n = Z(\mathcal{L}(\mathbf{1}_{B_n}\mu)) = \mathcal{L}(\mathbf{1}_{B_n}\mu)(1) = \int \mathbf{1}_{B_n}\, d\mu = \mu(B_n). \tag{49}$$

The term $\|\mathcal{L}(\mathbf{1}_{B_n}\mu)\|$ can be handled very easily using the Lasota-Yorke inequality which we proved in item **A2** above. In fact, it follows from (30) that there are two constants C_1, C_2 depending only on the map such that

$$\|\mathcal{L}(\mathbf{1}_{B_n}\mu)\| \le C_1 \|\mathbf{1}_{B_n}\mu\| + C_2 |\mathbf{1}_{B_n}\mu|_w.$$

Lemma 5.1. *There exists two constants $\hat{C}_1, \hat{C}_2$ independent of n such that*

$$\|\mathbf{1}_{B_n}\mu\| \le \hat{C}_1 \|\mu\| \ \ \text{and} \ \ |\mathbf{1}_{B_n}\mu|_w \le \hat{C}_2 |\mu|_w. \tag{50}$$

Proof. The proof follows closely the arguments given in Section 3.2.2 to bound the quantity $\mathbf{1}_{B^{(k)}} h$, but we should now be more careful in getting constants which do not depend about upon B_n. Notice that, contrary to Section 3.2.2, we take here the intersection of W with the ball B_n, not with its preimages. With this in mind, it is immediate to check that $\|\mathbf{1}_{B_n}\mu\|_s \le \|\mu\|_s$ and $|\mathbf{1}_{B_n}\mu|_w \le |\mu|_w$. It remains to compute the strong unstable norm, and this reduces to bound the difference for a smooth

h: $\mathfrak{D}_n := \frac{1}{\epsilon^\beta}\mathfrak{I}_n$, where $\mathfrak{I}_n := \left|\int_{W_1\cap B_n} h\varphi_1 dm - \int_{W_2\cap B_n} h\varphi_2 dm\right|$, and $W_1 W_2, \varphi_1$, and φ_2 verify the constraints given in (9). We split the argument into in two parts. We call r_n the radius of the ball B_n, and we begin to take $\epsilon \geq 0.5r_n$. Then, we have the rough bound

$$\mathfrak{I}_n \leq 2\ r_n\ \|h\|_s \ \Rightarrow\ \mathfrak{D}_n \leq 2^{1+\kappa}\|h\|_s\epsilon^{\kappa-\beta} < 2^{1+\kappa}\|h\|_s,$$

since $\beta < \kappa$. Then we consider $\epsilon < 0.5r_n$;. We split the difference in $\mathfrak{I}_n$ over unmatched and matched pieces. There could be at most one matched piece inside B_n giving the contribution $\|h\|_u$. If the matched piece is inside B_n, there could be at most two unmatched pieces. They have length $\leq \epsilon$ if they are on the extremities of the two stable manifolds inside B_n. Otherwise, they are generated when the two stable manifolds meet the boundary of B_n. It is a simple exercise to see that the sum of the lengths of those unmatched pieces is bounded by the maximum difference of the lengths of two horizontal chords whose vertical distance is ϵ, and that value is less then or equal to $2\sqrt{2r_n\epsilon} \leq 2\sqrt{2\epsilon}$.

Finally, there could be an unmatched piece when only one manifold crosses B_n, and for the same argument as above, its length is bounded by $2\sqrt{2\epsilon}$. Summing all those contributions, when $\epsilon < 0.5r_n$, we get

$$\mathfrak{I}_n \leq \epsilon^\beta\|h\|_u + 2\epsilon^\kappa\|h\|_s + 2^{1+\frac{\kappa}{2}}\epsilon^{\frac{\kappa}{2}}\|h\|_s \Rightarrow\ \mathfrak{D}_n \leq \|h\|_u + 2\|h\|_s + 2^{1+\frac{\kappa}{2}}\|h\|_s,$$

since $2\beta < \kappa$. □

By setting

$$C_3 := C_1\hat{C}_1\|\mu\| + C_2\hat{C}_2|\mu|_w,$$

we are led to prove that (see (38)) $\eta_n C_3 \leq \text{const}\ \Delta_n$, namely

$$\eta_n \leq\ \text{const}\ \Delta_n = \text{const}\ \mu(B_n). \tag{51}$$

Before continuing, we have to focus on $\mu(B_n) = \mu(B(z, e^{-u_n}))$. It is well known that for μ-almost z and by taking the radius sufficiently small, depending on the value ι, $e^{-u_n(d+\iota)} \leq \mu(B(z, e^{-u_n}) \leq e^{-u_n(d-\iota)}$, where $\iota > 0$ is arbitrarily small. This follows from the existence of the limit

$$\lim_{r\to 0^+}\frac{\log\mu(B(x,r))}{\log r} = d, \text{ for } x \text{ chosen } \mu\text{-a.e.}, \tag{52}$$

and the quantity d is the Hausdorff dimension of the measure μ, which, in our case reads as [27], equation (3.24):

$$d = 1 + d_s, \text{ where } d_s := \frac{\alpha\log\alpha^{-1} + (1-\alpha)\log(1-\alpha)^{-1}}{\log\gamma_a^{-1}}.$$

We now have the following lemmma.

Lemma 5.2. *Assume $\kappa > d_s$.*
Then,

$$\eta_n \leq 2\lceil\gamma_a^{-1}\rceil\mu(B_n).$$

Proof. By density, it will be enough to prove the lemma for $h \in C^1(X)$. We have

$$Z(\mathcal{L}(h\,\mathbf{1}_{B_n})) = \int h\,\mathbf{1}_{B_n} dm.$$

By disintegrating along the stable partition $\mathcal{W}^s$, we get,

$$\int h\, \mathbf{1}_{B_n}\, dm_L = \int_\xi d\lambda(\xi) \left[\int_{W_\xi} (\mathbf{1}_{B_n} h)(x)\, dm(x) \right]. \tag{53}$$

We now cut the stable manifold W_ξ in pieces of length γ_a in order to compute the strong stable norm on each of them, and we take $|\tilde{W}_\xi|$ as the largest intersection of such pieces with B_n. We immediately get

$$(53) \le \quad \int_\xi d\lambda(\xi) \left[\lceil \gamma_a^{-1} \rceil |\tilde{W}_\xi|^\kappa \|h\|_s \right] \le e^{-u_n \kappa} \lceil \gamma_a^{-1} \rceil \|h\|_s \lambda(\xi; B_n \cap W_\xi \neq \emptyset),$$

where λ is the quotient measure on the space of stable leaves W_ξ belonging to $\mathcal{W}^s$, and indexed by ξ; see for instance [28], Appendix A. By the definition of disintegration, we have that

$$\lambda(\xi; B_n \cap W_\xi \neq \emptyset) = m_L(\bigcup W_\xi, B_n \cap W_\xi \neq \emptyset) = 2e^{-u_n},$$

and therefore

$$\eta_n \le 2\lceil \gamma_a^{-1} \rceil e^{-u_n(\kappa+1)}.$$

We finally have

$$\eta_n \le 2\lceil \gamma_a^{-1} \rceil e^{-u_n(\kappa+1)} \le 2\lceil \gamma_a^{-1} \rceil e^{-u_n(d+\iota)} \le 2\lceil \gamma_a^{-1} \rceil \mu(B_n),$$

provided we choose

$$\kappa > d + \iota - 1 \tag{54}$$

which can be satisfied by assumption. □

Remark 5.3. *The local comparison between the Lebesgue and the SRB measure of a ball of center z obliged us to choose z μ-almost everywhere because, in this way, we have a precise value for the locally constant dimension d. We are therefore discarding several points, possibly periodic, where the limiting distribution for the Gumbel law (see the next section) could exhibit extremal indices different from* 1.

Remark 5.4. *For invertible, piecewise differentiable hyperbolic maps in dimension* 2*, the construction of the Banach space imposes that $\kappa < 1$; for billiard maps associated with Lorentz gases,* [12]*, it even verifies $\kappa \le 1/6$. This could make difficult to check condition (54) for invariant sets with large d, like Anosov diffeomorphisms for instance. In some sense, this difficulty was already raised in Section 4.5 in the Keller's paper* [23]*, where an estimate like ours in terms of the Hölder exponent κ was given, and the subsequent question of the comparison with the SRB measure was addressed.*

6. **The limiting law.**

6.1. **Gumbel law.**

We have now all the tools to compute the asymptotic behavior of $\mathcal{L}_n$. We need one more ingredient which will constitute our last assumption:

- **A5** Let us suppose that the following limit exists, for any $k \ge 0$

$$q_k = \lim_{n\to\infty} q_{k,n} := \lim_{n\to\infty} \frac{Z\left([(\mathcal{L} - \mathcal{L}_n)\mathcal{L}_n^k(\mathcal{L} - \mathcal{L}_n)]\mu\right)}{\Delta_n} \tag{55}$$

Notice that

$$q_{k,n} = \frac{\mu(B_n \cap T^{-1}B_n^c \cap \cdots \cap T^{-k}B_n^c \cap T^{-(k+1)}B_n)}{\mu(B_n)}$$

and therefore, by the Poincaré recurrence theorem,

$$\sum_{k=0}^{\infty} q_{k,n} = 1.$$

Therefore, if the limits (55) exist, the quantity

$$\theta = 1 - \sum_{k=0}^{\infty} q_k, \tag{56}$$

is well defined and verifies

$$0 \le \theta \le 1.$$

It is called the *extremal index*, and it modulates the exponent of the Gumbel law as we will see in a moment. We have, in fact, by Theorem 2.1 of [24]

$$\lambda_n = 1 - \theta\Delta_n = \exp(-\theta\Delta_n + o(\Delta_n)),$$

or equivalently

$$\lambda_n^n = \exp(-\theta n\Delta_n + no(\Delta_n)).$$

Therefore, we have

$$\mu(M_n \le u_n) = \mathcal{L}_n^n \mu(1) = \lambda_n^n[\mu_n(1)Z_n(\mu) + Q_n^n(\mu)(1)],$$

and consequently

$$\mu(M_n \le u_n) = \exp(-\theta n\Delta_n + no(\Delta_n))[O(1) + Q_n^n(\mu)(1)],$$

since $\mu_n(1) = 1$ and it has been proved in [24], Lemma 6.1, $Z_n(\mu) \to 1$ for $n \to \infty$. At this point, we need an important assumption, which basically reduces to fixing the sequence u_n and allow us to get a non-degenerate limit for the distribution of M_n. We in fact ask that

$$n\,\Delta_n \to \tau,\ n \to \infty, \tag{57}$$

where τ is a positive real number. With this assumption, using (10) and (11), we have

$$|Q_n^n(\mu)(1)| \le \text{const } sp(Q)^n \|\mu\| \to 0.$$

In conclusion, we get the Gumbel law

$$\lim_{n\to\infty} \mu(M_n \le u_n) = e^{-\theta\tau}.$$

6.2. **The extremal index.** We are now ready to compute the $q_{k,n}$, which will determine the extremal index. Let us first suppose that the center of the ball B_n is not a periodic point. Then, the points $T^j(z), j = 1, \cdots, k$ will be disjoint from z. Let us take the ball so small that is does not cross the set $T^j\Gamma, j = 1, \cdots, k$, where Γ is the discontinuity line ($y = \alpha$). In this way, the images of B_n will be ellipses with the long axis along the unstable manifold and the short axis stretched by a factor γ^k. By continuity and taking n large enough, we can manage that all the iterates of B_n up to T^k will be disjoint from B_n, and for such, n the numerator of $q_{k,n}$ will be zero. At this point, we can state the following result.

Proposition 6.1. *Let T be the baker's transformation and consider the function $M_n(x) := \max\{\Xi(x), \ldots, \Xi(T^{n-1}x)\}$, where $\Xi(x) = -\log d(x,z)$, and z is chosen μ-almost everywhere with respect to the SRB measure μ. Then, if z is not periodic, we have*

$$\lim_{n\to\infty} \mu(M_n \le u_n) = e^{-\tau},$$

where the boundary level u_n is chosen to satisfy $n\mu(B(z, e^{-u_n})) \to \tau$ for some positive τ.

Suppose now that z is a periodic point of minimal period p. Of course, the next considerations make sense if the limit (52) exists. By doing as above, we can stay away from the discontinuity lines up to p iterates and look simply to $T^{-p}(B_n) \cap B_n$. Since the map acts linearly, thepreimage p of B_n would be an ellipse with center z and symmetric w.r.t. the unstable manifold passing through z. So, we have to compute the SRB measure of the intersection of the ellipse with the ball shown in Figure 2.

It turns out that this computation is not easy. The natural idea would be to disintegrate the SRB measure along the unstable manifolds belonging to the unstable partition $\mathcal{W}^u$. We index such fibers as W_ν, and we set $\zeta(\nu)$ as the associated quotient measure. Let us recall that the conditional measures along leaves W_ν are normalized Lebesgue measures, which we denote with l_ν. If we call $\mathcal{E}_{in}$ the region of the ellipse inside the ball B_n, we have to compute

$$\frac{\int l_\nu(\mathcal{E}_{in} \cap W_\nu)\, d\zeta(\nu)}{\int l_\nu(B_n \cap W_\nu)\, d\zeta(\nu)}. \tag{58}$$

Although simple geometry allows us to compute easily the length of $\mathcal{E}_{in} \cap W_\nu$ and $B_n \cap W_\nu$, and since they vary with W_ν, it is not at the end clear how to perform the integral with respect to the quotient measure, especially because we need asymptotic estimates, not bounds. We therefore proceed by introducing a different metric, a nice trick which was already used in [8]. We use the l^∞ norm on $\mathbb{R}^2$ for which $|(x,y)|_\infty = \max\{|x|,|y|\}$. In this way, the ball B_n will become a square with sides of length $r_n := e^{-u_n}$ and $T^{-p}(B_n)$ will be a rectangle with the long side of length $\gamma_a^{-p} r_n$ and the short side of length $\alpha^p r_n$. This rectangle will be placed symmetrically with respect to the square as indicated in Figure 3. The ratio (58) can now be computed easily since the length in the integrals are constant, and we get α^p. We will see that the value computed in this way is the right one, see Proposition 7.3, but in principle we *cannot apply the spectral technique* since the geometric shape of the rectangles does not allow to show that the characteristic function of such rectangles is in $\mathcal{B}$, and also it does not fit the convexity requirement which we used to control the unmatched pieces in the strong unstable norm. We will introduce in Section 7.2 below a different technique which will allow us to get the extremal index even when the target sets are rectangles.

7. **Poisson statistics.**

7.1. **The spectral approach.**

As mentioned in the introduction, the spectral technique has been recently generalized to study the statistics of the number of visits in balls shrinking around a point [3]. We briefly introduce such an approach, and the reader will see that we can easily adapt it to the baker's map. The starting point is to consider the following counting function:

$$N_{B_n}^{\tau}(x) = \sum_{i=0}^{\lfloor \tau/\mu(B_n) \rfloor} \mathbf{1}_{B_n} \circ T^i(x),$$

where τ is a positive parameter and $x \in X$. The goal is to study the distribution of this discrete random variable in the limit $n \to \infty$. With the spectral approach, we will rather look at the characteristic function of such a variable.

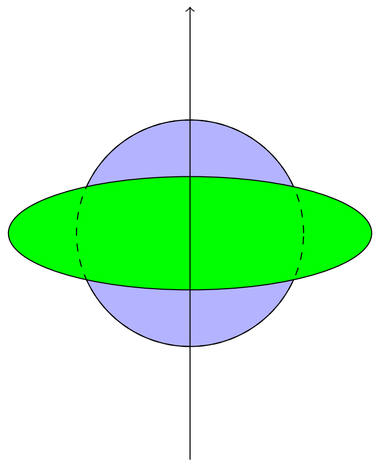

FIGURE 2. Computation of the extremal index around periodic point with the Euclidean metric. The vertical line is an unstable manifold. We should compute the green area inside the circle.

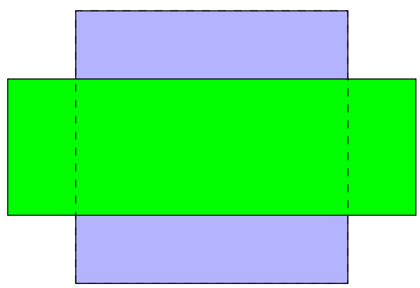

FIGURE 3. Computation of the extremal index around periodic point with the l^∞ metric. We should compute the green area inside the square.

We begin to define $S_{n,k} := \sum_{i=0}^{k} \mathbf{1}_{B_n} \circ T^i$ and take $S_{n,(\tau,n)} := N_{B_n}^{\tau}$. We then define the perturbed operator

$$\mathcal{L}_{n,s}(h) = \mathcal{L}(e^{is\mathbf{1}_{B_n}} h), \ s \in \mathbb{R}, \ h \in \mathcal{B}.$$

A simple computation shows that

$$\mathcal{L}_{n,s}^{k}(\mu)(1) = \int e^{isS_{n,k}} d\mu,$$

which suggests getting information on the characteristic function of $S_{n,k}$ by the behavior of the top eigenvalue $\lambda_{n,s}$ of the perturbed operator $\mathcal{L}_{n,s}$. At this point, the analysis proceeds in the same manner as for the perturbed operator $\mathcal{L}_n$, and we sketch here the main steps. The *difference* between the two operators is now quantified by

$$\Delta_{n,s} := Z(\mathcal{L} - \mathcal{L}_{n,s})(\mu) = (1 - e^{is})\mu(B_n),$$

and

$$\lambda_{n,s} = 1 - \tilde{\theta}(s)(1 - e^{is})\mu(B_n) + o(\mu(B_n)). \tag{59}$$

The quantity $\tilde{\theta}(s)$ plays the role of the extremal index, and is defined according to the formula (55), which in the present case reduces to $\tilde{\theta}(s) = 1-\sum_{k=0}^{\infty} q_k(s)$, where

$$q_k(s) = \lim_{n\to\infty} \frac{1}{1-e^{is}} \sum_{\ell=0}^{k} (1-e^{is})^2 e^{i\ell s} \beta_n^{(k)}(\ell) = (1-e^{is}) \sum_{\ell=0}^{k} e^{i\ell s} \beta_k(\ell), \tag{60}$$

$$\beta_n^{(k)}(\ell) := \frac{\mu(x; x\in B_n, T^{k+1}(x)\in B_n, \sum_{j=1}^{k} \mathbf{1}_{B_n}(T^j x) = \ell)}{\mu(B_n)}. \tag{61}$$

and we suppose that the limit $\beta_k(\ell) := \lim_{n\to\infty} \beta_n^{(k)}(\ell)$ exists. Then, we have

$$\tilde{\theta}(s) = 1-(1-e^{is}) \sum_{k=0}^{\infty} \sum_{\ell=0}^{k} e^{i\ell s} \beta_k(\ell),$$

and the exponential decay of correlation of the measure μ allows us to show that the series $\sum_{k=0}^{\infty} \sum_{\ell=0}^{k} \beta_k(\ell)$ converges absolutely[14], and therefore $\tilde{\theta}(s)$ is C^∞ in the neighborhood of 0. If we now return to the eigenvalue (59), we exponentiate it to the power n, and we again using the threshold condition (57), $n\mu(B_n) \to \tau$, we finally get

$$\lim_{n\to\infty} \int e^{isS_{n,(\tau,n)}} d\mu = e^{-\tilde{\theta}(s)(1-e^{is})\tau} := \Sigma(s).$$

Since $\Sigma(s)$ is continuous in $s = 0$, it is the characteristic function of some random variable $\mathfrak{Z}$, possibly defined on a different probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The variable $\mathfrak{Z}$ is clearly non-negative and integer valued and it is also infinitely divisible since $e^{-\tilde{\theta}(s)(1-e^{is})\tau} = (e^{-\tilde{\theta}(s)(1-e^{is})\tau/m})^m$, for any m. This implies that $\mathfrak{Z}$ has a compound Poisson (CP) distribution, see [14] or [3] for more references; namely, it may be written as $\mathfrak{Z} := \sum_{j=1}^{N} X_j$, where the X_j are i.i.d. random variables defined on same probability space, and N is Poisson distributed with *intensity* $\varkappa$ and X_j has distribution $\mathbb{P}(X_j = l) = \rho_l$; moreover, N is independent of all X_j. We call the sequence $(\rho)_{l\geq 1}$ the *cluster size distribution* of Z. Among the CP distributions, two are particularly important: the standard Poisson distribution and the Pòlya-Aeppli distribution. For the standard Poisson $\rho_1 = 1$, for Pòlya-Aeppli the distribution of X_j is geometrical, namely $\rho_l = \eta(1-\eta)^l, \eta \in (0,1)$. For such distributions the associated characteristic functions are perfectly known. To determine them for our baker's system, one should prove the existence and compute the quantities (61), which are of geometric and dynamical nature. This will be done in the next section in the context of a more probabilistic approach to Poisson-like statistics. Actually, the quantities computed in the next section are not exactly those in (61), but it is not difficult to modify their derivation to get (61) and therefore reprove Proposition 7.2 with the spectral approach. As we said in the introduction and in Section 6.2, we will present the alternative probabilistic approach since it will allow us to cover a wider class of target sets and also to get example 7.5, which shows a CP distribution different from the standard Poisson and the Pòlya-Aeppli.

7.2. **The probabilistic approach.** We now use a recent technique developed in [21] and apply it to our baker's map. We will recover the usual dichotomy and get a pure Poisson distribution when the points are not periodic, and a Pólya-Aeppli

[14] See Section 3 in [3] for the proof of this convergence which applies to our case as well.

distribution around periodic points with the parameter giving the geometric distribution of the size of the clusters which coincide with the extremal index computed in the preceding section. This last result is achieved in particular if we use the l^∞ metric. This result is not surprising. What is interesting is the great flexibility of the technique of the proof, which allows us to easily get the expected properties. In order to apply the theory in [21], we need to verify a certain number of assumptions, but we otherwise refer to the aforementioned paper for precise definitions. Here, we recall the most important requirements and prove in detail one of them.

Warning: The next considerations are carried over with the Euclidean metric which is more natural for applications. As for the visits to periodic points, we will use the l^∞ metric and the following computations are even easier.

Decay of correlation. There exists a decay function $\mathcal{C}(k)$ so that

$$\left|\int_M G(H\circ T^k)\,d\mu-\mu(G)\mu(H)\right|\le \mathcal{C}(k)\|G\|_{Lip}\|H\|_\infty \qquad \forall k\in\mathbb{N},$$

for functions H which are constant on local stable leaves W_s of T and the functions $G:M\to\mathbb{R}$ being Lipschitz continuous. This is ensured by Theorem 2.5 in [10], where the role of H is taken by the test functions in $C^\kappa(W,\mathbb{C})$ and $G\in\mathcal{B}$, which is the completion of Lipschitz functions on X. The decay is exponential.

Cylinder sets. The proof requires the existence, for each $n\ge 1$, of a partition of each unstable leaf in subsets $\xi_n^{(k)}$, called n-cylinders (or cylinders of rank n), and indexed with k, where T^n is defined and the image $T^n\xi_n^{(k)}$ is an unstable leaf of full length for each k. These cylinders are obtained by taking the 2^n preimages of $\Gamma=\{y=\alpha\}$ by the map T_Y restricted to each leaf. In the following, we will take $\alpha=1/2$ to simplify the exposition.

Exact dimensionality of the SRB measure. This uses the existence of the limit (52). We shall need the following result.

Lemma 7.1. (Annulus type condition) *Let $w>1$. If x is a point for which the dimension limit (52) exists for a positive d, then there exists a $\delta>0$ so that*

$$\frac{\mu(B(x,r+r^w)\setminus B(x,r))}{\mu(B(x,r))}=O(r^\delta),$$

for all $r>0$ small enough.

Now we can apply the results of Section 7.4 in [21] to prove the following result, which tracks the number of visits a trajectory of the point $x\in X$ makes to the set U on a suitable normalized orbit segment.

Proposition 7.2. *Consider the counting function*

$$N^\tau_{B_n}(z)=\sum_{i=0}^{\lfloor \tau/\mu(B_n)\rfloor}\mathbf{1}_{B_n}\circ T^i(x),$$

where τ is a positive parameter and z is a point for which limit (52) exists and $n\mu(B(z,e^{-u_n}))\to\tau$.

- *If z is not a periodic point, using the Euclidean metric we get a pure Poisson distribution:*

$$\mu(N^\tau_{B_n}=k)\to\frac{e^{-\tau}\tau^k}{k!},\ n\to\infty.$$

- *If z is a periodic point of minimal period p, using the l^∞ metric we get a compound Poisson distribution (Pólya-Aeppli):*

$$\mu(N_{B_n}^\tau = k) \to e^{-\Theta\tau} \sum_{j=1}^{k} (1-\Theta)^{k-j}\Theta^{2j}\frac{\tau^j}{j!}\binom{k-1}{j-1},\ n\to\infty,$$

where Θ is given by $\Theta = 1 - \lim_{n\to\infty}\frac{\mu(T^{-p}B_n\cap B_n)}{\mu(B_n)}$.

Proof of Lemma 7.1. We have to prove the lemma in the two cases when (I) the norm is ℓ^2 and (II) the norm is ℓ^∞ and the ball is geometrically a square.
(I) We now use the Euclidean metric and denote by $\mathcal{A}$ the annulus $\mathcal{A} = B(x, r + r^w)\setminus B(x,r)$ where $w > 1$. By disintegrating the SRB measure along the unstable manifolds, we have

$$\mu(\mathcal{A}) = \int l_\nu(\mathcal{A}\cap W_\nu)\, d\zeta(\nu).$$

We now split the subsets on each unstable manifold on the cylinders of rank n and condition with respect to the Lebesgue measure on them[15]:

$$l_\nu(\mathcal{A}\cap W_\nu) = \sum_{\xi_n;\xi_n\cap\mathcal{A}\neq\emptyset} \frac{l_\nu(\mathcal{A}\cap W_\nu\cap\xi_n)}{l_\nu(\xi_n)} l_\nu(\xi_n). \tag{62}$$

We then iterate forward each cylinder with T^n. They will become of full length equal to 1, and subsequently we get $l'_\nu(T^n\xi_n) = 1$ for some $W_{\nu'}$. Since the action of T is locally linear and expanding by a factor 2^n (with the given choice of $\alpha = \frac{1}{2}$) on the unstable leaves and therefore has zero distortion, we have

$$\frac{l_\nu(\mathcal{A}\cap W_\nu\cap\xi_n)}{l_\nu(\xi_n)} = \frac{l_{\nu'}(T^n(\mathcal{A}\cap W_\nu\cap\xi_n))}{l_{\nu'}(T^n\xi_n)} = l_{\nu'}(T^n(\mathcal{A})\cap W_{\nu'}),$$

so that $T^n(\mathcal{A}\cap W_\nu\cap\xi_n)\subset W_{\nu'}$. Therefore,

$$l_\nu(\mathcal{A}\cap W_\nu) = \sum_{\xi_n;\xi_n\cap\mathcal{A}\neq\emptyset} l_{\nu'}(T^n(\mathcal{A}\cap W_\nu\cap\xi_n)) l_\nu(\xi_n).$$

By elementary geometry, we see that the largest intersection of $\mathcal{A}$ with the unstable leaves will produce a piece of length $O(r^{\frac{w+1}{2}})$. Therefore, $l_{\nu'}(T^n(\mathcal{A}\cap W_\nu\cap\xi_n)) = O(2^n r^{\frac{w+1}{2}})$, and

$$\mu(\mathcal{A}) = O(2^n r^{\frac{w+1}{2}}) \int \sum_{\xi_n;\xi_n\cap\mathcal{A}\neq\emptyset} l_\nu(\xi_n)\, d\zeta(\nu).$$

We now observe that in order to have our result, it will be enough to get it with a decreasing sequence r_n, $n\to\infty$, of exponential type, $r_n = b^{-t(n)}, b > 1$, and $t(n)$ increasing to infinity. We put $r_n = 2^{-n}$. With this choice and remembering that 2^{-n} is also the length of the n-cylinders, we have

$$\bigcup_{\xi_n;\xi_n\cap\mathcal{A}\neq\emptyset} \xi_n \subset B(x, r_n + r_n^w + 2^{-n}) = B(x, 2r_n + r_n^w) \subset B(x, 3r_n),$$

which, as the cylinders ξ_n are disjoint, yields the estimate for the integral above:

$$\mu(\mathcal{A}) = O(2^n r_n^{\frac{w+1}{2}} r_n^{d-\epsilon}).$$

[15]We simply use here ξ_n instead of $\xi_n^{(k)}$ since the computation over k is replaced by the sum.

Now, by the exact dimensionality of the SRB measure, one has for any $\varepsilon > 0$ and by renaming r_n as r

$$(2r + r^w)^{d+\varepsilon} \leq \mu(B(x, 2r + r^w)) \leq (2r + r^w)^{d-\varepsilon}$$

for all r small enough, i.e. n large enough. With this, we can divide $\mu(\mathcal{A})$ by the measure of the ball of radius r, and obtain the estimate

$$\frac{\mu(\mathcal{A})}{\mu(B(x,r))} = O(r^{\frac{w-1}{2}+d-\varepsilon-d-\varepsilon}) = O(r^{\frac{w-1}{2}-2\varepsilon}) = O(r^{\frac{w-1}{4}}),$$

since $w > 1$, and provided ε is small enough.
(II) Now we shall use the ℓ^∞-distance and again denote by $\mathcal{A}$ the annulus $B(x, r + r^w) \setminus B(x, r)$. Since we are in two dimensions, we can cover the annulus by balls $B(y_j, 2r^w)$ of radii $2r^w$, with centers y_j for $j = 1, \ldots, N$. The number of balls needed N is bounded by $8\frac{r}{r^w}$. For any $\varepsilon > 0$, there exists a constant c_1 so that $\mu(B(y_j, 2r^w)) \leq c_1 r^{w(d-\varepsilon)}$ for all r small enough. Thus,

$$\mu(\mathcal{A}) \leq 8c_1 r^{1+w(d-1-\varepsilon)}$$

and since $\mu(B(x,r)) \geq c_3 r^{d+\varepsilon}$ for some $c_3 > 0$, we obtain

$$\frac{\mu(\mathcal{A})}{\mu(B(x,r))} \leq c_4 r^{(d-1)(w-1)-\varepsilon(w+1)}.$$

The exponent $\delta = (d-1)(w-1) - \varepsilon(w+1)$ is positive as $d, w > 1$ and $\varepsilon > 0$ can be chosen sufficiently small. □

Proof of Proposition 7.2. We can now prove the proposition by applying Theorem 1 from [21] and verify its Assumptions (I) to (VI) as follows:
(I) Let $\mathcal{I}_n$ be the collection of inverse branches φ of the n-th iterate T^n of the map T. Then, evidently, if $\varphi, \varphi' \in \mathcal{I}_n$ are two distinct inverse branches, their intersection $\varphi(X) \cap \varphi'(X)$ has zero measure. Therefore, the number of overlaps of 'n-cylinders' $\varphi(X)$ is bounded by $L = 1$.
(II) This condition is easily satisfied since decay of correlations is exponential as the transfer operator is quasi compact.
(III) The set $\mathcal{G}_n$ of uniform expansion covers the entire space X as there is no 'bad' set $\mathcal{G}_n^c$ of non-uniformly contracting inverse branches. Consequently, we get exponential contraction of $\max_{\varphi \in \mathcal{I}_n} \mathrm{diam}\varphi(X)$ of the n-cylinder sets $\varphi(X)$. Moreover, the distortion $\sup_{\varphi \in \mathcal{I}} \sup_{x,y \in \varphi(X)} \frac{J_n(x)}{J_n(y)}$ is uniformly bounded, where J_n is the Jacobian of T^n restricted to the unstable direction.
(IV) The dimension of the invariant measure is equal to $d = 1 + d_s$, where $d_s < 1$ is given above. So, we can choose $d_0 > 0$ and $d_1 < \infty$ so that $d_0 < d < d_1$.
(V) The dimension of the restricted measure on the unstable leaves equals $u_0 = 1$ as it is Lebesgue.
(VI) This condition was verified in Lemma 7.1.
This shows that the condition of Theorem 1 of [21] is satisfied.

If x is an aperiodic point, then $\min\{j \geq 1 : B_\rho(x) \cap T^j B_\rho(x) \neq \emptyset\}$ goes to infinity as $\rho = e^{-u_n} \to 0$. Thus, for the coefficients

$$\lambda_\ell(L) = \lim_{\rho \to 0} \frac{\mathbb{P}(Z^L = \ell)}{\mathbb{P}(Z^L \geq 1)}$$

we obtain that for every L, $\lambda_1 = 1$ and $\lambda_\ell = 0$ for all $\ell = 2, 3, \ldots$, where $Z^L = \sum_{j=1}^{L} \mathbf{1}_{B_\rho(x)} \circ T^j$ is the hit counter on the finite orbit segment of length L. This implies that $N^{\tau}_{B_n}$ converges in distribution to a standard Poisson random variable with parameter τ.

Let x be a periodic point with minimal period p, and let $\tilde{B}_\rho$ be a square of size ρ centered at x and whose sides are aligned with the stable and unstable directions, respectively. Then, for $\ell = 2, 3, \ldots$

$$\hat{\alpha}_\ell = \lim_{L\to\infty}\lim_{\rho\to 0}\mathbb{P}(\tilde{Z}^L \geq \ell|\tilde{B}_\rho) = \lim_{\rho\to 0}\frac{\mu(\tilde{B}_\rho \cap T^{-(\ell-1)p}\tilde{B}_\rho)}{\mu(\tilde{B}_\rho)} = \left(\lim_{\rho\to 0}\frac{\mu(\tilde{B}_\rho \cap T^{-p}\tilde{B}_\rho)}{\mu(\tilde{B}_\rho)}\right)^{\ell-1}$$

which implies that $\hat{\alpha}_\ell = \hat{\alpha}_2^{\ell-1}$, where $\tilde{Z}^L = \sum_{j=1}^{L}\mathbf{1}_{\tilde{B}_\rho(x)}\circ T^j$. Then, for $\alpha_\ell = \hat{\alpha}_\ell - \hat{\alpha}_{\ell+1}$, we thus obtain by [21] that $\lambda_\ell = \frac{\alpha_\ell - \alpha_{\ell+1}}{\alpha_1} = (1-\theta)\Theta^{\ell-1}$, where $1-\Theta = \alpha_1 = 1-\hat{\alpha}_2$ is the extremal index. Hence, $N^{\tau}_{B_n}$ converges in distribution to a Pólya-Aeppli distributed random variable. □

It is worth mentioning that the previous result gives also an alternative way to prove EVT for the baker's map, which is recovered when $k = 0$, as the limiting distribution of $\mu(N^{\tau}_{B_n} = 0)$. Let us state it explicitly.

Proposition 7.3. *Let T be the baker's transformation and consider the function $M_n(x) := \max\{\phi(x), \ldots, \phi(T^{n-1}x)\}$, where $\phi(x) = -\log d_\infty(x,z)$, and z is chosen μ-almost everywhere with respect to the SRB measure μ. Then, if z is a periodic point of minimal period p verifying (52), we have*

$$\lim_{n\to\infty}\mu(M_n \leq u_n) = e^{-\theta\tau},$$

where $n\mu(B(z, e^{-u_n})) \to \tau$ and

$$\theta = 1 - \alpha^p.$$

Remark 7.4. Propositions 6.1 and 7.3 show that for a typical (non-periodic) point z, the limiting distribution of the maximum is purely exponential. The baker's map is probably the easiest example of a singular attractor. It is annoying that we could not compute analytically the extremal index with respect to the Euclidean metric, which is the metric usually accessible in simulations and physical observations. With references to Figures 2 and 3 respectively, the area of an extremely thin green ellipse within the blue circle is asymptotically equivalent to the area of an extremely thin green rectangle within the blue square, so, taking into account the ratio between the areas of the blue circle and the blue square, the limit as $p \to \infty$ of the extremal index for the Euclidean holes can be calculated.[16]

Example 7.5. The second statement of Proposition 7.2 about periodic points requires the neighborhoods B_n to be chosen in a dynamically relevant way. Here, they turn out to be squares (or rectangles). If the measure has some mixing properties with respect to a partition, then the sets B_n can be taken to be cylinder sets as it was done in [20] for periodic points, and in [19] Corollary 1 for non-periodic points. Here, we show that for Euclidean balls, one cannot in general expect the limiting distribution at periodic points to be Pólya-Aeppli, and therefore cannot be described by the single value of the extremal index.

We assume that all parameters are equal, that is $\gamma_a = \gamma_b = \alpha = \beta = \frac{1}{2}$. This is the *fat* baker's map for which the Lebesgue measure on $[0,1]^2$ is the SRB measure μ.

[16] We thank the anonymous referee for this observation.

Let x be a periodic point with minimal period p. Then, $\mu(B(x,r)) = r^2\pi$ and

$$\mu\left(\bigcap_{i=0}^{k} T^{-ip}B(x,r)\right) = 4r^2 2^{-kp}(1+\mathcal{O}(2^{-2kp})).$$

This yields

$$\hat{\alpha}_{k+1} = \lim_{r\to 0}\frac{\mu\left(\bigcap_{i=0}^{k} T^{-ip}B(x,r)\right)}{\mu(B(x,r))} = \frac{4}{\pi}\arctan 2^{-kp} = \frac{4}{\pi}2^{-kp}(1+\mathcal{O}(2^{-2kp}))$$

for $k = 1,2,\ldots$. According to [21] Theorem 2, we then define the values $\alpha_k = \hat{\alpha}_k - \hat{\alpha}_{k+1}$ where the value α_1 is the extremal index, i.e. $\theta = \alpha_1$. If the limiting distribution is Pólya-Aeppli, then the probabilities $\lambda_k = \frac{\alpha_k - \alpha_{k+1}}{\alpha_1}$, $k = 1,2,\ldots$, are geometrically distributed and must satisfy $\lambda_k = \theta(1-\theta)^{k-1}$, which is equivalent to saying that $\hat{\alpha}_{k+1} = (1-\theta)^k$ for $k = 0,1,2,\ldots$ (see [21] Theorem 2). Evidently, this condition is violated in the present case, and we conclude that the limiting distribution given by the values $\hat{\alpha}_k$ is not Pólya-Aeppli and in fact obeys another compound Poisson distribution.

7.3. **Compound point processes.** The compound Poisson distribution could be enriched by defining the *rare event point process* (REPP). Let us first introduce a few objects. Take $I_l = [a_l, b_l), l = 1,\ldots,k, a_l, b_l \in \mathbb{R}_0^+$, a finite number of disjoint semi-open intervals of the non-negative real axis, and call $J = \cup_{l=1}^k I_l$ their union. If r is a positive real number, we write $rJ = \cup_{l=1}^k rI_l = \cup_{l=1}^k [ra_l, rb_l)$. We denote with $|I_l|$ the length of the interval I_l, which coincides with its Lebesgue measure $\mathrm{Leb}(I_l)$. The REPP counts the number of visits to the set B_n during the rescaled time period v_nJ

$$N_n(\cdot)(J) = \sum_{l\in v_nJ\cap\mathbb{N}_0} \mathbf{1}_{B_n}(T^l\cdot), \tag{63}$$

where v_n is taken as

$$v_n = \left\lfloor \frac{\tau}{\mu(B_n)}\right\rfloor,\ \tau > 0.$$

Our REPP belongs to the class of the point processes on $\mathbb{R}_0^+$; see [22] for all the properties of point processes used below. They are given by any measurable map $N : (X, \mathcal{F}_X, \mu) \to \mathcal{N}_p([0,\infty))$, where $(X, \mathcal{F}_X, \mu)$ is the probability space of our original dynamical system with the invariant measure μ and the Borel σ-algebra $\mathcal{F}_X$, and $\mathcal{N}_p([0,\infty))$ denotes the set of counting measures $\mathfrak{c}$ on $\mathbb{R}_0^+$ endowed with the σ-algebra $\mathcal{M}_p(\mathbb{R}_0^+)$, which is the smallest σ-algebra making all evaluation maps $\mathfrak{c} \to \mathfrak{c}(B)$, from $\mathcal{N}_p([0,\infty)) \to [0,\infty]$ measurable for all $B \in \mathbb{R}_0^+$. Any counting measure $\mathfrak{c}$ has the form $\mathfrak{c} = \sum_{i=1}^\infty \delta_{x_i}$, $x_i \in [0,\infty)$. The distribution of N, denoted μ_N, is the measure $(\mu\circ N^{-1}) = (\mu[N\in\cdot])$ on $\mathcal{M}_p(\mathbb{R}_0^+)$. The set $\mathcal{N}_p([0,\infty))$ becomes a topological space with the vague topology, i.e. the sequence $\mathfrak{c}_n$ converges to $\mathfrak{c}$ whenever $\mathfrak{c}_n(\phi) \to \mathfrak{c}(\phi)$ for any continuous function $\phi : \mathbb{R}_0^+ \to \mathbb{R}$ with compact support. We also say that the sequence of point processes N_n converges in distribution to the point process N, eventually defined on another probability space $(X', \mathcal{F}'_{X'}, \mu')$, if μ_{N_n} converges weakly to μ'_N, that is, for every continuous function φ defined on $\mathcal{N}_p([0,\infty))$, we have $\lim_{n\to\infty}\int \varphi\, d\mu\circ N_n^{-1} = \int \varphi\, d\mu'\circ N^{-1}$. In this case, we will write $N_n \xrightarrow{\mu} N$.

If we now return to our REPP (63), we will see that a very common result is to get $N_n \xrightarrow{\mu} \tilde{N}$, where

$$\mu(x, \tilde{N}(x)(I_l) = k_l, 1 \le l \le n) = \prod_{l=1}^{n} e^{-\tau \mathrm{Leb}(I_l)} \frac{\tau^{k_l} \mathrm{Leb}(I_l)^{k_l}}{k_l!}, \tag{64}$$

for any disjoint bounded sets $I_1, \ldots, I_n$ and non-negative integers $k_1, \ldots, k_n$, which is called the *standard Poisson point process*. In general, our REPP processes converge in distribution to a *compound point process* (CPP). We say that the point process $N : (X', \mathcal{F}'_{X'}, \mu') \to \mathcal{N}_p([0, \infty))$ is a CPP with intensity parameter t, and cluster size distribution $(\lambda_l)_{l \ge 1}$ if it satisfies:

- For any finite sequence of measurable sets $B_1, \ldots, B_k$ in $\mathcal{F}'_{X'}$ and mutually disjoint, the random variables $N(\cdot)(B_i), i = 1, \ldots, k$, are independent.
- For any measurable set $B \in \mathcal{F}'_{X'}$, the random variable $N(\cdot)(B)$ is a CP random variable with intensity $t\mathrm{Leb}(B), t > 0$ and cluster size distribution $(\rho_l)_{l \ge 1}$, see the definition in Section 7.

From now on we will simply write $N(\cdot)$ instead of $N(x)(\cdot)$, and we consider it as a CPP. In order to study the convergence of our REPP N_n to the CPP N, two equivalent criteria are available. Before stating them, we should recall the definition of the Laplace transform for a general point process $R : (X', \mathcal{F}'_{M'}, \mu') \to \mathcal{N}_p([0, \infty))$

$$\psi_R(y_1, \ldots, y_k) = \mathbb{E}_{\mu'}\left(e^{-\sum_{l=1}^{k} y_l R(I_l)}\right), \tag{65}$$

for all non-negative values $y_1, \ldots, y_k$, each choice of k disjoint intervals $I_i = [a_i, b_i)$, $i = 1, \ldots, k$. In the case of a CPP N with intensity parameter t and cluster size distribution $(\rho_l)_{l \ge 1}$, we get

$$\psi_N(y_1, \ldots, y_k) = e^{-t \sum_{l=1}^{k} (1 - \varphi(y_l)) \mathrm{Leb}(I_l)}, \tag{66}$$

where $\varphi(y) = \sum_{i=0}^{\infty} e^{-yi} \rho_i$ is the Laplace transform of the cluster size distribution $(\rho_l)_{l \ge 1}$.

Therefore in order to establish the convergence in distribution of the REPP N_n toward the CPP N, it will be sufficient to show [22]:
- (C1): that for any k disjoint intervals $I_i = [a_i, b_i), i = 1, \ldots, k$, the joint distribution of N_n converges to the joint distribution of N, namely

$$(N_n(I_1), \ldots N_n(I_k)) \to (N(I_1), \ldots N(I_k)).$$

-(C2): the convergence of the Laplace transforms

$$\psi_{N_n}(y_1, \ldots, y_\zeta) = \mathbb{E}\left(e^{-\sum_{l=1}^{k} y_l N_n(I_l)}\right) \to \psi_N(y_1, \ldots, y_k) = e^{-t \sum_{l=1}^{k} (1 - \varphi(y_l)) \mathrm{Leb}(I_l)},$$

as $n \to \infty$.

The criterion (C1) lends itself to being studied with the probabilistic approach of [21] as two of us recently showed in ([1], Theorem 3); see also [15] for a different method. The criterion (C2) is *naturally* adapted to the spectral approach (just replacing characteristic functions with Laplace transforms), and the complete treatment, involving two of us, will appear soon [4]. Both criteria allow to extend immediately Proposition 7.2 to the point process framework, giving the following propostion.

Proposition 7.6. *Consider the counting measure*

$$N_n(\cdot)(J) = \sum_{l \in v_n J \cap \mathbb{N}_0} \mathbf{1}_{B_n}(T^l \cdot)$$

where τ is a positive parameter, $v_n = \left\lfloor \frac{\tau}{\mu(B_n)} \right\rfloor$, and z is a point for which the limit (52) exists and $n\mu(B(z, e^{-u_n})) \to \tau$.

- *If z is not a periodic point, using the Euclidean metric, then N_n converges in distribution to a standard Poisson point process of intensity τ; see (64) for the finite size distributions.*
- *If z is a periodic point of minimal period p, using the l^∞ metric, we get a compound point process of Pólya-Aeppli type, namely a CPP with intensity $\tau\theta$ and cluster size distribution $\theta(1-\theta)^l, l \geq 1$, where θ is given as above by $\theta = 1 - \lim_{n\to\infty} \frac{\mu(T^{-p}B_n \cap B_n)}{\mu(B_n)}$.*

8. **Generalization to other observable and connection with hitting time.** One could possibly wonder if the observable (13), $\Xi(x) = -\log d(x,z)$, plays an essential role in the theory. The answer is more nuanced. Let us consider a measurable function $\Phi : X \to \mathbb{R} \cup \pm\{\infty\}$ and construct the new process $\Phi \circ T^j, j \geq 0$. We are interested in the extreme value distribution (EVD):

$$\mathcal{W}_n = \mu(\mathcal{M}_n \leq \mathfrak{u}_n),$$

where

$$\mathcal{M}_n(x) := \max_{0 \leq j \leq n-1} \{\Phi(T^j x)\}.$$

We will return in a moment on the choice for sequence $\mathfrak{u}_n$. Let us introduce the set

$$\mathfrak{B}_n := \{\Phi \geq \mathfrak{u}_n\},$$

which we continue to call a *ball*. For instance, another commonly used observable is

$$\Phi(x) = x^{-\frac{1}{\alpha}}, \ \alpha > 0;$$

in this case, $\mathfrak{B}_n$ is simply a closed euclidean ball around zero of radius $\mathfrak{u}_n^{-\alpha}$; see Appendix A for a brief account of EVD for different types of observables.

Define now the quantity for any $x \in X$:

$$\mathfrak{t}_{\mathfrak{B}_n}(x) := \inf\{j \geq 1; \Phi(T^j x) \in \mathfrak{B}_n\},$$

which gives the first *hitting time* to the ball $\mathfrak{B}_n$ when we start from the point x. By the invariance of the SRB measure μ, it is easy to show that

$$\mu(\mathfrak{t}_{\mathfrak{B}_n} > n) = \mathcal{W}_n = \mu(\mathcal{M}_n \leq \mathfrak{u}_n), \tag{67}$$

which establishes an important link between the law of extremes and the distribution of the hitting times (this also enlightens again why the EVD is recovered when $k = 0$ in the Poisson distribution, see Proposition 7.3).

Given the general observable Φ, the spectral analysis of Section 3 proceeds formally as we did in the previous chapters, but in order to get the final results, we need to check the following points:
(i) $\mathbf{1}_{\mathfrak{B}_n} \in \mathcal{B}$, and we saw that the geometrical shape of $\mathfrak{B}_n$ matters.
(ii) Condition **A2** must be checked taking into account again the geometrical nature of $\mathfrak{B}_n$.
(iii) In dealing with condition, **A4** we now have to compare $\sup_{\|h\|\leq 1} |Z(\mathcal{L}(h\mathbf{1}_{\mathfrak{B}_n}))|$,

see equation (36), with $\Delta_n = \mu(\mathfrak{B}_n)$. In doing that with the help of Lemma 5.1, we use again the local structure of the set $\mathfrak{B}_n$.
(iv) Finally, we have to prove the convergence of q_k, (55), to get the extremal index. This last condition requires the important scaling

$$n\mu(\mathfrak{B}_n) \to \tau,$$

for some positive τ, which fixes as well the choice of $\mathfrak{u}_n$.

Appendix A. **Observables and corresponding extreme value laws.** The main classical result of extreme value theory is given in the next theorem due to Gnedenko [17] and Fisher and Tippett [16]. The theorem deals with a sequence of i.i.d. random variables, and we denote again with M_n the maximum over the first n variables.

Theorem A.1. *If $X_0, X_1, \dots$ is a sequence of i.i.d. random variables and there exists linear normalizing sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$, with $a_n > 0$ for all n, such that*

$$\mathbb{P}(a_n(M_n - b_n) \le y) \to G(y),$$

where the convergence occurs at continuity points of G, and G is nondegenerate, then $G(y) = e^{-\tau(y)}$, where $\tau(y)$ is one of the following three types (for some $\beta, \gamma > 0$):
(1) $\tau_1(y) = e^{-y}, y \in \mathbb{R}$;
(2) $\tau_2(y) = y^{-\beta}, y > 0$;
(3) $\tau_3(y) = (y)^{\gamma}, y > 0$.

We now give conditions on the choice of the observable to get sufficient and necessary conditions in order to get a nondegenerate EVD still in the i.i.d. setting. For the reader's convenience, we quote Section 4.2.1 of the book [26], for the choice of the function Φ introduced in Section 8. It has the form, for $x \in X$,

$$\Phi(x) = g(\mathrm{dist}(x,\zeta)),$$

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

- Type g_1: There exists some strictly positive function $h : W \to \mathbb{R}$ such that, for all $y \in \mathbb{R}$,
$$\lim_{s\to g_1(0)} \frac{g_1^{-1}(s + yh(s))}{g_1^{-1}(s)} = e^{-y}.$$
- Type g_2: $g_2(0) = +\infty$, and there exists $\beta > 0$ such that, for all $y > 0$,
$$\lim_{s\to+\infty} \frac{g_2^{-1}(sy)}{g_2^{-1}(s)} = y^{-\beta}.$$
- Type g_3: $g_3(0) = D < +\infty$, and there exists $\gamma > 0$ such that, for all $y > 0$,
$$\lim_{s\to+\infty} \frac{g_3^{-1}(D - sy)}{g_3^{-1}(D - s)} = y^{\gamma}.$$

Examples of each one of the three types are as follows:

1. $g_1(x) = -\log x$, in this case $h = 1$.
2. $g_2(x) = x^{-1/\alpha}$ for some $\alpha > 0$, in this case $\beta = \alpha$.
3. $g_3(x) = D - x^{1/\alpha}$ for some $D \in \mathbb{R}$ and $\alpha > 0$, in this case $\gamma = \alpha$.

Type 1 gives the *Gumbel* law, type 2 gives the *Fréchet* law, and finally type 3 furnishes the *Weibull* law.

A great amount of work has been done to extend such a result from the i.i.d. setting first to stationary, and then to non-stationary processes. Whenever the latter arise in the dynamical systems setting, we refer to the book [26] for an exhaustive presentation of the results: The spectral approach used in this article is one of them. As a final remark, we notice that, by expressing our scaling sequence u_n as $u_n = b_n + \frac{y}{a_n}$, we will recover one of the previous three distributions as a function of the parameter y.

Acknowledgments. S.V. thanks the Laboratoire International Associé LIA LYSM, the INdAM (Italy), the UMI-CNRS 3483, Laboratoire Fibonacci (Pisa) where this work has been completed under a CNRS delegation and the *Centro de Giorgi* in Pisa for various supports. SV thanks M. Demers, D. Dragičević C. Gonzalez-Tokman and Y. Nakano for enlightening discussions.

N.H. was supported by a Simons Foundation Collaboration Grant (ID 526571).

J.A. was supported by an ARC Discovery Project and thanks the Centro de Giorgi in Pisa and CIRM in Luminy for their support and hospitality.

We finally thank the anonymous referees whose very careful and accurate reading of the paper, helped us to considerably improve it by adding several new proofs and explications.

REFERENCES

[1] L. Amorim, N. Haydn and S. Vaienti, Compound Poisson distributions for random dynamical systems using probabilistic approximations, *Stochastic Processes and Their Applications*, **179** (2025), 104511.

[2] J. Atnip, G. Froyland, C. Gonzalez-Tokman and S. Vaienti, Thermodynamic formalism and perturbation formulae for quenched random open dynamical systems, to appear in Dissertationes Mathematicae, `https://arxiv.org/pdf/2307.00774.pdf`.

[3] J. Atnip, G. Froyland, C. Gonzalez-Tokman and S. Vaienti, Compound Poisson statistics for dynamical systems via spectral perturbation, submitted, `https://arxiv.org/pdf/2308.10798.pdf`.

[4] J. Atnip, M. Nicol and F. Pène S. Vaienti, in preparation

[5] H. Aytac, J. Freitas and S. Vaienti, Laws of rare events for deterministic and random dynamical systems, *Trans. Amer Math. Soc.*, **367** (2015), 8229-8278, `http://arxiv.org/pdf/1207.5188`.

[6] Th. Caby, D. Faranda, G. Mantica, S. Vaienti and P. Yiou, Generalized dimensions, large deviations and the distribution of rare events, *Physica D*, **40015** (2019), 132143, 15 pp.

[7] Th. Caby, D. Faranda, S. Vaienti and P. Yiou, On the computation of the extremal index for time series, *Journal of Statistical Physics*, **179** (2020), 1666-1697.

[8] M. Carvalho, A. C. M. Freitas, J. M. Freitas, M. Holland and M. Nicol, Extremal dichotomy for uniformly hyperbolic systems, *Dyn. Syst.*, **30** (2015), 383-403,

[9] M. F. Demers, Escape rates and physical measures for the infinite horizon Lorentz gas with holes, *Dynamical Systems: An International Journal*, **28** (2013), 393-422.

[10] M. Demers, A gentle introduction to anisotropic banach spaces, *Chaos, Solitons and Fractals*, **116** (2018), 29-42.

[11] M. F. Demers and C. Liverani, Stability of statistical properties in two-dimensional piecewise hyperbolic maps, *Trans. Amer. Math. Soc.*, **360** (2008), 4777-4814.

[12] M. Demers and H. Zhang, Spectral analysis of the transfer operator for the Lorentz gas, *Journal of Modern Dynamical Systems*, **5** (2011), 665-709.

[13] D. Faranda, H, Ghoudi, P. Guiraud and S. Vaienti, Extreme value theory for synchronization of coupled map lattices, *Nonlinearity*, **31** (2018), 3326-3358.

[14] W. Feller, *An Introduction to Probability Theory and its Applications. Vol. I*, John Wiley and Sons, Inc., New York-London-Sydney, third edition, 1968.

[15] A. C. M. Freitas, J. M. Freitas and M. Todd, The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics, *Comm. Math. Phys.*, **321** (2013), 483-527.

[16] A. Fisher and L. Tippett, Limiting foms of the frequency distribution of the largest or smallest member of a sample, *Mathematical Proceedings of the Cambridge Philosophical Society*, **24** (1928), 180-190.

[17] B. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Mathematics*, **2** (1943), 423-453.

[18] P. Giulietti, P. Koltai and S. Vaienti, Targets and holes, *Proc. Amer. Math. Soc.*, **149** (2021), 3293-3306.

[19] N. Haydn and Y Psiloyenis, Return times distribution for Markov towers with decay of correlations, *Nonlinearity*, **27** (2014), 1323-1349.

[20] N. Haydn and S. Vaienti, The compound Poisson distribution and return times in dynamical systems, *Prob. Th. & Related Fields*, **144** (2009), 517-542.

[21] N. Haydn and S. Vaienti, Limiting entry and return times distributions fir arbitrary null sets, *Communication Mathematical Physics*, **378** (2020), 149-184.

[22] O. Kallenberg, *Random Measures*, Akademia-Verlag Berlin, 1986.

[23] G. Keller, Rare events, exponential hitting times and extremal indices via spectral perturbation, *Dyn. Syst.*, **27** (2012), 11-27.

[24] G. Keller and C. Liverani, Rare events, escape rates and quasistationarity: Some exact formulae, *J. Stat. Phys.*, **135** (2009), 519-534.

[25] G. Keller and C. Liverani, Stability of the spectrum for transfer operators, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, **1999**, 141-152.

[26] V. Lucarini, D. Faranda, A. C. M. Freitas, J. M. Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd and S. Vaienti, *Extremes and Recurrence in Dynamical Systems*, Pure Appl. Math. (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2016.

[27] E. Ott, *Chaos in Dynamical Systems*, 2nd ed. Cambridge University Press, 2002.

[28] M. Viana, *Stochastic dynamics of deterministic systems*, Brazillian Math. Colloquium 1997, IMPA.

Received February 2024; 1st revision January 2025; 2nd revision February 2025; early access March 2025.