HAL
open science

Extreme value theory for synchronization of coupled map lattices

D. Faranda, H. Ghoudi, P. Guiraud, Sandro Vaienti

To cite this version:

D. Faranda, H. Ghoudi, P. Guiraud, Sandro Vaienti. Extreme value theory for synchronization of coupled map lattices. Nonlinearity, 2018, 31 (7), pp.3326-3358. 10.1088/1361-6544/aabc8e . hal-01768182

HAL Id: hal-01768182

https://hal.science/hal-01768182v1

Submitted on 22 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1708.00191v2 [math.DS] 7 Apr 2018

EXTREME VALUE THEORY FOR SYNCHRONIZATION OF COUPLED MAP LATTICES

Abstract. We show that the probability of the appearance of synchronization in chaotic coupled map lattices is related to the distribution of the maximum of a certain observable evaluated along almost all orbits. We show that such a distribution belongs to the family of extreme value laws, whose parameters, namely the extremal index, allow us to get a detailed description of the probability of synchronization. Theoretical results are supported by robust numerical computations that allow us to go beyond the theoretical framework provided and are potentially applicable to physically relevant systems.

D. Faranda [1] [2], H. Ghoudi[3], P. Guiraud[4], S. Vaienti[5]

1. Introduction

Coupled Map Lattices (CML) are discrete time and space dynamical systems introduced in the mid 1980's by Kaneko and Kapral as suitable models for the study and the numerical simulation of nonlinear phenomena in spatially extended systems. The phase space of a CML is a set of scalar (or vector) sequences indexed by a lattice L, e.g. $L = \mathbb{Z}^d$, $\mathbb{Z}$ or $\mathbb{Z}/n\mathbb{Z}$. For instance, a configuration $\overline{x} \in I^L$ of the lattice may represent a spacial sample of a mesoscopic quantity with value in an interval I, such as a chemical concentration, the velocity of a fluid, a population density or a magnetization. The dynamics of the lattice is given by a map $\hat{T} : I^L \to I^L$ which is usually written as the composition of two maps, i.e. $\hat{T} := \Phi_\gamma \circ \hat{T}_0$, where $\hat{T}_0 : I^L \to I^L$ is called the *uncoupled dynamics* and $\Phi_\gamma : I^L \to I^L$ the *coupling operator*. The uncoupled dynamics acts on a configuration $x \in I^L$ as the product dynamics of a *local map* $T : I \to I$, that is $\hat{T}_0(x)_i := T(x_i)$ for every $i \in L$. The coupling operator models spacial interactions, which intensity is given by the parameter $\gamma \in [0, 1]$. In particular, in the absence of interaction

Date: April 10, 2018.

[1]LSCE-IPSL, CEA Saclay l'Orme des Merisiers, CNRS UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France.

[2]London Mathematical Laboratory, London, UK . E-mail: `davide.faranda@lsce.ipsl.fr`.

[3] Laboratory of Dynamical Systems and Combinatorics, Sfax University Tunisia and Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France and Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde. Email: `ghoudi.hamza@gmail.com` .

[4]CIMFAV, Facultad de Ingeniera, Universidad de Valparaso, Valparaìso, Chile. E-mail: `pierre.guiraud@uv.cl`.

[5]Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France. E-mail: `vaienti@cpt.univ-mrs.fr`.

$\gamma = 0$ and $\Phi_0 = Id$. For example, for $L = \mathbb{Z}$ or $L = \mathbb{Z}/n\mathbb{Z}$ the coupling operator often writes as

$$(\Phi_\gamma(x))_i := \sum_{j \in L} c_{\gamma,j} x_{i-j} \quad \forall i \in L, \tag{1.1}$$

where $c_{\gamma,j} \geq 0$, $\sum_{j \in L} c_{\gamma,j} = 1$ and $c_{0,0} = 1$.

In the huge literature about CML, one can find many possible choices for the local map and the coupling operator. For instance, the dynamics of CML of bistable, unimodal, or chaotic maps have been studied for different kind and range of coupling, revealing a rich phenomenology including spatial chaos, stable periodic points, space-time chaos, clusters, traveling waves and synchronization, (see [6, 4, 10] and references therein). In this paper, we will consider a system of n coupled chaotic local maps (the precise properties are given in Section 2) defined for any $\overline{x} := (x_1, \ldots, x_n) \in I^n$ by:

$$(\hat{T}(\overline{x}))_i = (1-\gamma)T(x_i) + \frac{\gamma}{n}\sum_{j=1}^{n} T(x_j) \qquad \forall i \in \{1, \ldots, n\}.$$

Note that the study of this system is equivalent to that of a CML on a periodic lattice where the coupling operator is defined by (1.1) with $L = \mathbb{Z}/n\mathbb{Z}$, $c_{\gamma,0} = (1 - \gamma + \frac{\gamma}{n})$ and $c_{\gamma,j} = \frac{\gamma}{n}$ for all $j \in \{1, \ldots, n-1\}$. The chaotic and synchronization properties of CML of logistic local maps with this mean-field-type global coupling were observed and studied by Kaneko in [23] and then, among others, by P. Ashwin [2] (and references therein).

The first contribution which looked at CML in the framework and with the tools of ergodic theory, was the work by Bunimovich and Sinai. In the famous paper [3], using thermodynamic formalism, they proved the existence of mixing SRB measures for infinite CML with chaotic local map and weak (nearest neighbor) coupling. Since then, the progress in the study of the statistical properties of chaotic CML has been enormous, with the contribution of several people, and the development of a spectral theory [15, 16]. We defer to the book [4] for a wide panorama on the different approaches to CML and for exhaustive references.

In this paper, we present a new application of Extreme Value Theory (EVT) to CML on a finite (or periodic) lattice. Our aim is to provide a first approach to CML by using EVT and to show how to get a certain number of rigorous results about the statistics of some rare events, such as the synchronization in chaotic CML. We say that the CML is synchronized when it is near a homogeneous configuration (in a small neighborhood of the diagonal of the phase space). Synchronization is usually intended to last for a while once it has started and this is what usually happens for some kinds of chains of synchronized oscillators. This is not the case of course for chaotic CML, since almost every orbit is recurrent by the Poincaré Theorem. What we actually investigate is therefore the *probability of a first synchronization and how long we should wait to get it with a prescribed accuracy.* EVT provides this kind of quantitative information, since synchronization

processes can be interpreted and quantified by computing the asymptotic distribution of the maximum of a suitable random process, see Sections 3 and 4.

Although we could not get a global synchronization persisting in time, we could ask about the distribution of the number of successive synchronization events when the systems evolves up to a certain time. We will see that after a suitable rescaling, the distribution of that number follows a compound Poisson statistics: it is worth mentioning that for two *uncoupled* expanding maps of the circle, this result dates back to a paper by Coelho and Collet, [5].

Actually a first result in our direction was given in the paper [13], although not explicitly related to EVT, where the authors considered two coupled interval maps and applied their spectral theory of open systems with holes to investigate the first entrance of the two components into a small strip along the diagonal, which is equivalent to the synchronization of the two-components lattice up to a certain accuracy. In more general situations, we will present arguments about the spectral properties of the transfer operator of the system to sustain the existence of a limit distribution for the maxima of some observables related to synchronization, and we will discuss a formula approximating the *extremal index* (a parameter of the distribution) for lattices with an arbitrary number of components. We therefore estimate the behavior of such an index when the number of components is large. We will then generalize the theory to CML which are randomly perturbed with additive noise and show, in particular with numerical evidence, that the extremal index is 1 for any dimension of the lattice. We hope that our approach could be helpful to understand and quantify those phenomena, like in neuronal spikes or in business cycles of financial markets, where bursts of synchronization happen, disappear, happen again, apparently in a disordered manner, but very often following the extreme distributions arising in chaotic systems.

In Section 2, we present a powerful and general approach based on perturbation of the transfer operator, and which has the advantage of being applicable to a large class of observables arising in the study of EVT. In Section 3, we give a short insight into basic notions of EVT, especially when it is applied to recurrence in dynamical systems. In particular, we define the extremal index and show that it goes to one when the size of the lattice goes to infinity or in presence of noise. In Section 4, we apply EVT to compute the probability of synchronization events, and sustain the results by computing the extremal index in Section 5. This computation depends on the behavior of the invariant density in the neighborhood of the diagonal; our formula (5.35) can be proved under the assumption **P8** which we believe to be unavoidable. In Section 6, we study the distribution of the number of successive synchronization events. In Section 7, we show that our

analytic results and estimates are supported by numerical computations. They confirm the existence of an extreme value distribution for a different kind of synchronization, which we called *local*, and they validate the expected compound Poisson statistics for the distribution of the number of successive visits. The fact that the extremal index for local synchronization seems not to depend on the size of the lattice is an interesting numerical discovery. In forthcoming papers we will study more general CML with non-local form of coupling including the important case of diffusive or Laplacian interaction. A few other possible developments are presented at the end of the paper (see section 7.2).

2. The map and the operators

As mentioned in the Introduction, we consider a finite CML of size $n \geq 2$ with a local map $T : I \to I$ and a global coupling. It is defined for any $\overline{x} = (x_1, \ldots, x_n) \in I^n$ and $\gamma \in [0, 1]$ by [6]:

$$\hat{T}(\overline{x})_i = (1-\gamma)T(x_i) + \frac{\gamma}{n}\sum_{j=1}^{n} T(x_j) \qquad \forall\, i \in \{1, 2, \ldots, n\}, \tag{2.2}$$

where $\overline{x} = (x_1, \ldots, x_n) \in I^n, \gamma \in [0, 1]$. We suppose that T is a piece-wise expanding map of the unit interval onto itself, with a finite number of branches, say q, and which we take of class C^2 on the interiors of the domains of injectivity $A_1, \ldots, A_q$, and extended by continuity to the boundaries. The C^2 assumption is used in the proof of Propositions (5.5) and (5.6), although it could be relaxed with a $C^{1+\alpha}$ condition. Instead the finitness of the number of branches is widely used in almost all the arguments. Let us denote by $U_k, k = 1, \ldots, q^n$, the domains of local injectivity of $\hat{T}$. By the previous assumptions on T, there exist open sets $W_k \supset U_k$ such that $\hat{T}_{|W_k}$ is a C^2 diffeomorphism (on the image). We will require that

$$s_n := \sup_k \sup_{\overline{x} \in \hat{T}(W_k)} ||D\hat{T}^{-1}_{|W_k}(\overline{x})|| < \lambda < 1,$$

where $\lambda := \sup_i \sup_{x \in T(A_i)} |DT^{-1}_{|A_i}(x)|$, and $||\cdot||$ stands for the euclidean norm. We will write *dist* for the distance with respect to this norm.

An important tool for our further considerations is the transfer, or Perron-Frobenius (PF), operator. The PF operator $\hat{P}$ of the map $\hat{T}$ is simply defined by the duality integral relation

$$\int \hat{P}(f) g d\mathrm{Leb} = \int f g \circ \hat{T} d\mathrm{Leb},$$

where Leb denotes the Lebesgue measure on I^n, $f \in L^1$ and $g \in L^\infty$.[7] The spectral properties of the PF operator become interesting when it acts on suitable Banach spaces.

[6] We will not index the map $\hat{T}$ with n, hoping it will be clear from the context.

[7] In the following we will use the same symbol Leb for any n. Moreover L^1, L^p and L^∞ will be taken with respect to Leb. Finally the integral with respect to Lebesgue measure will be denoted with $\int d\mathrm{Leb}(x)$ or $\int dx$.

Let us therefore suppose that there exists a Banach space $\mathcal{B}$ with norm $||\cdot||_{\mathcal{B}}$, which is compactly injected in L^1 and the following properties hold[8]:

- **P1** (Lasota-Yorke inequality) For any $f \in \mathcal{B}$ there exists $\eta < 1$ and $C > 0$ such that
$$||\hat{P}f||_{\mathcal{B}} \leq \eta||f||_{\mathcal{B}} + C||f||_1.$$

The Lasota-Yorke inequality implies that $\hat{P}$ has an isolated eigenvalue equal to 1 which is also the spectral radius of $\hat{P}$ (*spectral gap property*). We will often call η the contraction factor in the Lasota-Yorke inequality.

Remark 2.1. *By iterating the previous inequality one easily get that*
$$||\hat{P}^k f||_{\mathcal{B}} \leq \eta^k||f||_{\mathcal{B}} + \frac{C}{1-\eta}||f||_1, \quad \forall\, k > 1. \tag{2.3}$$
This last inequality is actually needed in the perturbation theory used below. If one cannot achieve it because **P1** *fails, it is enough to get* **P1** *for an iterate of $\hat{T}$. In this case a standard argument allows us to get again (2.3).*

- **P2** The eigenvalue 1 is simple and $\hat{P}$ has no other eigenvalue on the unit circle. This implies that $\hat{P}$ preserves a mixing measure $\hat{\mu}$ which is the unique absolutely continuous invariant measure with respect to Lebesgue. We moreover assume that the associated density $\hat{h} \in L^\infty$.

It is well known that with our assumptions on T, the uncoupled dynamics $\hat{T}_0$, i.e $\gamma = 0$ in (2.2), satisfies **P1** on any *reasonable* functional space $\mathcal{B}$. We will give examples of such spaces just below. Therefore, the spectral decomposition theorem of Ionescu-Tulcea-Marinescu, see for instance [21], guarantees the existence of a finite number of absolutely continuous ergodic components. They reduce to a unique absolutely continuous mixing measure, which is **P2**, with some topological transitivity condition on the map T, which could be achieved by asking, for instance, T to be Bernoulli, Markov, covering, etc (see, e.g., Example 2.2).

In order to transfer the properties **P1** and **P2** to the map $\hat{T}$ with $\gamma > 0$, we invoke the perturbation theory by Keller and Liverani developed in [14]. According to that theory, one should previously show the persistence of the Lasota-Yorke inequality (2.3) for the map $\hat{T}$ and then check that, for any $f \in \mathcal{B}$, we have
$$||(\hat{P} - P_0)f||_1 \leq p_\gamma||f||_{\mathcal{B}}, \tag{2.4}$$
where P_0 is the PF operator of the uncoupled system ($\gamma = 0$), and p_γ is a monotone upper semi-continuous function converging to 0 when γ goes to 0. We defer again to Example 2.2 for a particular case, where this technique can be applied.

[8] We will call a Banach space with this property *adapted* (to L^1).

The aforementioned perturbation theory was successively improved in [13] by the same authors, in order to deal with *open* systems which produce a different kind of perturbation for the transfer operator. This perturbation arises naturally in the context of the EVT, as we will see in the next section. In order to introduce and define it, let $\{D_l\}_{l\in\mathbb{N}}$ be an increasing collection of nested subsets of I^n such that $\mathrm{Leb}(D_l) \to 1$ when $l \to \infty$. Moreover, suppose that the sets D_l are the closures of their interiors and have piece-wise C^∞ and co-dimension 1 boundaries. According to the observable used for the application of EVT, the sets D_l have a specific definition, and they will be given by (3.18) and (4.22). The EVT can be related to the spectral theory by considering the perturbed transfer operator $\widetilde{P}_l$, which is defined for any $h \in \mathcal{B}$ by:

$$\widetilde{P}_l(h) := \hat{P}(h\mathbf{1}_{D_l}).$$

We now add new assumptions this operator must satisfy in order to apply the perturbation theory for open systems. The goal is to compare the operators $\hat{P}$ and $\widetilde{P}_l$ and get an asymptotic expansion for the spectral radius of $\widetilde{P}_l$ close to 1 for large values of l. We will see that it will give us the extremal index in the limiting distribution of Gumbel's law. We follow in particular the scheme proposed by Keller in [12], that we also summarized in [1], Section 5, and in Chapter 7 of the book [25] to which we defer for more details. There are 6 assumptions in [12], Section 2. The first three ask for uniform (in the "noise" parameter l) quasi-compactness for the operator $\widetilde{P}_l$. We summarize them in the following single assumption:

- **P3** The operators $\widetilde{P}_l$ satisfy a Lasota-Yorke inequality, uniform in l, on the space $\mathcal{B}$, namely, the factors η and C are the same for every sufficiently large l.

The next two properties **P4** and **P5** cover assumptions (5) and (6) in Keller [12]. We also notice that **P4**, together with **P2**, implies assumption (4) in [12], as explained in Remark 3 still in [12].

- **P4** For any $h \in \mathcal{B}$, the quantity

$$r_l := \sup_{h,||h||_{\mathcal{B}}\leq 1} |\int (\hat{P}h - \widetilde{P}_l h) d\mathrm{Leb}|$$

  goes to zero when $l \to \infty$.
- **P5** The density $\hat{h}$ of the (unique mixing) invariant measure $\hat{\mu}$ of T verifies

$$r_l||(\hat{P} - \widetilde{P}_l)\hat{h}||_{\mathcal{B}} \leq C'\hat{\mu}(D_l^c), \tag{2.5}$$

  where C' is a constant independent of l and D_l^c denotes the complement of D_l. We moreover assume that the density $\hat{h}$ is strictly positive, namely its infimum is larger than $\hat{h}^{(\inf)} > 0$ on a set of full measure.

We finally assume that

- **P6** The following limit

$$q_k := \lim_{l\to\infty} q_{k,l} := \lim_{l\to\infty} \frac{\int (\hat{P} - \widetilde{P}_l)\widetilde{P}_l^k(\hat{P} - \widetilde{P}_l)(\hat{h})d\mathrm{Leb}}{\hat{\mu}(D_l^c)} \tag{2.6}$$

  exists for any $k \in \mathbb{N} \cup \{0\}$.

Under the assumptions **P1-P6**, it has been proved in [13] that

$$\theta := 1 - \sum_{k=0}^{\infty} q_k, \tag{2.7}$$

exists and is equal to $\lim_{l\to\infty} \frac{1-\rho_l}{\hat{\mu}(D_l^c)}$, where ρ_l is the spectral radius of $\widetilde{P}_l$. Therefore we have the following asymptotic expansion for ρ_l:

$$1 - \rho_l = \hat{\mu}(D_l^c)\theta(1 + o(1)), \text{ in the limit } l \to \infty. \tag{2.8}$$

We stress that ρ_l is the largest eigenvalue of $\widetilde{P}_l$, that there are no other eigenvalues on the circle of radius ρ_l, and that there exist functions $\hat{g}_l \in \mathcal{B}$ and measures $\hat{\mu}_l$ for which the operators $\widetilde{P}_l$ satisfy

$$\widetilde{P}_l h = \rho_l \hat{g}_l \int h d\hat{\mu}_l + Q_l h \tag{2.9}$$

for all $h \in \mathcal{B}$. Moreover $\int \hat{g}_l d\hat{\mu} = 1$, $\int h d\hat{\mu}_l \to \int h d\hat{\mu}$ when $l \to \infty$ and finally Q_l is a linear operator with spectral radius strictly less than ρ_l and satisfying: $||Q_l^n||_{\mathcal{B}} \leq \varsigma_l^n$, for a suitable $0 < \varsigma_l < 1$, see again [13] for the derivation of these formulas.

It is a remarkable fact that this approach automatically provides the scaling exponent θ for the asymptotic distribution of the maxima, see (4.25) below, and therefore it gives a new proof of the existence of that distribution. The quantity θ is called the extremal index (EI) and it will play an important role in the following. We will see in particular that it gives a correction to the pure exponential law for the distribution of the maxima. In that respect it coincides with the extremal index as it is defined in EVT, see [8], [25]. Our next task will therefore be to look for a Banach space which verifies the preceding six properties.

One natural candidate would be the *space $BV(I^n)$ of functions of bounded variation on $\mathbb{R}^n$ restricted to the L^1 functions supported on $\widetilde{I}^n :=$ interior(I^n)*. This space was used in [13] in dimension 2, but it seems difficult to use it in higher dimensions to obtain **P5**. The reason is that in order to get **P5** one needs first to compute the quantity r_l in **P4**. Since h may not be necessarily in L^∞, we should use Sobolev's inequality to estimate the integral and we get r_l of order $\mathrm{Leb}(D^c)^{\frac{1}{n}}$. This is not enough to recover **P5**, since the Banach norm $||(\hat{P} - \widetilde{P}_l)\hat{h}||_{\mathcal{B}}$ is simply bounded by a constant as a consequence of the Lasota-Yorke inequality. Instead for $n = 2$ the characterization of the total variation as the maximum of sectional variations along the coordinate axis is sufficient to get (**P5**), and it was just used in [13]. By referring to (2.12) below, we can in fact bound the integral

$\int |h\mathbf{1}_{D_l^c}|d\text{Leb}$ by 1/2 times the total variation of the density times the Lebesgue measure of the section of D_l^c along one of the two coordinate axis (we are using here the corollary 2.1 in [15]). But that sectional measure is of the same order of the Lebesgue measure of the whole D_l^c, just because we are on the unit square. We therefore turn our attention to another functional space, the quasi-Hölder space, whose importance for expanding dynamical systems was stressed in the seminal works by Keller [11] and Saussol [27].

We start by defining for all functions $h \in L^1(I^n)$ a semi-norm, which given two real numbers $\varepsilon_0 > 0$ and $0 < \alpha \leq 1$, writes

$$|h|_\alpha := \sup_{0<\varepsilon\leq\varepsilon_0} \frac{1}{\varepsilon^\alpha} \int \text{osc}(h, B_\varepsilon(\overline{x}))d\text{Leb},$$

where $\text{osc}(h, A) := \text{Esup}_{\overline{x}\in A} h(\overline{x}) - \text{Einf}_{\overline{x}\in A} h(\overline{x})$ for any measurable set A. We say that $h \in V_\alpha(I^n)$ if $|h|_\alpha < \infty$. Although the value of $|h|_\alpha$ depends on ε_0, the space $V_\alpha(I^n)$ does not. Moreover the value of ε_0 can be chosen in order to satisfy a few geometric constraints, like distortion, and to guarantee the forthcoming bound (2.10)[9]. We equip V_α with the Banach norm

$$||h||_\alpha := |h|_\alpha + ||h||_1,$$

and from now on V_α will denote the Banach space $\mathcal{B} = (V_\alpha(I^n), ||\cdot||_\alpha)$. With the assumptions we put on the map $\hat{T}$, in particular for the nature and smoothness of the boundaries of the domains U_k, it can be shown that the transfer operator $\hat{P}$ leaves V_α invariant with $\alpha = 1$, and moreover a Lasota-Yorke inequality (**P1**) holds, whenever

$$\eta := s_n + \frac{4s_n}{1-s_n} Z \frac{Y_{n-1}}{Y_n} < 1, \tag{2.10}$$

where Y_n is the volume of the unit ball in $\mathbb{R}^n$ and Z is the maximal number of the boundaries of the domains of local injectivity that meet in one point, see [27]. Also, one can show that $\mathcal{B}$ can be continuously injected into L^∞ and in particular, [27], $||h||_\infty \leq C_H ||h||_\alpha$, where $C_H = \frac{\max(1,\varepsilon_0^\alpha)}{Y_n \varepsilon_0^n}$.

Our next step is to show that $\mathcal{B}$ is invariant under the perturbed operator $\widetilde{P}_l$. By comparing with the computations in [27], we see that the new term we should take care of is:

$$|h\mathbf{1}_{D_l}|_\alpha = \sup_{0<\varepsilon\leq\varepsilon_0} \frac{1}{\varepsilon^\alpha} \int \text{osc}(h\mathbf{1}_{D_l}, B_\varepsilon(\overline{x}))d\text{Leb}.$$

Using the results in [27] and with $B_\varepsilon(D_l)$ denoting the ε-neighborhood of the set D_l[10] we have:

$$\text{osc}(h\mathbf{1}_{D_l}, B_\varepsilon(\overline{x})) \leq \text{osc}(h, D_l \cap B_\varepsilon(\overline{x}))\mathbf{1}_{D_l} +$$
$$2\left[\text{Esup}_{B_\varepsilon(\overline{x})\cap D_l}|h|\right] \mathbf{1}_{B_\varepsilon(D)\cap(B_\varepsilon(D_l^c))}(\overline{x}).$$

[9]For explicit computations of ε_0 on concrete examples, see [27] and [22]; for the example (2.2) below, that value was computed in Proposition 6 in [28].

[10]To be more precise we have $B_\varepsilon(D_l) := \{\overline{x} \in \mathbb{R}^n : \text{dist}(\overline{x}, D_l) \leq \varepsilon\}$.

By integrating and dividing by $\varepsilon^{-\alpha}$ we get

$$|h\mathbf{1}_{D_l}|_\alpha \leq |h|_\alpha + \sup_{0<\varepsilon\leq\varepsilon_0} \frac{2}{\varepsilon^\alpha} \int_{B_\varepsilon(\overline{x})\cap D} \sup |h(\overline{x})| \mathbf{1}_{B_\varepsilon(D_l)\cap(B_\varepsilon(D_l^c))}(\overline{x}) d\mathrm{Leb} \leq$$

$$|h|_\alpha + 2||h||_\infty \frac{\mathrm{Leb}(B_\varepsilon(D_l)\cap(B_\varepsilon(D_l^c))}{\varepsilon^\alpha}.$$

Before continuing we must say what really the set D_l is in our case. Its complement, D_l^c is given in (4.22) and with the actual notation reads

$$D_l^c = \{\overline{x} \in I^n : \max_{i\neq j} |x_i - x_j| \leq \nu_l\},$$

where ν_l goes to zero when $l \to \infty$. In this case it is easy to see that

$$\mathrm{Leb}(B_\varepsilon(D_l)\cap(B_\varepsilon(D_l^c)) \leq C_n \varepsilon \nu_l, \tag{2.11}$$

see Appendix 1 for the proof. Therefore we can continue the previous bound as:

$$|h\mathbf{1}_{D_l}|_\alpha \leq |h|_\alpha [1 + 2C_H C_n \varepsilon^{1-\alpha} \nu_l].$$

This computation shows that $\mathcal{B}$ is preserved by $\widetilde{P}_l$, but if we want to get a Lasota-Yorke inequality for it, and therefore satisfy (**P3**), we should multiply η by $(1 + 2C_H C_n \nu_l)$ and ask that $\eta(1 + 2C_H C_n \nu_l) < 1$, which is surely satisfied by taking l large enough. Alternatively, one could take higher iterates of $\hat{T}$. In this case the backward images of D_l will grow as well, but linearly with the power of the map and their contribution will be dominated by the exponential decay of the contraction factor.
As we said above property (**P2**) requires that the invariant measure of the unperturbed map be mixing; we will give an explicit example below.

Since quasi-Hölder functions h are essentially bounded, it is easy to get Property (**P4**) estimating as:

$$|\int (\hat{P}h - \widetilde{P}_l h) d\mathrm{Leb}| \leq \int |h\mathbf{1}_{D_l^c}| d\mathrm{Leb} \leq ||h||_\infty \mathrm{Leb}(D_l^c) \leq C_H ||h||_\alpha \mathrm{Leb}(D_l^c). \tag{2.12}$$

To check (**P5**), we begin to observe that the Banach norm $||(\hat{P} - \widetilde{P}_l)\hat{h}||_{\mathcal{B}}$ is bounded by a constant, say $\hat{C}$ depending on $\hat{h}$ as a consequence of the Lasota-Yorke inequality. Since the density is bounded away from zero, we immediately have $r_l ||(\hat{P} - \widetilde{P}_l)\hat{h}||_{\mathcal{B}} \leq \frac{C_H \hat{C}}{\hat{h}^{(\inf)}} \mu(D_l^c)$.

Example 2.2. We now give an easy example which satisfies **P1** to **P3** with $\mathcal{B}$ the space of quasi-Hölder functions; **P4** and **P5** follow from the above arguments and finally Property (**P6**) will be proved in Section 5 under the additional assumption **P0** and for a much larger class of maps. We stress that our example will be used for the numerical simulations in Section 7. Moreover the techniques we are using could be easily extendable to other transformations not necessarily affine. As the one-dimensional map T we will take $T(x) = 3x \bmod 1$. By coupling n of them as in (2.2) we get a piece-wise linear uniformly expanding higher dimensional map. We first notice that this map is not necessarily

continuous on the n-torus, but it satisfies the assumption (**P0**) in Section 5. The Lasota-Yorke inequality (2.1) can be proved for l large enough, say for $l > l_0$ if we verify the condition (2.10). If it does not hold for the map $\hat{T}$ it will be enough to get it for an iterate of $\hat{T}$ and this is surely possible thanks to Theorem 11 in Tsujii's paper [28], which holds for expanding piecewise linear maps whose locally domains of injectivity are bounded by polyhedra. The constants η and C in (2.1) depend in our case (local affine maps), simply on the contraction rate $s_n = 3^{-n}$ to the power l. The next step is to prove the bound (2.4).

This can be easily achieved by adapting our proofs of Proposition 4.3 in [1], or of Lemma 7.5 in [20]. The basic ingredients of such proofs are: (i) the control of the distance between the preimages of the same point $z \in I^n$ with the maps $\hat{T}_0$ and $\hat{T}$ (for a given γ); (ii) the distortion, involving the two determinants $|\det(D\hat{T}_0)|$ and $|\det(D\hat{T})|$ (for a given γ). By the structure of the map (2.2), one immediately sees that the distance at point (i) is of order γ times a constant depending on the dimensionality n of ambient space. The ratio of the determinants at point (ii) is instead of order $(1-\gamma)^n$ as it follows from the proof of Proposition 3.2 below. This is enough to obtain the bound (2.4); we left the details to the reader. We should finally check that the invariant density is bounded away from zero for the map $\hat{T}$. We dispose of, at least, two criteria of *covering* type for that. The first is taken from Section 7.3.1 and Lemma 7.5 in our paper [20] and requires the existence of a domain of local injectivity U_k (see Section 2), whose image is the full hypercube I^n. The second is described in Sublemma 5.3 in [22] and requires the so-called *topological exactness*, namely the existence for any $\overline{x} \in I^n$ and $\varepsilon > 0$, of an integer $N_\varepsilon = N_\varepsilon(x, \varepsilon) > 0$ such that $\hat{T}^{N_\varepsilon} B_\varepsilon(\overline{x}) = I^n$. Both results rely on an interesting property of the quasi-Hölder functions, namely the existence of a ball where the (essential) infimum of such a function is bounded away from zero, see [27]. We believe such covering conditions are satisfied in our cases. As an example, we report the computation of the density for two and three coupled maps; it is also interesting to observe that the density does not oscillate too much in the vicinity of the diagonal, which is required by our assumption **P8**, see Figure 1 for $n = 2$ and Figure 2 for $n = 3$.

3. Extreme values and localizations

In this section and in the next one, we apply EVT to the study of a few recurrence behaviors for our system of CML.

There are, at least, two approaches to EVT. The first, which we call the pure probabilistic one (PPA) uses strong mixing properties to get fast decay of correlations for a suitable class of observables and to control short returns around a given point. It is worth mentioning that the PPA covers cases where there is no spectral gap and therefore

the correlations do not decay exponentially fast, see for instance [25] for a rich variety of examples.

The second approach, developed by Keller [12] and which we name the spectral approach (SA), is based on the perturbation technique discussed in the preceding section and which allow us to get Gumbel's law directly by a smooth perturbation of the spectral radius of the operator $\widetilde{P}_l$. We will show explicitly in Section 4 how this method works. The SA seems particularly adapted to investigate synchronization, while the PPA is not suited, for the moment, to study observables which become infinite on sets with uncountably many points, which is what happens when we consider synchronization (along the diagonal). As we have already pointed out in the previous section, the issue in the SA is to verify the properties **P1-P6**.

Let us suppose the vector $\overline{z}$ is given. When the orbit of a point $\overline{x}$ enters in a sufficiently small ball centered at $\overline{z}$ we will say that there is *localization* of the orbit around the point $\overline{z}$.

Let us introduce the observable

$$\varphi(\overline{x}) := -\log\{\sum_{i=1}^{n} |x_i - z_i|\}, \tag{3.13}$$

and consider the maximum

$$M_m(\overline{x}) := \max\{\varphi(\overline{x}), \varphi(\hat{T}\overline{x}), \dots \varphi(\hat{T}^{m-1}\overline{x})\}. \tag{3.14}$$

By adopting the point of view of EVT, we will fix a positive number τ and we will ask for the existence of a sequence u_m for which the following limit exists

$$m\ \hat{\mu}(\varphi > u_m) \to \tau,\ m \to \infty. \tag{3.15}$$

We will say that *the sequence M_m has an Extreme Value Law, (EVL)*, if there exists a non-degenerate distribution function $H : \mathbb{R} \to [0, 1]$, with $H(0) = 0$ such that

$$\hat{\mu}(M_m \leq u_m) \to 1 - H(\tau),\ m \to \infty. \tag{3.16}$$

By using the expression of φ we can rewrite (3.15) as

$$m\ \hat{\mu}(U_m^{(n)}) \to \tau. \tag{3.17}$$

where

$$U_m^{(n)} := \{\overline{x} \in I^n : \sum_{i=1}^{n} |x_i - z_i| \leq \nu_m\},\ \text{with } \nu_m := e^{-u_m} \tag{3.18}$$

and consequently (3.16) can be restated as

$$\hat{\mu}(\overline{x} \in I^n : \hat{T}^k(\overline{x}) \notin U_m^{(n)}, k = 0, \dots, m-1) \to 1 - H(\tau). \tag{3.19}$$

We call ν_m the *accuracy of the localization* and we use the symbol a_c to denote it. Of course it depends on m, but as we will see soon, it is sometimes convenient to fix the

value of a_c and choose m accordingly. If we see $\{\hat{T}^k\}_{k\geq 1} : I^n \to I^n$ as a vector valued random variable on the space $\{I^n, \hat{\mu}\}$ associating to the point $\overline{x} \in I^n$ its orbit, then the limit (3.19) could also be interpreted as the *probability that each component* $\{\hat{T}_i^k\}_{k\geq 1}$ *is localized with accuracy* $a_c = e^{-u_m}$ *around* z_i *for the first time when* $k > m$. In order to get the probability of such an event, we have to insure a few assumptions, which were already anticipated in the previous section, and which will allow us to apply Proposition 3.3 in [1] that we restate in the following proposition:

Proposition 3.1. *Suppose that the system* $(I^n, \hat{T}, \hat{\mu})$ *has a unique absolutely continuous invariant and mixing measure* $\hat{\mu}$ *with density bounded away from zero and exponential decay of correlations on an adapted Banach space. Let* $(X_0, X_1, \cdots)$ *be the process given by* $X_m = \varphi \circ \hat{T}^m, m \in \mathbb{N}$, *where* φ *achieves a global maximum at some points* $\overline{z}$. *Then we have an EVL for the maximum* M_m *and:*
(1) if $\overline{z}$ *is not a periodic point, then the EVL is such that* $H(\tau) = 1 - e^{-\tau}$*;*
(2) if $\overline{z}$ *is a (repelling) periodic point of prime period* p*, then the EVL is such that* $H(\tau) = 1 - e^{-\theta\tau}$ *, where the* extremal index (EI) *is given by* $\theta(\overline{z}) = 1 - |\det D(\hat{T}^p)(\overline{z})|^{-1}$.

We notice that eventually (repelling) periodic points fall in part (1). Our observable (3.13) satisfies the assumption of the Proposition. On the other hand, by using Theorem 1.7.13 in [24], we have a sufficient condition to guarantee the existence of the limit (3.15) for $0 < \tau < \infty$. Such a condition requires that $\frac{1-F(x)}{1-F(x-)} \to 1$, as $x \to u_F$, where F is the distribution function of X_0, the term $F(x-)$ in the denominator denotes the left limit of F at x and $u_F = \sup\{x : F(x) < 1\}$. For the observable just introduced $u_F = \infty$ and if the probability $\hat{\mu}$ is not atomic at $\overline{z}$, then it is easy to conclude that F is continuous at $\overline{z}$ and therefore the above ratio goes to 1.

This general result will not allow us to explicitly compute the sequence u_m. Let us take the affine sequence: $u_m = \frac{y}{a_m} + b_m$, with $a_m > 0$, and $y \in \mathbb{R}$. This suggests that we redefine $u_m(y)$ as a one parameter family in y. When the sequence $\hat{\mu}(M_m \leq u_m) = \hat{\mu}(a_m(M_n - b_m) \leq y)$ converges to a non-degenerate distribution function $G(y)$, in the point of continuity of the latter, then we have an EVL. It is a beautiful result of EVT, just related to the affine choice for the sequence u_n [11], that such a $G(y)$ could be only of three types, called Gumbel, Fréchet and Weibull, see [24], and what determines it in a particular situation is just the common distribution given by the function F.

For instance and in our case, if we suppose that the invariant measure behaves like Lebesgue, $\hat{\mu}(U_m^{(n)}) = O(\nu_m^n)$[12], then $e^{-u_m} \sim \left(\frac{\tau}{m}\right)^{\frac{1}{n}}$, or equivalently $u_m \sim \frac{1}{n}\log m - \frac{1}{n}\log\tau$

[11] For other choices of the sequence u_n, see [24].

[12] Actually we have $\hat{\mu}(U_m^{(n)}) = O(2^n\nu_m^n)$, but the factor 2^n will become negligeable by taking large m.

and therefore *the probability of the first localization after m iterations with m large and with accuracy a_c of order $\left(\frac{\tau}{m}\right)^{\frac{1}{n}}$, is $e^{-\tau}$*, or equivalently $e^{-e^{-y}}$, having set $\tau = e^{-y}$. The distribution function $e^{-e^{-y}}, y \in \mathbb{R}$ is just the Gumbel law. In this easy example $a_m = n, b_m = \frac{1}{n}\log m$, but we used very crude approximation in estimating the $\hat{\mu}$-measure of the parallelepiped $U_m^{(n)}$ since we simply forgot the local density of the measure at the point $\overline{z}$. Very often it is a difficult task to get an explicit expression for the scaling coefficients a_m, b_m. In a few cases one succeeds, see the results in [25], Propositions 7.2.4, 7.4.1, 7.5.1. Otherwise and for practical purposes, the distribution function $\hat{\mu}(M_m \leq y)$ is modeled, for m sufficiently large, by the so-called *generalized extreme value (GEV)* distribution which is a function depending upon three parameters $\xi \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0$: $\mathrm{GEV}(y;\mu,\sigma,\xi) = \exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$.

The parameter ξ is called the tail index. When it is 0, the GEV corresponds to the Gumbel type, when the index is positive, it corresponds to a Fréchet and finally when it is negative, it corresponds to a Weibull. The parameter μ is called the location parameter and σ is the scale parameter: for m large the scaling constant a_m is close to σ^{-1} and b_m is close to μ.

The proof of Proposition (3.1) can be done with the SA or the PPA approaches and the latter uses the approximation of our process with an i.i.d. process, this being guaranteed by the exponential rate of mixing of the measure $\hat{\mu}$ on functions in $\mathcal{B}$. It is interesting to point out the dichotomy in the choice of the target point $\overline{z}$: there is only two functional expressions for the distribution $H(\tau)$ and what determines such a difference is the possible periodicity of $\overline{z}$. We now focus on the EI θ. Suppose we have successive entrances in the neighborhood of $\overline{z}$, namely consecutive occurrences of an exceedance of our threshold u_n. We interpret it as a memory of the underlying random process, and we quantify it with the parameter θ. In particular, see [25], p. 34, when $\theta > 0$ and for most of the times, the inverse of the EI defines the mean number of exceedances in a cluster of large observations, i.e., is the mean size of the clusters. We now show that in our model and whenever the number of components of the lattice goes to infinity, the EI of periodic points goes to 1, so there are no clusters in the limit of an infinitely large lattice.

We now have (from now on we write the EI as θ_n to signify the dependence on n)

Proposition 3.2. *Let $\hat{T}$ be the CML with n sites given by (2.2) and take $\gamma < 1-\lambda$. Fix $p \geq 1$, if $\overline{z}_n^{(p)} \in I^n$ is a periodic point of prime period p, the EI $\theta_n(\overline{z}_n^{(p)})$ satisfies*

$$\lim_{n\to\infty} \theta_n(\overline{z}_n^{(p)}) = 1.$$

Proof. Recall the definition of the uncoupled dynamics

$$\hat{T}_0(\overline{x}) = \big(T(x_1), T(x_2), \ldots, T(x_n)\big) \qquad \forall\, \overline{x} = (x_1, x_2, \ldots, x_n) \in I^n,$$

and let C_γ be the real $n \times n$ matrix, whose coefficients $(C_\gamma)_{ij}$ are defined by $(C_\gamma)_{ij} = \gamma/n$ if $i \neq j$ and $(C_\gamma)_{ij} = (1-\gamma) + \gamma/n$ if $i = j$. It is easy to check that $\hat{T} = \Phi_\gamma \circ \hat{T}_0$, where the coupling operator $\Phi_\gamma : I^n \to I^n$ is the linear map associated to the matrix C_γ (i.e $\Phi_\gamma(\overline{x}) := C_\gamma \overline{x}$).

Let $p \geq 1$, $\overline{z} \in I^n$ and let us compute the determinant of the Jacobian matrix of $\hat{T}^p$ evaluated in the point $\overline{z}$. We have,

$$\det(D_{\overline{z}}\hat{T}^p) = \prod_{t=0}^{p-1} \det(D_{\hat{T}^t(\overline{z})}\hat{T}) = \prod_{t=0}^{p-1} \det(D_{\hat{T}_0(\hat{T}^t(\overline{z}))}\Phi_\gamma D_{\hat{T}^t(\overline{z})}\hat{T}_0)$$

$$= \prod_{t=0}^{p-1} \det(C_\gamma D_{\hat{T}^t(\overline{z})}\hat{T}_0) = \det(C_\gamma)^p \prod_{t=0}^{p-1} \det(D_{\hat{T}^t(\overline{z})}\hat{T}_0).$$

It is an easy exercise in linear algebra to show that the determinant of the symmetric matrix C_γ is $\det(C_\gamma) = (1-\gamma)^{n-1}$.

On the other hand $D_{\overline{z}}\hat{T}_0$ is a diagonal matrix with diagonal entries $T'(z_1), \cdots, T'(z_n)$ and corresponding Jacobian determinant in $\overline{z}$ given by $\det\left(D_{\overline{z}}\hat{T}_0\right) = \prod_{k=1}^n T'(z_k)$. It follows that

$$|\det\left(D_{\overline{z}}\hat{T}^p\right)| = (1-\gamma)^{p(n-1)} \prod_{t=0}^{p-1}\prod_{k=1}^{n} |T'((\hat{T}^t(\overline{z}))_k)| = (1-\gamma)^{p(n-1)} \prod_{t=0}^{p-1}\prod_{k=1}^{n} |T'(z_k^t)|.$$

According to [1], if $\overline{z}_n^{(p)}$ is a periodic point of period p, the EI satisfies

$$\theta_n(\overline{z}_n^{(p)}) = 1 - \frac{1}{|\det\left(D_{\overline{z}_n^{(p)}}\hat{T}^p\right)|}.$$

Since

$$|\det\left(D_{\overline{z}_n^{(p)}}\hat{T}^p\right)| \geq (1-\gamma)^{p(n-1)}(\frac{1}{\lambda})^{np} = \left(\frac{(1-\gamma)^n(\frac{1}{\lambda})^n}{(1-\gamma)}\right)^p,$$

and as $\gamma < 1 - \lambda$, we have that $\lim_{n\to\infty} \theta_n(\overline{z}_n^{(p)}) = 1$. □

3.1. **Random perturbations.** There is another situation which produces an extremal index equal to 1. We can perturb the map $\hat{T}$ with additive noise, see [1], by defining a family of maps $\hat{T}_{\underline{\omega}} = \hat{T} + \underline{\omega}$, with each vector $\underline{\omega}$ belonging to the set Ω and chosen in such a way that each $\hat{T}_{\underline{\omega}}$ sends I^n into itself. The iteration of $\hat{T}$ will be now replaced by the concatenation

$$\hat{T}_{\overline{\omega}}^n := \hat{T}_{\underline{\omega}_n} \circ \cdots \circ \hat{T}_{\underline{\omega}_1}, \text{ with } \overline{\omega} := (\underline{\omega}_1, \cdots, \underline{\omega}_n, \cdots) \in \Omega^{\mathbb{N}},$$

and the $\underline{\omega}_k$ chosen in an i.i.d. way in Ω according to some (common) distribution $\mathbb{P}$. If we now take any measurable real observable φ, the process $\{\varphi \circ \hat{T}_{\overline{\omega}}^n\}_{n\geq 1}$ will be stationary with respect to the product measure $\hat{\mu}_s \times \mathbb{P}^{\mathbb{N}}$, where $\hat{\mu}_s$ is the so-called *stationary measure*, verifying, for any real measurable bounded function f: $\int f d\hat{\mu}_s = \int f \circ \hat{T}_\omega d\hat{\mu}_s$: see [25] Chapter 7 for a general introduction to the matter. We call the couple $\{\hat{T}_{\overline{\omega}}, \hat{\mu}_s \times \mathbb{P}^{\mathbb{N}}\}$ a *random dynamical system*. In the framework of EVT we could therefore consider the

process $\{X_{m,\overline{\omega}}(\cdot)\}_{n\geq 1} = \{\varphi \circ \hat{T}^m_{\overline{\omega}}(\cdot)\}_{n\geq 1}$, where φ is the observable introduced in (3.13), and consider accordingly the distribution of the maximum (3.14) with respect to the probability measure $\hat{\mu}_s \times \mathbb{P}^{\mathbb{N}}$. By adopting for $\hat{T}$ the same assumptions as in Proposition 3.1, it is not difficult to show that $\hat{\mu}$ is equivalent to Lebesgue and we finally proved in [1], Corollary 4.4, that for *any* choice of the target point $\overline{z}$, an extreme value distribution holds with $H(\tau) = 1 - e^{-\tau}$.

4. Extreme values and synchronization

We now introduce a new observable which allows us to consider synchronization of the n components of an initial state iterated by $\hat{T}$. Let us therefore define

$$\psi(\overline{x}) := -\log\{\max|x_i - x_j|, i \neq j : i, j = 1, \ldots, n\} \tag{4.20}$$

and consider the maximum

$$M_m(\overline{x}) := \max\{\psi(\overline{x}), \psi(\hat{T}\overline{x}), \ldots \psi(\hat{T}^{m-1}\overline{x})\}.$$

By adopting the point of view of EVT, we fix again a positive number τ and we ask for a sequence u_m for which the following limit exists $m\ \hat{\mu}(\psi > u_m) \to \tau,\ m \to \infty$. We say again that *the sequence M_n has an Extreme Value Law*, if there exists a non-degenerate distribution function $H : \mathbb{R} \to [0, 1]$, with $H(0) = 0$ such that $\hat{\mu}(M_m \leq u_m) \to 1 - H(\tau),\ m \to \infty$. By using the expression of ψ we can rewrite (3.15) as

$$m\ \hat{\mu}(S^{(n)}_m) \to \tau \tag{4.21}$$

$$S^{(n)}_m := \{\overline{x} \in I^n : \max_{i\neq j}|x_i - x_j| \leq \nu_m\},\ \text{where } \nu_m := e^{-u_m} \tag{4.22}$$

and consequently (3.16) can be restated as

$$\hat{\mu}(\overline{x} \in I^n : \hat{T}^k(\overline{x}) \notin S^{(n)}_m, k = 0, \ldots, m-1) \to 1 - H(\tau). \tag{4.23}$$

The limit (4.23) could also be interpreted as the *probability that the n components have synchronized for the first time after m iterations with accuracy a_c of order e^{-u_m}.*

We cannot use the PPA to prove the existence of the limit (4.23). The reason is that our new observable becomes infinite on a line (the diagonal), and for the moment rigorous results are avalaible when the set of points where the observable is maximised is at most countable, see [7] for a discussion of these problems.

The SA will bypass that issue by using the Banach space $\mathcal{B}$ given by quasi-Höder functions, since for such a space we can check properties **P1-P5**. Nevertheless there is still a problem remaining, namely prove the existence of the limits (2.6). We will return to that in the next section.

We now show how to get the asymptotic distribution functions of the extreme value theory by using the SA. Let us begin by rewriting the maximum given in (4.23) using the density $\hat{h}$ of the measure $\hat{\mu}$:

$$\hat{\mu}(M_n \leq u_m) = \int \hat{h}(\overline{x})\mathbf{1}_{(S_m^{(n)})^c}(\overline{x})\mathbf{1}_{(S_m^{(n)})^c}(\hat{T}(\overline{x}))\ldots\mathbf{1}_{(S_m^{(n)})^c}(\hat{T}^{m-1}(\overline{x}))d\text{Leb} = \int \widetilde{P}_m^m(\hat{h})d\text{Leb}, \tag{4.24}$$

where, from now on,

$$\widetilde{P}_m(\cdot) := \hat{P}(\mathbf{1}_{(S_m^{(n)})^c}\cdot).$$

Notice that $(S_m^{(n)})^c$ plays the role of the set D_l in Section 2. By invoking the spectral representation (2.9) we have with obvious interpretation of the symbols

$$\int \widetilde{P}_m^m(\hat{h})d\text{Leb} = \rho_m^m \int \hat{h}d\hat{\mu}_m + \int Q_m^m\hat{h}d\text{Leb},$$

where $\int \hat{h}d\hat{\mu}_m \to \int \hat{h}d\text{Leb} = 1$, as $m \to \infty$, and the spectral radius of Q_m is strictly less than ρ_m. We now need to bound ρ_m, the largest eigenvalue of $\widetilde{P}_m$, for increasing m and it is given by (2.8). Let us now denote the exponent θ_Δ the EI along the diagonal set $\Delta := \{\overline{x} \in \mathbb{R}^n; x_1 = x_2, \cdots = x_n\}$ and its existence will follow if we prove limit (2.6). We then write:

$$1 - \rho_m = \hat{\mu}(S_m^{(n)})\theta_\Delta(1 + o(1)), \text{ in the limit } m \to \infty,$$

then

$$\int \widetilde{P}_m^m(\hat{h})d\text{Leb} = e^{-(\theta_\Delta m\hat{\mu}(S_m^{(n)}) + mo(\hat{\mu}(S_m^{(n)}))} \int \hat{h}d\hat{\mu}_m + O(\rho_m^{-m}||Q_m^m||_{\mathcal{B}}) \tag{4.25}$$

which converges to $e^{-\tau\theta_\Delta}$ under the assumptions on $\hat{\mu}_m$, the spectral radius of Q_m and the condition (4.21). From now on we will simply write θ_n for the EI along the diagonal set for lattices with n components.

We now return to (4.23) since we now know that $1 - H(\tau) = e^{-\theta_n\tau}$. If we suppose that $\hat{\mu}(S_m^{(n)}) = O(\nu_m^{n-1})$[13], then $e^{-u_m} \sim \left(\frac{\tau}{m}\right)^{\frac{1}{n-1}}$ and therefore *the probability of the first synchronization after m iterations with accuracy $a_c \sim \left(\frac{\tau}{m}\right)^{\frac{1}{n-1}}$, is $e^{-\theta_n\tau}$.*[14] If the components of the vector $\hat{T}^k(\overline{x})$ are seen as the positions of different particles on a lattice, we have a quantitative estimate of the probability of synchronization of the lattice after a prescribed time and with a given accuracy.

[13] Actually this is a very crude approximation. In fact what is possible to prove easily is an upper bound on the Lebesgue measure of the domain $\{\overline{x} \in I^n, |x_i - x_j| < \nu_m : i \neq j\}$ which is simply $(2\nu_m)^{n-1}$. We sketch the argument for $n = 3$. In this case, the measure we are looking for is $\int dx_1 \int dx_2 \mathbf{1}_{\{|x_1-x_2|\leq\nu_m\}}(\overline{x}) \int dx_3 \mathbf{1}_{\{|x_1-x_3|\leq\nu_m\}}(\overline{x})\mathbf{1}_{\{|x_2-x_3|\leq\nu_m\}}(\overline{x})$. The last integral will contribute with $2\nu_m$ and so the second one.

[14] We defer to the discussion after Proposition 3.1 for the validity of this argument and its approximations.

Example 4.1.

- Ex. 1. *Suppose we use the data in Section 7, with an EI* $\theta_3 \sim 1-(\frac{10}{27})^2 \sim 0.86$, *having chosen* $\lambda = 1/3$ *and* $\gamma = 0.1$, *and take* 3 *particles each living on the unit interval. If we want to synchronize them with a probability larger than* $1/2$ *and an accuracy* $a_c = 0.01$ *before* m *iterations, then we have to iterate the lattice around* $m = 8\ 100$ *times.*
- Ex. 2. *Analogously, if we want to observe with a probability larger than* $1/2$ *the synchronization of* 100 *particles each living on the unit interval with an accuracy* $a_c = 0,01$ *and before* m *iterations of the CML, then* m *has to be larger than* 100^{100}.

5. Computation of the extremal index

The extremal index is given by formula (2.7). Keller showed in [12] that it coincides with that given in Proposition 3.1 for the process $X_m = \varphi \circ \hat{T}^m$ and the proof is exactly the computation we performed in the previous section. As we said in the introduction, the rigorous computation of the EI for two coupled maps was given in [13]. Their map was slightly different from ours in the sense that for the i-th component the averaged term $\frac{\gamma}{n}\sum_{j=1}^{n} T(x_j)$ does not contain the contribution of $T(x_i)$. They first observed that in (2.6), all the q_k but q_0, are zero due to the fact that the diagonal is invariant and q_0 reads:

$$q_0 = \lim_{m\to\infty} \frac{\hat{\mu}(S_m^{(2)} \cap \hat{T}^{-1} S_m^{(2)})}{\hat{\mu}(S_m^{(2)})} \tag{5.26}$$

This quantity was explicitly computed giving the formula [13]:

$$\theta_2 = 1 - \frac{1}{1-2\delta}\frac{1}{\int \hat{h}(x,x)dx}\int \frac{\hat{h}(x,x)}{|DT(x)|}dx,$$

where the density $\hat{h}$ has bounded variation and for almost every $x \in I$ the value $\hat{h}(x,x)$ is the average of the limits of $\hat{h}(x-u, x+u)$ and $\hat{h}(x+u, x-u)$ as $u \to 0$.

We get a similar result and still for $n = 2$, with a modification due to the fact that our map is different, see formula (5.35) in the remark below. Instead the density along the diagonal is defined again as a bounded variation function. It seems difficult to extend such a result in higher dimensions without much stronger assumptions. Before doing that, we will explore how the EI θ_n behaves for large n in a quite general setting with the objective to show that for large n such an index approaches 1 and therefore the Gumbel's law will emerge as the extreme value distribution.

We will index with n the invariant densities $\hat{h}_n$, while we continue to use the symbol $\hat{\mu}$ for the invariant measure, despite the fact that $\hat{\mu}$ depend on n too, *via* the density $\hat{h}_n$. Our next objective is to show that all the q_k but q_0 are zero. Such a result is claimed in [13] in dimension 2 and without proof; we sketch it below for the reader's convenience in any dimension and asking for a few assumptions.

We first notice that the quantities $q_{k,l}$ introduced in (2.6), read:

$$q_{k,l} := \frac{\int (\hat{P} - \widetilde{P}_l)\widetilde{P}_l^k(\hat{P} - \widetilde{P}_l)(\hat{h})d\mathrm{Leb}}{\hat{\mu}(D_l^c)} = \hat{\mu}_{D_l^c}\{x \in D_l^c : \mathbf{t}_{D_i^c}(x) = k+1\} \tag{5.27}$$

where $\hat{\mu}_{D_l^c}$ is the conditional measure to D_l^c, and $\mathbf{t}_{D_i^c}(x)$ denotes the first return time of the point $x \in D_l^c$ to D_l^c (we will come back on this equality in the next section). Additional properties are necessary; for that let us first denote with $V_\varepsilon(\Delta)$ an ε-neighborhood of the diagonal Δ.

- **P01**: The boundaries of the domains of local injectivity $U_1, \cdots, U_q$ (see Section 2) are union of finitely many discontinuity surfaces $\mathcal{D}_j, j = 1, \cdots, p$[15], which are co-dimension 1 embedded submanifolds. We denote by $\mathcal{D}$ the union of those discontinuity sets. Moreover $\forall \varepsilon > 0$ and $k \in \mathbb{N}$, let us denote with $F_{d,\varepsilon,k}$ the set of points $x \in V_\varepsilon(\Delta)$ for which there is a neighborhood $\mathcal{O}(x)$ such that $\mathcal{O}(x) \cup \Lambda \neq \emptyset$, and $\mathcal{O}(x) \cap (\mathcal{D} \cup \hat{T}^{-1}(\mathcal{D}) \cup \cdots \hat{T}^{-k}(\mathcal{D})) = \emptyset$. We require the existence of a constant C_k independent on ε such that $\hat{\mu}(F_{d,\varepsilon,k}^c) \leq C_k \sigma(\varepsilon)\hat{\mu}(V_\varepsilon(\Delta))$, where $\sigma(\varepsilon)$ goes to zero when $\varepsilon \to 0$.

- **P02**: Let us denote with $G_{d,\varepsilon}$ the set of points in $V_\varepsilon(\Delta)$ for which the segment of minimal lenght connecting one of this point to the diagonal intersects one component of $\hat{T}\mathcal{D}_j$. For ε small enough, we will assume that there is a constant C_d independent of ε such that $\hat{\mu}(G_{d,\varepsilon}) \leq C_d\ \kappa(\varepsilon)\ \hat{\mu}(V_\varepsilon(\Delta))$, where $\kappa(\varepsilon)$ goes to zero when $\varepsilon \to 0$.

Remark 5.1. *The condition* **P01** *means that for a large portion of points in the vicinity of the diagonal, we can find a neighborhood which intersects the diagonal but does not cross the discontinuity lines up to a certain order. The condition* **P02** *means that the piece of $V_\varepsilon(\Delta)$ which is crossed by an element of $\hat{T}\mathcal{D}_j$ has a length along the direction of Δ of order $\kappa(\varepsilon)$. Both situations happen when the crossing of the discontinuities are "transversal": it is easy to produces pictures in dimension $n = 2$ and $n = 3$. See Figure 3 for* **P01** *and Figure 4 for* **P02**. *In both cases we took ε small enough in such a way that the discontinuity behaves locally, when it intersects $V_\varepsilon(\Delta)$, as a line for $n = 2$ and as a plane for $n = 3$; moreover $\kappa(\varepsilon) = O(\varepsilon)$. We notice that condition* **P02** *requires the control only of the first images of $\mathcal{D}$ and also it is not necessary if the map $\hat{T}$ is onto on each U_l.*

We sketch the argument for $k = 1$, the others being similar. By replacing l with m in (5.27) we show that:

[15] We observe that the map could be continuous on such boundaries, but the first derivative surely is not.

Lemma 5.2. *The quantity*

$$\frac{\int(\hat{P}-\widetilde{P}_m)\widetilde{P}_m(\hat{P}-\widetilde{P}_m)(\hat{h})fdLeb}{\hat{\mu}(S_m^{(n)})}=\frac{\hat{\mu}(S_m^{(n)}\cap\hat{T}^{-1}(S_m^{(n)})^c\cap\hat{T}^{-2}S_m^{(n)})}{\hat{\mu}(S_m^{(n)})} \tag{5.28}$$

goes to zero when $m\to\infty$.

Proof. Let us take a point $x\in F_{d,\varepsilon,2}$. With these assumptions $\hat{T}$ and $\hat{T}^2$ are open maps on $\mathcal{O}(x)$. In particular, $\hat{T}^2(\mathcal{O}(x))$ will be included in the interior of one of the U_l and it will intersect Δ by the forward invariance of the latter. We now suppose that $\hat{T}^2(x)$ is in $V_\varepsilon(\Delta)$ and we try to prove that $\hat{T}(x)$ must be in $V_\varepsilon(\Delta)$ too. Let us call D_* the domain of the function $\hat{T}_*^{-1}$, namely the inverse branch of the map sending $\hat{T}(x)$ to $\hat{T}^2(x)$. If the distance between $\hat{T}^2(x)$ and any point $z\in\hat{T}^2(\mathcal{O}(x))\cap\Delta$, such that the segment $[\hat{T}^2(x),z]$ is included in D_*, is less than ε, we have done since $\mathrm{dist}(\hat{T}_*^{-1}(z),\hat{T}_*^{-1}(\hat{T}^2(x))=\mathrm{dist}(\widetilde{z},\hat{T}(x))\le\lambda\varepsilon$, where $\widetilde{z}=\hat{T}_*^{-1}(z)\in\Delta$. Notice that such a point $z\in\Delta$ should not be necessarily in $\hat{T}^2(\mathcal{O}(x))$, provided the segment $[\hat{T}^2(x),z]\in D_*$ and $\mathrm{dist}(z,\hat{T}^2(x))\le\ \varepsilon$. What could prevent the latter conditions to happen is the presence of the boundaries of the domains of definition of the preimages of $\hat{T}$, which are the images of $\mathcal{D}$. We should therefore avoid that $\hat{T}^2(x)$ lands in the set $G_{d,\varepsilon}$, which means we have to discard those points $x\in V_\varepsilon(\Delta)$ which are in $\hat{T}^{-2}G_{d,\varepsilon}$, and, by invariance, the measure of those point is bounded from above by $C_d\ \kappa(\varepsilon)\hat{\mu}(V_\varepsilon(\Delta))$. We now choose $\nu_m<\varepsilon$ and work directly with the sets $S_m^{(n)}$. The points which are not in $\hat{T}^{-2}G_{d,\nu_m}\cap S_m^{(n)}\cap F^c_{d,\nu_m,2}$ gives zero contribution to the quantity $\hat{\mu}(S_m^{(n)}\cap\hat{T}^{-1}(S_m^{(n)})^c\cap\hat{T}^{-2}S_m^{(n)})$, while the measure of the remaining points divided by $\hat{\mu}(S_m^{(n)})$ goes to zero for m tending to infinity.

□

Proposition 5.3. *Let us suppose our CML satisfies properties* **P1**-**P5** *on a Banach space* $\mathcal{B}$ *with* $\lambda=\inf|DT|^{-1}<1-\gamma$, *the density* $\hat{h}_n\in L^\infty$ *and* $\hat{h}_n^{(\inf)}:=\inf_{I^n}\hat{h}_n>0$. *Then*

$$\limsup_{m\to\infty}\frac{\hat{\mu}(S_m^{(n)}\cap\hat{T}^{-1}S_m^{(n)})}{\hat{\mu}(S_m^{(n)})}\le\frac{\lambda^{n-1}||\hat{h}_n||_\infty}{(1-\gamma)^{n-1}\hat{h}_n^{(\inf)}}$$

Remark 5.4. *The upper bound makes sense of course when the right hand side of the above inequality is less or equal to* 1. *Moreover the EI* θ_n *will converge to* 1, *under the additional* **P0** *assumptions, when* $n\to\infty$ *if the ratio* $\frac{||\hat{h}_n||_\infty}{\hat{h}_n^{(\inf)}}$ *does not grow faster than* υ^{n-1} *with* $\upsilon>(\lambda/(1-\gamma))^{-1}$.

Proof. We start by writing

$$\hat{\mu}\big(\mathcal{S}_m^{(n)}\cap\hat{T}^{-1}\mathcal{S}_m^{(n)}\big)=\int_{I^n}d\overline{x}\hat{h}_n(\overline{x})\mathbf{1}_{\mathcal{S}_m^{(n)}}(x)\mathbf{1}_{\mathcal{S}_m^{(n)}}(\hat{T}\overline{x})$$

$$=\int_I dx_1\int_{I^{n-1}}dx_2\dots dx_n\hat{h}_n(x_1,\dots,x_n)\mathbf{1}_{\mathcal{S}_m^{(n)}}(\overline{x})\cdot$$

$$\mathbf{1}_{\mathcal{S}_m^{(n)}}\Big((1-\gamma)T(x_1)+\frac{\gamma}{n}\sum_{i=1}^{n}T(x_i),\dots,(1-\gamma)T(x_n)+\frac{\gamma}{n}\sum_{i=1}^{n}T(x_i)\Big).$$

W now have to reduce the domain of integration of $I \ni x_1$ in two steps: the first, changing I into I'_m, consists in removing intervals of length $2\nu_m$ on the left and on the right on each boundary point of the $A_l, l = 1, \cdots, q$. Clearly the difference between the integrals over I and I'_m will converge to zero when $m \to \infty$ since the integrand functions are bounded (remember the density is in L^∞); this argument is made more precise in Appendix 2 together with the reason of that reduction. For the moment we simply write $\mathbf{I}(I \setminus I'_m)$ for the integral over $I \setminus I'_m$. By introducing the operator P_l acting on the variable x_l, $l \geq 2$, we could continue as:

$$\hat{\mu}\big(\mathcal{S}_m^{(n)}\cap\hat{T}^{-1}\mathcal{S}_m^{(n)}\big) = \mathbf{I}(I\setminus I'_m)+\int_{I'_m}dx_1\int_{I^{n-1}}dx_2\dots dx_nP_2\circ\cdots\circ P_n\Big[\hat{h}_n(x_1,\dots,x_n)\mathbf{1}_{\mathcal{S}_m^{(n)}}(\overline{x})\Big]\cdot$$

$$\mathbf{1}_{\mathcal{S}_m^{(n)}}\Big((1-\gamma)T(x_1)+\frac{\gamma}{n}\big(T(x_1)+x_2+\cdots+x_n\big),\dots,(1-\gamma)x_n+\frac{\gamma}{n}\big(T(x_1)+x_2+\cdots+x_n\big)\Big).$$

If we now introduce the sets

$$S_{m,\gamma}^{(n)}(Tx_1) = \{(x_2,x_3,\cdots,x_n)\in I^n :\mid T(x_1)-x_j\mid\leq\frac{\nu_m}{1-\gamma}, j=2,\dots,n,\ \mid x_i-x_j\mid\leq\frac{\nu_m}{1-\gamma}\quad i\neq j\neq 1\},$$

and

$$S_m^{(n)}(x_1) = \{(x_2,x_3,\cdots,x_n)\in I^n : |x_1-x_j|\leq\nu_m, j=2,\cdots,n,\mid x_i-x_j\mid\leq\nu_m\quad i\neq j\neq 1\},$$

we have

$$\frac{\hat{\mu}(S_m^{(n)}\cap\hat{T}^{-1}S_m^{(n)})}{\hat{\mu}(S_m^{(n)})}\leq$$

$$\frac{\int_{I'_m}dx_1\int_{S_{m,\gamma}^{(n)}(Tx_1)}dx_2\dots dx_nP_2\circ\cdots\circ P_n\Big[\hat{h}_n(x_1,\dots,x_n)\mathbf{1}_{\mathcal{S}_m^{(n)}}(\overline{x})\Big]+\mathbf{I}(I\setminus I'_m)}{\int_{I'''_m}dx_1\int_{S_m^{(n)}(x_1)}dx_2\cdot dx_n\hat{h}(x_1,\cdots,x_n)}$$

We reduced the domain of integration in the integral in the denominator from I to I'''_m : this kind of reduction will also affect I'_m and it will be explained in the Appendix 2. Let us now consider for simplicity the structure of the operators when $n = 3$:

$$P_2\circ P_3\big[\hat{h}_3(x_1,x_2,x_3)\mathbf{1}_{\mathcal{S}_m^{(3)}}(x_1,x_2,x_3)\big] =$$

$$\sum_j\sum_k\frac{\hat{h}_3(x_1,T_j^{-1}x_2,T_k^{-1}x_3)\mathbf{1}_{\mathcal{S}_m^{(3)}}(x_1,T_j^{-1}x_2,T_k^{-1}x_3)}{\mid DT(T_j^{-1}x_2)\mid\mid DT(T_k^{-1}x_3)\mid}\mathbf{1}_{TA_j}(x_2)\mathbf{1}_{TA_k}(x_3),\tag{5.29}$$

where $\{A_k\}$ denotes the intervals of monotonicity of the map T. The preceding constraints and the assumption $\gamma < 1-\lambda$ imply that: $\mid T_j^{-1}x_2 - x_1\mid < \nu_m, \mid T_k^{-1}x_3 - x_1\mid < \nu_m$. Since the original partition is finite, if we take first m large enough and having removed the intervals of length $2\nu_m$ around the boundary point of the domain of monotonicity of T, for any $x_1 \in I'_m$ there will be only one preimage which can contribute in each sum. By generalizing to n components we could therefore bound the term (5.29) by $\lambda^{n-1}||h||_\infty$. Moreover a simple geometrical inspection shows that the Lebesgue measures of the sets

$S^{(n)}_{m,\gamma}(Tx_1)$ and $S^{(n)}_m(x_1)$ are independent of the point x_1 and also the ratio of the two measures is independent of m and gives

$$\frac{\mathrm{Leb}(S^{(n)}_{m,\gamma})}{\mathrm{Leb}(S^{(n)}_m)} = \frac{1}{(1-\gamma)^{n-1}}, \tag{5.30}$$

see Appendix 2. We therefore get

$$\frac{\hat{\mu}\big(\mathcal{S}^{(n)}_m \cap \hat{T}^{-1}\mathcal{S}^{(n)}_m\big)}{\hat{\mu}(\mathcal{S}^{(n)}_m)} \leq \frac{\mathrm{Leb}(S^{(n)}_{m,\gamma})\lambda^{n-1}||\hat{h}_n||_\infty + \mathbf{I}(I\setminus I''_m)}{\mathrm{Leb}(S^{(n)}_m)\hat{h}^{(\inf)}_n}. \tag{5.31}$$

We now notice that $\mathbf{I}(I\setminus I''_m)$ can be immediately bounded by $||\hat{h}||_\infty \mathrm{Leb}(S^{(n)}_{m,\gamma})\mathrm{Leb}(I\setminus I''_m)$. This allows us to factorize the term $\mathrm{Leb}(S^{(n)}_{m,\gamma})$ in the denominator and divide it by $\mathrm{Leb}(S^{(n)}_m)$. By taking the lim sup we finally get our result. □

We can now strengthen the previous result by adding further assumptions. We start first with a stronger hypothesis on the invariant density which we will relax later on.

- **P7** The density $\hat{h}$ is continuous on I^n.

This condition is for instance satisfied in the uncoupled case for smooth and locally onto maps T of the unit circle.

Proposition 5.5. *Let us suppose that our CML satisfies properties* **P1**-**P5** *and* **P7** *on a Banach space $\mathcal{B}$ with $\lambda = \inf|DT|^{-1} < 1-\gamma$, then*

$$\lim_{m\to\infty} \frac{\hat{\mu}(S^{(n)}_m \cap \hat{T}^{-1}S^{(n)}_m)}{\hat{\mu}(S^{(n)}_m)} = \frac{1}{(1-\gamma)^{n-1}} \frac{\int_I \frac{\hat{h}_n(x,\cdots,x)}{|DT(x)|^{n-1}}dx}{\int_I \hat{h}_n(x,\cdots,x)dx}.$$

Proof. We will write the proof for $n=3$, the generalization being immediate, and this will allows us to use the simple formulas in the previous demonstration. By the same arguments in the latter and by denoting with $T^{-1}_{x_1}$ the inverse branch of T such that $T^{-1}_{x_1}(T(x_1)) = x_1$, we have

$$\hat{\mu}\big(\mathcal{S}^{(3)}_m \cap \hat{T}^{-1}\mathcal{S}^{(3)}_m\big) = \int_{I''_m} dx_1 \int_{S^{(3)}_{m,\gamma}(Tx_1)} \frac{\hat{h}_3(x_1, T^{-1}_{x_1}x_2, T^{-1}_{x_1}x_3)}{|DT(T^{-1}_{x_1}x_2)||DT(T^{-1}_{x_1}x_3)|}dx_2dx_3 + \mathbf{I}(I\setminus I''_m) \tag{5.32}$$

and we have a lower bound for $\hat{\mu}\big(\mathcal{S}^{(3)}_m \cap \hat{T}^{-1}\mathcal{S}^{(3)}_m\big)$ without the $\mathbf{I}(I\setminus I''_m)$ term. We call $\mathbf{I}(I''_m)$ the first integral on the right hand side.

Since $\hat{h}_3$ is continuous on I^3 and therefore uniformly continuous, having fixed $\widetilde{\varepsilon} > 0$, it will be enough to choose ν_m small enough (remember that $|T^{-1}_{x_1}x_2 - x_1| \leq \nu_m$, $|T^{-1}_{x_1}x_3 - x_1| \leq \nu_m$), to have $\hat{h}_3(x_1, T^{-1}_{x_1}x_2, T^{-1}_{x_1}x_3) = \hat{h}_3(x_1,x_1,x_1) + O(\widetilde{\varepsilon})$.

For the derivative we can use the fact that our map is C^2 on the interior of the $A_l, l = 1,\cdots,q$ and extendable with continuity on the boundaries to get by the mean value theorem

$$DT(T^{-1}_{x_1}x_2) = DT(x_1) + D^2T(\hat{x}_2)|T^{-1}_{x_1}x_2 - x_1|,\ DT(T^{-1}_{x_1}x_3) = DT(x_1) + D^2T(\hat{x}_3)|T^{-1}_{x_1}x_3 - x_1|$$

where $\hat{x}_2$ belongs to the interval with endpoints $T_{x_1}^{-1}x_2$ and x_1, and $\hat{x}_3$ belongs to the interval with endpoints $T_{x_1}^{-1}x_3$ and x_1 and these two intervals are in the domains where T is locally injective. By inserting these formulas in the definition of $\mathbf{I}(I''_m)$ we have:

$$\mathbf{I}(I''_m) = \int_{I''_m} dx_1 \frac{\hat{h}_3(x_1,x_1,x_1)}{|DT(x_1)|^2} \int_{S^{(3)}_{m,\gamma}(Tx_1)} \frac{dx_2dx_3}{[1+\frac{D^2T(\hat{x}_2)}{DT(x_1)}|T_{x_1}^{-1}x_2-x_1|][1+\frac{D^2T(\hat{x}_3)}{DT(x_1)}|T_{x_1}^{-1}x_3-x_1|]}+$$

$$\int_{I''_m} dx_1 \frac{1}{|DT(x_1)|^2} \int_{S^{(3)}_{m,\gamma}(Tx_1)} \frac{O(\widetilde{\varepsilon})}{[1+\frac{D^2T(\hat{x}_2)}{DT(x_1)}|T_{x_1}^{-1}x_2-x_1|][1+\frac{D^2T(\hat{x}_3)}{DT(x_1)}|T_{x_1}^{-1}x_3-x_1|]} dx_2dx_3$$

We now rewrite the first summand as

$$\mathcal{I}_{1,m} := \mathrm{Leb}(S^{(3)}_{m,\gamma}) \int_{I''_m} dx_1 \frac{\hat{h}_3(x_1,x_1,x_1)}{|DT(x_1)|^2} \frac{1}{\mathrm{Leb}(S^{(3)}_{m,\gamma})}$$

$$\int_{S^{(3)}_{m,\gamma}(Tx_1)} \frac{dx_2dx_3}{[1+\frac{D^2T(\hat{x}_2)}{DT(x_1)}|T_{x_1}^{-1}x_2-x_1|][1+\frac{D^2T(\hat{x}_3)}{DT(x_1)}|T_{x_1}^{-1}x_3-x_1|]} \tag{5.33}$$

where we have suppressed the dependence on Tx_1 in the Lebesgue measure of the external $S^{(3)}_{m,\gamma}$, which are independent of Tx_1 when $x_1 \in I''_m$, and the second summand as

$$\mathcal{I}_{2,m} := \mathrm{Leb}(S^{(3)}_{m,\gamma}) \int_{I''_m} dx_1 \frac{1}{|DT(x_1)|^2} \frac{1}{\mathrm{Leb}(S^{(3)}_{m,\gamma})}$$

$$\int_{S^{(3)}_{m,\gamma}(Tx_1)} \frac{O(\widetilde{\varepsilon})\ dx_2dx_3}{[1+\frac{D^2T(\hat{x}_2)}{DT(x_1)}|T_{x_1}^{-1}x_2-x_1|][1+\frac{D^2T(\hat{x}_3)}{DT(x_1)}|T_{x_1}^{-1}x_3-x_1|]}$$

Using same arguments we have:

$$\hat{\mu}\big(\mathcal{S}^{(3)}_m\big) = \mathrm{Leb}(S^{(3)}_m) \int_{I'''_m} dx_1 \hat{h}_3(x_1,x_1,x_1) \frac{1}{\mathrm{Leb}(S^{(3)}_m)} \int_{S^{(3)}_m(x_1)} dx_2dx_3+$$

$$\mathrm{Leb}(S^{(3)}_m) \int_{I'''_m} dx_1 \frac{1}{\mathrm{Leb}(S^{(3)}_m)} \int_{S^{(3)}_m(x_1)} O(\widetilde{\varepsilon})dx_2dx_3+\mathbf{I}(I\setminus I'''_m) = \mathcal{I}_{3,m}+\mathcal{I}_{4,m}+\mathbf{I}(I\setminus I'''_m) \tag{5.34}$$

and with a lower bound for $\hat{\mu}\big(\mathcal{S}^{(3)}_m\big)$ without the $\mathbf{I}(I\setminus I'''_m)$ term. Hence we get

$$\frac{\mathcal{I}_{1,m}+\mathcal{I}_{2,m}}{\mathcal{I}_{3,m}+\mathcal{I}_{4,m}+\mathbf{I}(I\setminus I'''_m)} \le \frac{\hat{\mu}(S^{(n)}_m \cap \hat{T}^{-1}S^{(n)}_m)}{\hat{\mu}(S^{(n)}_m)} \le \frac{\mathcal{I}_{1,m}+\mathcal{I}_{2,m}+\mathbf{I}(I\setminus I''_m)}{\mathcal{I}_{3,m}+\mathcal{I}_{4,m}}$$

As in the proof of Proposition 5.3, we have that $\mathbf{I}(I\setminus I''_m) \le ||\hat{h}||_\infty \mathrm{Leb}(S^{(3)}_{m,\gamma})\mathrm{Leb}(I\setminus I''_m)$ and $\mathbf{I}(I\setminus I'''_m) \le ||\hat{h}||_\infty \mathrm{Leb}(S^{(3)}_m)\mathrm{Leb}(I\setminus I'''_m)$. We can then factorize in the numerator and in the denominator the Lebesgue measures of the sets $S^{(3)}_{m,\gamma}$ and $S^{(3)}_m$ and remember that $\frac{\mathrm{Leb}(S^{(3)}_{m,\gamma})}{\mathrm{Leb}(S^{(3)}_m)} = \frac{1}{(1-\gamma)^2}$. After this factorization and when m goes to infinity, the remaining part of $\mathcal{I}_{1,m}$ converges to $\int_I dx_1 \frac{\hat{h}_3(x_1,x_1,x_1)}{|DT(x_1)|^2}$ by the dominated convergence theorem and the fact that $|T_{x_1}^{-1}x_j - x_1| \le \nu_m, j=2,3$, while the remaining part of $\mathcal{I}_{2,m}$ converges to an $O(\widetilde{\varepsilon})$ term. Still after the previous factorization, the remaining part of $\mathcal{I}_{3,m}$ goes to $\int_I dx_1 \hat{h}_3(x_1,x_1,x_1)$, while the remaining part of $\mathcal{I}_{4,m}$ goes to an $O(\widetilde{\varepsilon})$ term. The result then follows sending $\widetilde{\varepsilon}$ to zero. □

It is possible to relax the continuity assumption **P7** on the density by asking a much weaker property. It seems to us that this condition is natural, and probably unavoidable, in the sense that it controls the oscillations of the density in the neighborhood of the diagonal.

- **P8** Let us suppose the density $\hat{h}$ is in $V_1(I^n)$ and moreover

$$\hat{h}_D := \sup_{0<\varepsilon\leq\varepsilon_0} \frac{1}{\varepsilon}\int \mathrm{osc}(\hat{h}, B_\varepsilon(x,\cdots,x)))dx < \infty.$$

Proposition 5.6. *Let us suppose that our CML satisfies properties* **P1**-**P5** *and* **P8** *on the Banach space $\mathcal{B} = V_1(I^n)$ with $\lambda = \inf |DT|^{-1} < 1-\gamma$, then the statement in Proposition 5.5 holds.*

Proof. The proof follows the line of Proposition 5.5, with an essential change when we compare the density in the neighborhood of the point (x_1, x_1, x_1). In fact, we can now write

$$|\hat{h}_3(x_1, T_{x_1}^{-1}x_2, T_{x_1}^{-1}x_3) - \hat{h}_3(x_1,x_1,x_1)| \leq \mathrm{osc}(\hat{h}_3, B_{\nu_m}(x_1,\cdots,x_1)).$$

An quick inspection of the previous proof shows immediately that the integral $\int_I dx_1\, O(\widetilde{\varepsilon})$ will be now replaced with $\int_I dx_1\, \mathrm{osc}(\hat{h}_3, B_{\nu_m}(x_1,\cdots,x_1))$, and this last integral is bounded by $\hat{h}_D\nu_m$, which goes to zero when m tends to infinity. □

Corollary 5.7. *As a consequence of Propositions 5.5 and 5.6, the extremal index θ_n for maps satisfying* **P0** *too, is given by*

$$\theta_n = 1 - \frac{1}{(1-\gamma)^{n-1}} \frac{\int_I \frac{\hat{h}_n(x,\cdots,x)}{|DT(x)|^{n-1}}dx}{\int_I \hat{h}_n(x,\cdots,x)dx}. \tag{5.35}$$

and it will converge to 1 *when $n \to \infty$.*

5.1. **Random perturbations.** As for localization, we expect that the extremal index be one when we keep n fixed and we add noise to the system. In the paper [1] we extended the SA to randomly perturbed dynamical systems, mostly with additive noise. Even if we assume properties (**P1**)-(**P6**) on some Banach space $\mathcal{B}$, there will be a new difficulty related to the computation of the quantities q_k in (2.6) in the random setting. Such a computation as it was done in Proposition 5.3 in [1] strongly relies on the fact that the observable becomes infinite in a single point, the center of a ball: we do not know how to adapt it in the neighborhood of the diagonal Δ. We will present nevertheless numerical evidences in Section 7 that in presence of noise the EI is 1.

6. Distribution of the number of successive visits

We anticipated in the introduction that once the synchronization is turned on for the first time, it cannot last since almost every orbit is recurrent. However the orbit $\hat{T}^n(\overline{x}_0)$ will visit for almost every point $\overline{x}_0$ infinitely often the neighborhood of the diagonal.

We could therefore expect that the exponential law $e^{-\tau}$ given by the EVT describes the time between successive events in a Poisson process. To formalize this, let us take a neighborhood $S_\varsigma^{(n)}$ of the diagonal Δ with accuracy $a_c = \varsigma$ and introduce the following quantity (remember that the map $\hat{T}$ and the measure $\hat{\mu}$ depend on n too):

$$N_\varsigma^{(n)}(t) = \sum_{l=1}^{\left\lfloor \frac{t}{\hat{\mu}(S_\varsigma^{(n)})} \right\rfloor} \mathbf{1}_{S_\varsigma^{(n)}}(\hat{T}^l(\overline{x})),$$

where $\lfloor \cdot \rfloor$ is the floor function, and consider the following distribution

$$\mathcal{N}(n, \varsigma, t, k) := \hat{\mu}(N_\varsigma^{(n)}(t) = k)$$

If the target set was a ball of radius ς around a generic point $\overline{z}$ or a dynamical cylinder set converging to this point, one can prove under the mixing assumptions of our paper, that in the limit of vanishing radius or infinite length for the cylinder, $\mathcal{N}(n, \varsigma, t, k)$ converges to the Poisson distribution $\frac{t^k e^{-t}}{k!}$, see for instance [17],[18]. Instead if we take the target point z periodic of minimal period q, one get the so-called compound Poisson distribution, see [22] and [9], which in our situation reads, for $k \geq 1$:

$$\mathcal{N}(n, \varsigma, t, k) = e^{-t(1-p)} \sum_{j=0}^{k} p^{k-j}(1-p)^{j+1} \frac{t^j (1-p)^j}{j!} \binom{k-1}{j-1} \tag{6.36}$$

where

$$p = \frac{1}{|\det(D_{\overline{z}} \hat{T}^q)|} \tag{6.37}$$

Remark 6.1. *We do not dispose for the moment of analogous formulas when a ball is replaced by a strip along our diagonal set Δ. To the best of our knowledge the only known result is in dimension 2 for the* uncoupled *systems given by the direct product of two piece-wise expanding and smooth maps of the circle, see* [5]*, and it is consistent with our results. Nevertheless a few preliminary considerations*[16] *seem to indicate that the compound distribution (6.36) still holds with p in (6.37) replaced by $1 - \theta_n$ in (5.35), and more generally with the EI given by formulas (2.7), (2.6), with the quantities $q_{k,l}$ given by the right hand side of (5.27) when the transfer operator is not available. In particular one should recover a pure Poisson distribution when the size n of the lattice tends to infinity.*

Example 6.2. *(Ex. (4.1) revisited)*

- *Suppose we consider as in the example (4.1), Ex. 2, $n = 100$ particles living in the unit interval and take the accuracy $\varsigma = 0.01$. With that value of n and taking the coupling γ sufficiently small, we could consider that the previous number of visits $N_\varsigma^{(n)}(t)$ follows a Poisson distribution. Since the probability of entering the neighborhood of the diagonal is of order 100^{-100}, the probability to observe*

[16] At this regard see also the discussion in the last part of Section 7.

exactly 5 *synchronization events during* m *iterations of the lattice is maximal for* $m = 5\ 100^{100}$ *and is of order* 18%.

- *If instead we consider Ex. 1 with* 3 *particles and the same accuracy, the probability to observe* 5 *synchronizations is maximal after* 50 000 *iterations and it is again of order* 18%.

Comment 6.3. *In the case of large* n *the extremely high number of iterations needed to get synchronization or a given number of successive synchronizations could surprise. One reason is surely due to the fact that we considered lattices which are globally coupled and we looked at global synchronization. It would be interesting, and it will be the objects of future investigations, to explore CML where only the nearest-neighbors of a given site contribute to the coupling term (*diffusive coupling*), and also synchronization of the closest neighbors. About the latter we will give a few preliminary numerical results in the next section.*

7. Extensions and Numerical computations

The goal of numerical computations will be to show that in the situations considered above we have effective convergence toward an extreme value law and moreover the extremal index satisfies the behavior we predicted theoretically. We will be mostly interested in synchronization, since for localization we have plenty of analytic results. But there is one aspect where the comparison with localization is particularly useful. In order to explain that, we first have to introduce a new observable to depict a different kind of synchronization.

7.1. **Local synchronization.** Up to now synchronization was defined by asking that all the components of the evolutionary state become close to each other with a given accuracy a_c. We could ask instead that each component synchronize only with the close neighbors. This is done by introducing the following observable

$$\Theta(\overline{x}) := -\log\{\max|x_i - x_j|, i \neq j : j = i \pm 1\} \tag{7.38}$$

(of course on the extreme points of the period of the lattice, j will take only one value). We could generalize to more than one neighbor $j = i \pm 2, \pm 3$, etc., but we limit ourselves here to the case ± 1. It is not immediately obvious to have a geometrical description of the set that the orbit will visit for the first time (and therefore to give analytic results in terms of the EI), although the "physical" interpretation will be the same, namely we get the probability that the lattice will have for the first time and after a given number of iterations m, all the components synchronized with the close neighbors and with a given accuracy a_c. We call this *local synchronization*, to distinguish from the *global synchronization* described in the preceding sections. It seems intuitive from a physical

point of view, that for m large enough and for a given accuracy a_c, the probability to get local synchronization for the first time (from now on we write it as $\mathbb{P}_1(\cdot)$ for the different cases), is larger than that to get global synchronization, $\mathbb{P}_1(\text{glob. sync.}) \leq \mathbb{P}_1(\text{loc. sync.})$, and this will be confirmed by the numerical simulation as we will see in a moment. On the other hand as soon as the global synchronization occurs, all the components of the lattice will be aligned in a narrow strip around all of them, and this is close to localization. Therefore we will expect that the probability to get localization is larger than the probability of global synchronization. This is also confirmed by an easy application of the theory. Suppose we fix m and the accuracy a_c; we have also fixed n. By supposing a pure exponential law for the asymptotic distribution of the maximum, we have

- For localization: $a_c \sim (\frac{\tau}{m})^{\frac{1}{n}}$, which gives $\mathbb{P}_1(\text{local.}) \sim e^{-\tau} \sim e^{-ma_c^n}$.
- For global synchronization: $a_c \sim (\frac{\tau}{m})^{\frac{1}{n-1}}$, which gives $\mathbb{P}_1(\text{glob. sync.}) \sim e^{-\tau} \sim e^{-ma_c^{n-1}}$.

We see that $\mathbb{P}_1(\text{glob. sync.}) \leq \mathbb{P}_1(\text{local.})$.

7.2. **Blocks of synchronization.** The observable (7.38) could be modified further by introducing a new one which we are going to define. Let us first construct N blocks of L successive integer indices: $B_q := \{i_q, \cdots, i_q + L\}$ and take these blocks disjoint and possibly scattered along the lattice. Then we define:

$$\Upsilon(\overline{x}) := -\log\{\max|x_i - x_j|, i \neq j : (i,j) \in B_q, q = 1, \cdots, N\}.$$

The distribution of the maximum of this observable will give us the probability that the particles in the N blocks will synchronise for the first time with a given accuracy. On the other hand we do not require any synchronization of the particles outside those blocks. If such a limiting distribution would exist, it could be consistent with the appearance of *chimeras* in chains of coupled particles, namely patterns of synchronized sets which emerge as a consequence of the self-organization of the entire lattice, see e.g. [26]. If our claim would be confirmed, such a self-organization would be another statistical property of chaotic systems with several degrees of freedom.

7.3. **Simulations.** Let us now analyze the results of numerical procedure. The experiment performed is the following: we consider the one-dimensional map T in (2.2) as

$$T(x) = 3x \mod 1.$$

Once we have constructed the CML $\hat{T}$ we will perturb it with additive noise:

$$\hat{T}_{\underline{\omega}}(\overline{x})_i = \hat{T}(\overline{x})_i + \varepsilon\omega_i \mod 1$$

where ε is here the noise intensity and $\underline{\omega}$ with components ω_i is a random variable drawn from a uniform distribution between -0.5 and +0.5. The stationary measure for such a map will be L^1 close to that for $\gamma = 0$ which is the direct product of the uniform Lebesgue measures on the unit circle for each component and this independently of the value of ε. Let us notice that we are considering now a one-dimensional map on the circle. This is not a restriction to our previous considerations and moreover it allows us to define correctly the additive noise. Numerically we produce trajectories of 10^4 iterations for $\gamma < 2/3$ and 0.02 increments. The range $3 < n < 53$ is analyzed. We consider the two observables ψ, see (4.20) and Θ, see (7.38), corresponding to global and local synchronization cases respectively and in the following we will refer to them as the *global and local cases*. We analyze also the role of small noise $\varepsilon = 10^{-4}$ and moderate noise $\varepsilon = 10^{-2}$.

We first assess the convergence of the maxima of ψ and Θ to the Gumbel law by analyzing the tail index ξ, see Section 3. Here we chose to consider the complementary approach to the block-maxima selection, i.e. the peak over threshold. The two approaches are equivalent in chaotic systems as shown in [25]. The maxima of the observable are defined as the exceedances over the 0.98 quantile of ψ and Θ distributions. If a good convergence towards the Gumbel law is reached, then $\xi \simeq 0$. The values of ξ as a function of γ and n are reported in Figure 5. A maximum likelihood estimator has been used for computation. The left panels show the global case ψ while the local case Θ is reported on the right. From top to bottom we switch on the noise. In general, the convergence towards the Gumbel law is satisfactory although some differences exist between global and local cases. For the global case the convergence is slower as the global synchronization event is more rare then the local one.

Moreover, the quality of the fits is lower when n and γ are larger. The addition of noise helps the convergence to the Gumbel law as for the systems analyzed in [25].

We now study the implications of global and local synchronization on the extremal index θ. For the analysis presented in this paper, we adopt the estimator by Süveges (see the book [25] for explanation and to retrieve the codes for the computation). For fixed quantile q, Süveges' estimator reads:

$$\theta = \frac{\sum_i^{N_c}(1-q)S_i + N + N_c - \sqrt{\left(\sum_i^{N_c}(1-q)S_i + N + N_c\right)^2 - 8Nc\sum_i^{N_c}(1-q)S_i}}{2\sum_i^{N_c}(1-q)S_i},$$

where N is the number of recurrences above the chosen quantile, N_c is the number of observations which form a cluster of at least two consecutive recurrences, and S_i the length of each cluster i. From the numerical point of view, this estimator is the expected

value of the compound distribution $\mathcal{N}(n, \varsigma, t, k)$ with S_i being the empirical equivalent of the quantity $N_\varsigma^{(n)}(x)$.

We begin by checking the theoretical results predicted in Remark 5.4: for the $3x$ mod 1 map, θ_n can now be estimated by taking the trace of the density on the diagonal reasonably of order 1 in (5.35), so that in dimension 2: $\theta_2 \sim 1 - \frac{1}{(1-\gamma)}\frac{1}{3}$ and in dimension 3: $\theta_3 \sim 1 - \frac{1}{(1-\gamma)^2}\frac{1}{9}$.

The comparison between the theoretical curves and the numerical computations are shown in Figure 6. For each case $n = 2, 3$ and $\gamma < 2/3$ we produce 10 simulations of the map consisting of 10^4 iterations and we estimate the extremal index as a function of γ. The numerical estimates indeed match the theoretical curves (bold magenta lines).

We now check the asymptotic formula for large n and still with the same assumption on the trace of the density, namely $\theta_n \sim 1 - (\frac{\lambda}{1-\gamma})^{n-1}$, with $\lambda = 1/3$. For each $3 < n < 53$ and $\gamma < 2/3$ we perform one simulation of the deterministic $3x$ mod 1 map and compare the obtained extremal index θ_n with the previous asymptotic formula. Results are shown in Figure 7. There is indeed very good agreement between our asymptotic and numerical results. The largest divergence is obtained for $\gamma \simeq 2/3$ which correspond to the limit value for the map.

We then perform a numerical analysis of the extremal index in the cases not covered by the theory, namely for the observable Θ. The results are presented in Figure 8. The top-left panel is repeated for convenience and show the global case results. The latter show that global and local cases are substantially different. For the global, the synchronization depends on both n and γ: in particular, it is easier to synchronize systems with n small because the probability of finding all the particles in the same state decreases quickly with n. On the other hand, in the local case the extremal index θ_n is substantially independent of n. In fact, whether n is small or large, the particle *sees* only the nearest neighbors for synchronization so that it is insensitive to the size of the lattice. The only dependence left is in γ: in particular, for all n, we see the dependence is compatible with the case $n = 2$ of the global coupling case: $\theta_2 \sim 1 - \frac{1}{(1-\gamma)}\frac{1}{3}$. The addition of the noise destroys clusters as observed in [1]. Qualitatively, the structure of the extremal index is quite robust with respect to small perturbations. To fully destroy the clusters, large intensity of the noise are needed. The results for $\varepsilon = 10^{-4}$ also demonstrate that our results are stochastically stable because one recovers the deterministic structure of the extremal index for low noise values.

Although the numerical estimates of the extremal index are done by computing the expected values of the compound Poisson distribution (Süveges' estimator), we can also

check that the waiting times $q_{k,\varsigma}$[17] defined in (5.27), between consecutive entrances in the neighborhood of the diagonal with accuracy ς, provides the same information. Actually this is what we get for recurrence in balls as we discussed above, see [22] and [9]. Therefore we give some examples of time series of ψ and Θ in Figures 9 and 10 respectively. The noise increases from top to bottom. The histograms of the waiting times in cluster are normalized to sum-up to 1 (empirical probability density function EPDF) and are in y-log scale. No clustering corresponds to an exponential law (sequence of linearly decreasing boxes in log scale), whereas the clustering case is characterized by an higher EPDF for lower waiting times. As one can see from the deterministic cases, the higher the EPDF for short waiting times, the lower θ. Effectively the fraction of waiting times equals to 1 which exceed the standard exponential law is exactly the extremal index θ. We stress again that although we cannot demonstrate this identity theoretically, the numerical evidence suggests that one can use directly $q_{k,\varsigma}$ as defined in (5.27), for the estimation of the extremal index θ.

8. Appendices

8.1. **Appendix 1: proof of (2.11).** The argument is the following. The quantity we are interested in is bounded by $\int_{I^n} d\overline{x} |\prod_{i\neq j} \mathbf{1}_{\{|x_i-x_j|<\nu_l+\varepsilon\}}(\overline{x}) - \prod_{i\neq j} \mathbf{1}_{\{|x_i-x_j|<\nu_l-\varepsilon\}}(\overline{x})|$. If at least one factor in the first product is zero, the same is true for the second product, so we will suppose that all the factors in the first product are 1. Therefore the difference of the two products will be maximum if at least one factor in the second product is zero. There will be at most $\sum_{k=1}^{n} \binom{k}{n}$ such possibilities. We now proceed with a very rough bound. Each term in $\binom{k}{n}$, with $1 \leq k \leq (n-1)$ contributes with k measures of values $4^k\varepsilon^k$ and with $(n-k)$ measures of values $4^{n-k}\nu_l^{4-k}$, having chosen $\varepsilon < \nu_l$. When $k = n$ we simply write $\varepsilon^n < \varepsilon^{n-1}\nu_l$. In conclusion, we bound the quantity we are interested in by $\varepsilon\nu_l C_n$, with $C_n = 4^n \sum_{k=1}^{n} \binom{k}{n}$.

8.2. **Appendix 2: proof of (5.30).** Take for simplicity $n = 2$. There is in fact dependence of the two sets on x_1 since they intersect $I^2 \ni (x_2, x_3)$ and as a consequence their measure will depend on the location of x_1. It will therefore be enough to evaluate the external integrals in x_1 on a even smaller domain $I''_m \subset I'_m$ and on I'''_m in the denominator, in such a way they will not contain a (disconnected) neighborhood $\mathcal{U}$ of 0 and 1 and its preimages $T^{-1}\mathcal{U}$. As a consequence, we can keep the full amount of the area of the two sets $S^{(2)}_{m,\gamma}(Tx_1)$ and $S^{(2)}_m(x_1)$, which from now on we simply write as $\mathrm{Leb}(S^{(2)}_{m,\gamma})$ and $\mathrm{Leb}(S^{(2)}_m)$. Clearly the difference between the integrals over I and I''_m, I'''_m will converge again to zero when $m \to \infty$. About the other issue: write $S^{(n)}_{m,\gamma}(Tx_1)$ as the integral

[17] We now index this quantity with the size $\varsigma \to 0$ of the neighborhood of the diagonal.

of obvious characteristic functions in the variables $x_2, \ldots, x_n$. Then make the change of variables: $x'_k = x_k(1-\gamma) + \gamma T(x_1)$, in this way we get the measure of $S_m^{(n)}(x_1)$ multiplied by $(1-\gamma)^{1-n}$.

Acknowledgments

SV and PG were supported by the MATH AM-Sud Project Physeco. SV was supported by the Leverhulme Trust for support thorough the Network Grant IN-2014-021 and by the project APEX Systèmes dynamiques: Probabilités et Approximation Diophantienne PAD funded by the Réegion PACA (France). DF was partially supported by the ERC grant A2C2 (No. 338965). PG thanks FONDECYT project 1171427. SV warmly thanks J. M. Freitas, P. Giulietti, N. Haydn and B. Saussol for illuminating discussions. We finally thank the referees for the careful reading of the paper which helped us to greatly improve it.

References

[1] H. Aytach, J.M. Freitas, S. Vaienti, *Laws of rare events for deterministic and random dynamical systems*, Trans. Amer Math. Soc., 36 , 8229-8278, 2015.

[2] P. Ashwin, *Riddled Basis and Coupled Dynamical System*, in [4].

[3] L.A. Bunimovich, Ya.G. Sinai, *Space-time chaos in coupled map lattices*, Nonlinearity, 1, 491-519, 1988.

[4] J-R. Chazottes, B. Fernandez Editors, *Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems*, Volume 671 of the series Lecture Notes in Physics pp 265-284, 2005.

[5] Z. Coelho, P. Collet,Asymptotic limit law for the close approach of two trajectories in expanding maps of the circle, Probability Theory and Related Fields, 99, 237-250, 1994.

[6] J.P. Crutchfield and K. Kaneko, Phenomenology of spatiotemporal chaos (in Directions in Chaos, ed Hao Bai Lin) World Scientific, Singapore, 1987.

[7] D. Azevedo, A.C.M. Freitas, J.M. Freitas, F.B. Rodrigues, *Extreme Value Laws for dynamical systems with countable extremal sets*, J. Stat. Phys., 167, no. 5, 1244-1261, 2017.

[8] A. C.M,Freitas, J. M. Freitas and M. Todd, *The extremal index, hitting time statistis and periodicity,* Adv. Math., 231(5): 2626-2665, 2012.

[9] 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., pages 1-45, 2013.

[10] K. Kaneko ed. Theory and Applications of Coupled Map Lattices Wiley, Chichester (1993).

[11] G. Keller, *Generalized bounded variation and applications to piecewise monotonic transformations*, Z. Wahr. verw. Geb. **69** 461-478 (1985).

[12] G. Keller, *Rare events, exponential hitting times and extremal indices via spectral perturbation*, Dyn. Syst. 27, no. 1, 1127 2012.

[13] G. Keller, C. Liverani, *Rare events, escape rates and quasistationarity: some exact formulae,* J. Stat. Phys. 135 (2009), no. 3, 519534.

[14] G. Keller, C. Liverani, *Stability of the spectrum for transfer operators,* Ann. Scuola Norm. Sup. Pisa Cl. Sci.(4), 28:141-152, 1999.

[15] G. Keller, C. Liverani *A spectral gap for a one-dimensional lattice of coupled piecewise expanding interval maps*,in "Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems", J.Chazottes and B.Fernandez Eds., Lecture Notes in Physics, Vol. 671, pp. 115-151, Springer (2005).

[16] G. Keller and C. Liverani, Uniqueness of the SRB measure for piecewise expanding weakly coupled map lattices in any dimension, Commun. Math. Phys. 262 (2006) 33?50.

[17] M. Hirata, B. Saussol, S. Vaienti, *Statistics of return times: A general framework and new applications*, Communication in Mathematical Physics, 206, 33-55, 1999.

[18] N.Haydn, S. Vaienti, *The limiting distribution and error terms for return time of dynamical systems*, Discrete and Continuous Dynamical Systems A, 10, 584-616, 2004.

[19] N. Haydn, S. Vaienti, *The compound Poisson distribution and return times in dynamical systems*, Probability Theory and Related Fields, 144, 517-542, 2009.

[20] N. Haydn, M. Nicol, A. Török, S. Vaienti, *Almost sure invariance principle for sequential and non-stationary dynamical systems*, Trans. Amer Math. Soc, **369**,p. 5293-5316 (2017).

[21] H. Hennion and L. Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, **1766**; Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2001.

[22] H. Hu, S. Vaienti, *Lower Bounds for the Decay of Correlations in Non-uniformly Expanding Maps*, to appear on Ergodic Theory and Dynamical Systems.

[23] K. Kaneko,*On the Strength of Attractors in a High-dimensional System: Milnor Attractor Network, Robust Global Attraction, and Noise-induced Selection*, Physica D, 124, 322-344, 1998

[24] M. R. Leadbette G. Lindgren and H Rootzn, *Extremes and related properties of random sequences and processes,* Springer Series in Statistics. Springer-Verlag, New York, 1983.

[25] V. Lucarini, D. Faranda, A. M. Freitas, J. M. Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd, S. Vaienti, *Extremes and Recurrence in Dynamical Systems,* Wiley, New York, 2016, ISBN: 978-1-118-63219-2

[26] Y. Sul Cho, T. Nishikawa, A. E. Motter, *Stable Chimeras and Independently Synchronizable Clusters*, to appear in Phys. Rev. Lett., https://arxiv.org/pdf/1707.06657.pdf.

[27] B. Saussol, *Absolutely continuous invariant measures for multidimensional expanding maps,* Israel J. Math. 116: 223248, 2000,

[28] M. Tsujii, *Absolutely continuous invariant measures for expanding piecewise linear maps*, Inventiones mathematicae, **143**, 349-373 (2001).

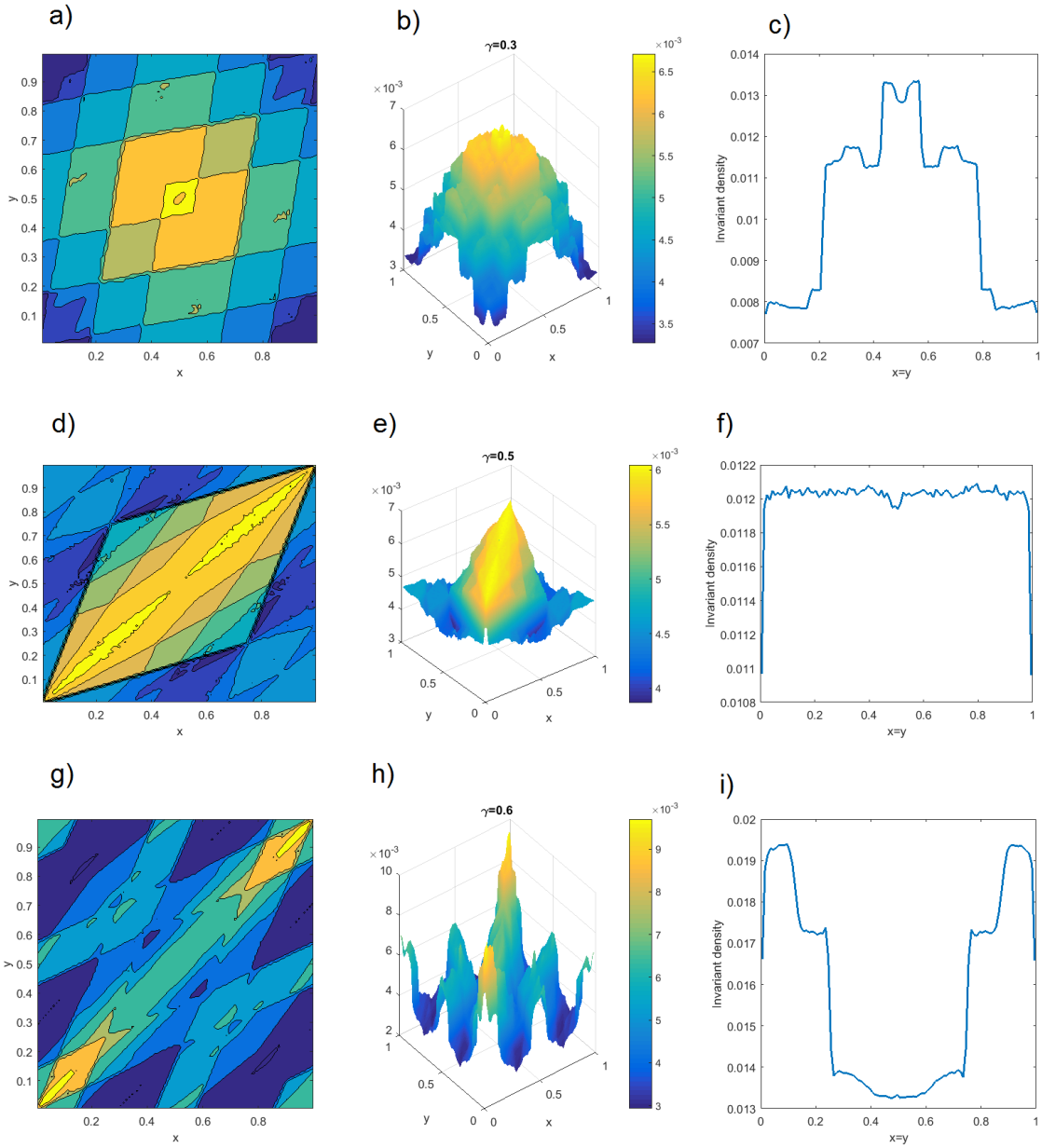

FIGURE 1. Invariant density for the map 2.2 with $n = 2$ for $\gamma = 0.3$ (a,b,c), $\gamma = 0.5$ (d,e,f), $\gamma = 0.6$ (g,h,i). The plots show the density in colorscale (a,d,g) with a view from the top and (b,e,h) for a three dimensional view. The plots (c,f,i) show the behavior of the map on the diagonal. The figure is obtained by averaging the density over 300 realizations each consisting of 10^7 iterations of the trajectory.

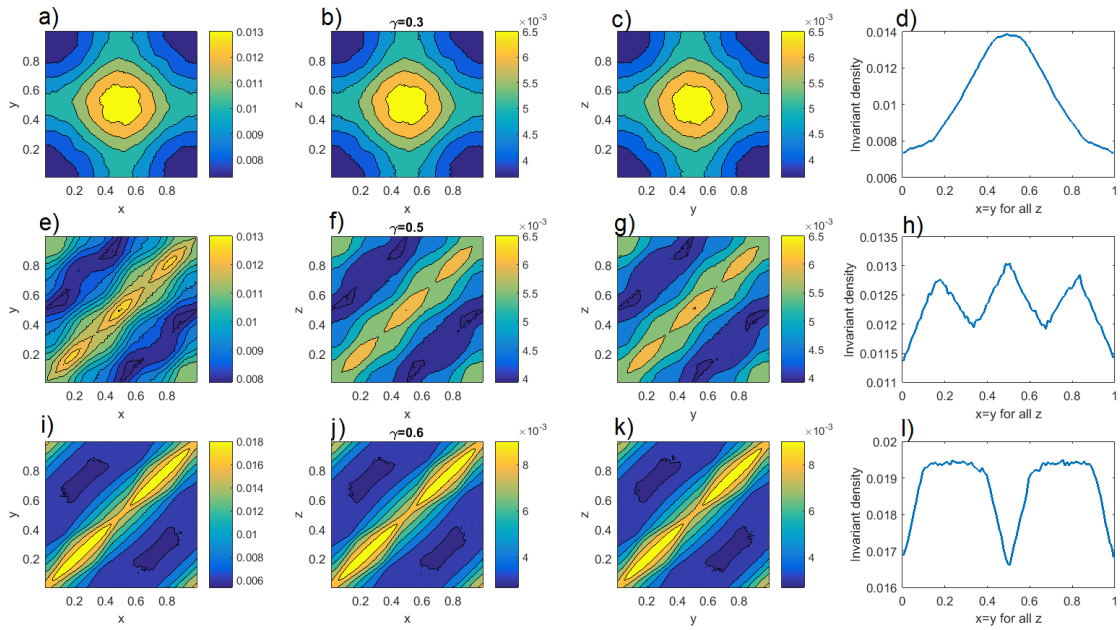

FIGURE 2. Invariant density for the map 2.2 with $n = 3$ for $\gamma = 0.3$ (a,b,c,d), $\gamma = 0.5$ (e,f,g,h), $\gamma = 0.6$ (i,j,k,l). The plots show the density in colorscale (a-c,e-g,i-k). The plots (d,h,l) show the behavior of the map on the diagonal $x = y$. The figure is obtained by averaging the density over 300 realizations each consisting of 10^6 iterations of the trajectory.

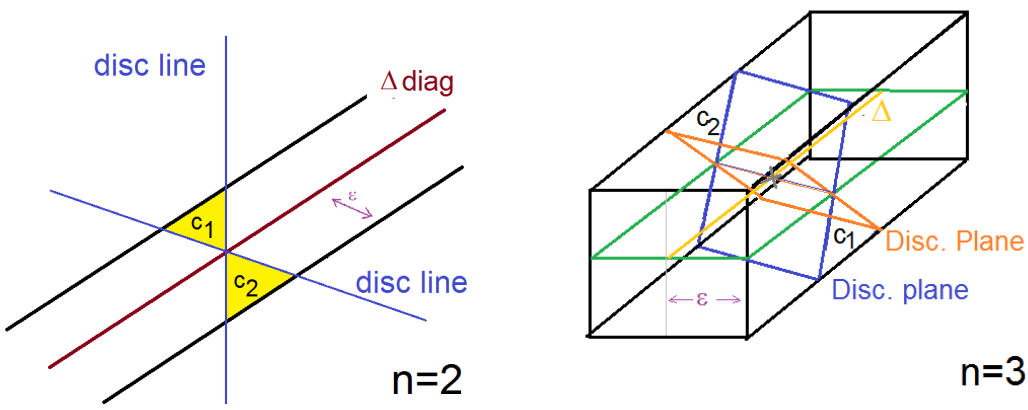

FIGURE 3. Crossing of the discontinuity line, $n = 2$ (left) and surface, $n = 3$ (right), of the neighborhood of the diagonal Δ, for Property **P01**. C_1 and C_2: triangular (left) and pyramidal (right) regions belonging respectively to $F^c_{d,\varepsilon,2}$ and $F^c_{d,\varepsilon,3}$. We remove the shaded regions on the left and the pyramidal regions C_1 and C_2 on the right.

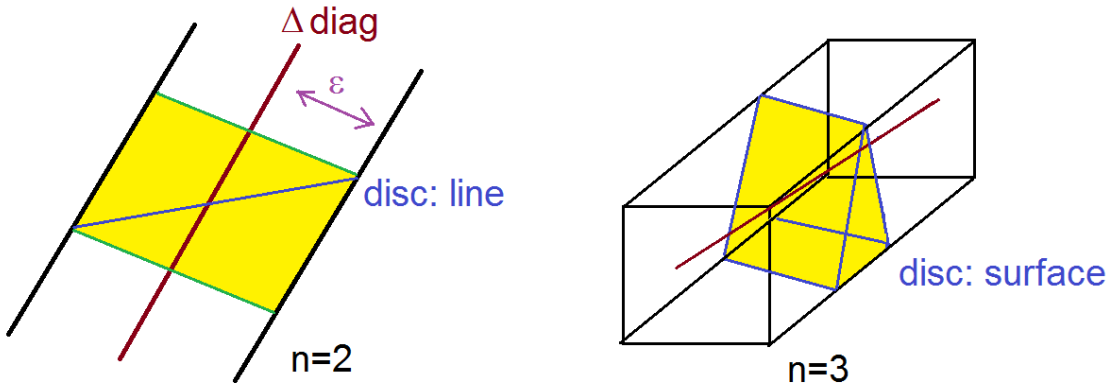

FIGURE 4. Crossing of the discontinuity line, $n = 2$ (left) and surface, $n = 3$ (right), of the neighborhood of the diagonal Δ, for Property **P02**. We remove the shaded regions.

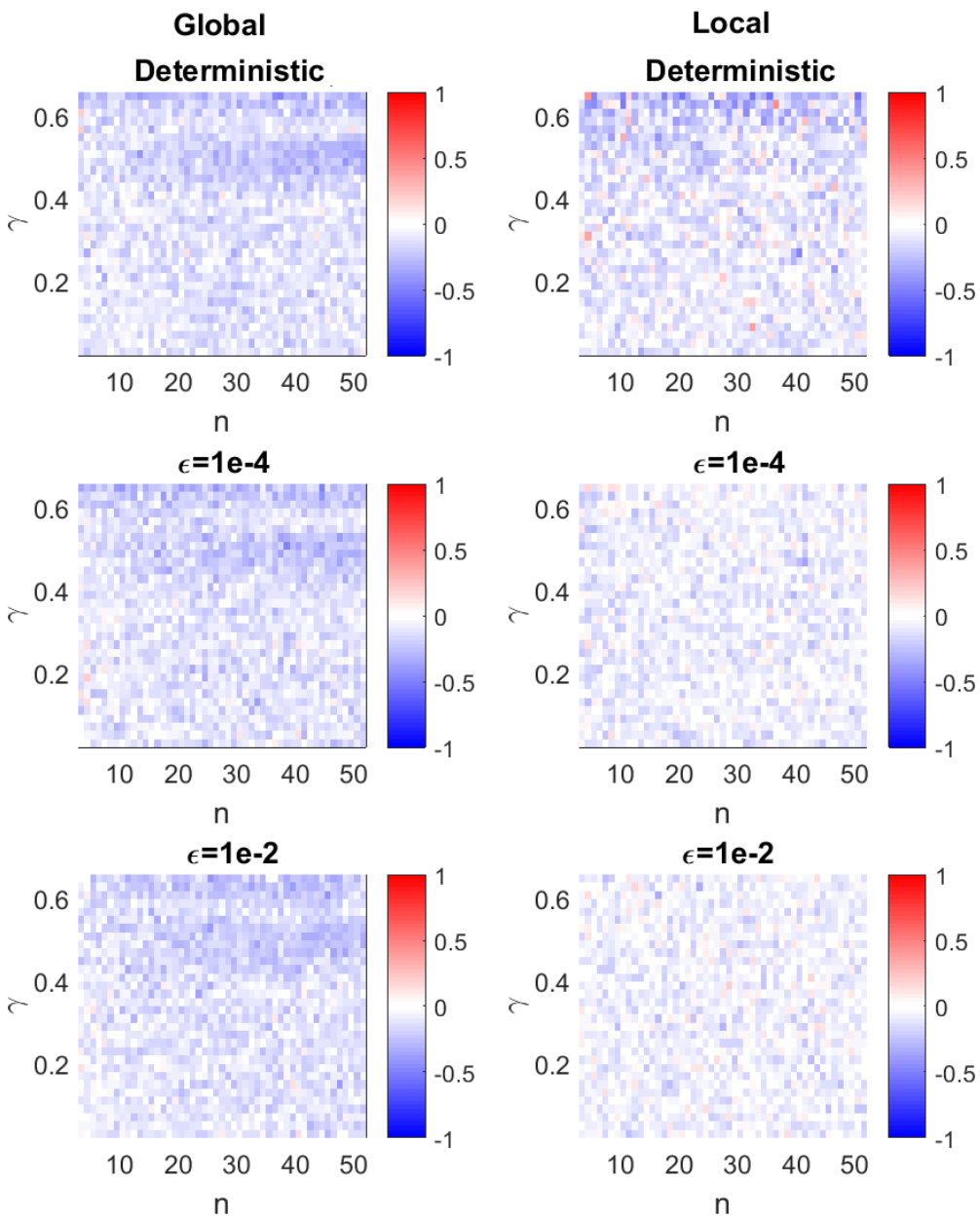

FIGURE 5. Shape parameter ξ of the Generalized Pareto distribution as a function of the number of variables n and the coupling parameter γ. Left: global case ψ. Right: local case Θ. From top to bottom: Deterministic, additive noise with intensity $\varepsilon = 10^{-4}$, additive noise with intensity $\varepsilon = 10^{-2}$.

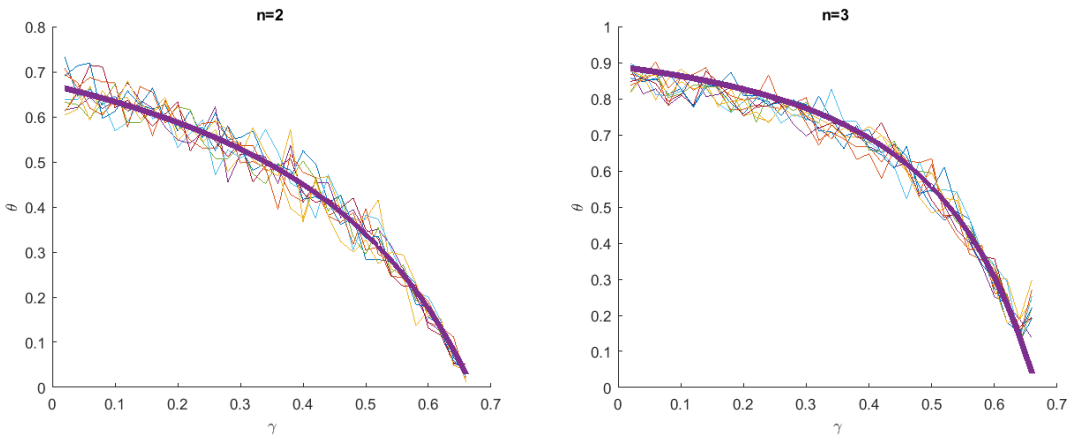

FIGURE 6. Extremal index θ of the Generalized Pareto distribution as a function of the coupling parameter γ. Thin lines indicate estimates for 10 different realization of the maps $3x$-mod1. Bold magenta lines indicate the expected theoretical values. Left: $n = 2$, Right: $n = 3$.

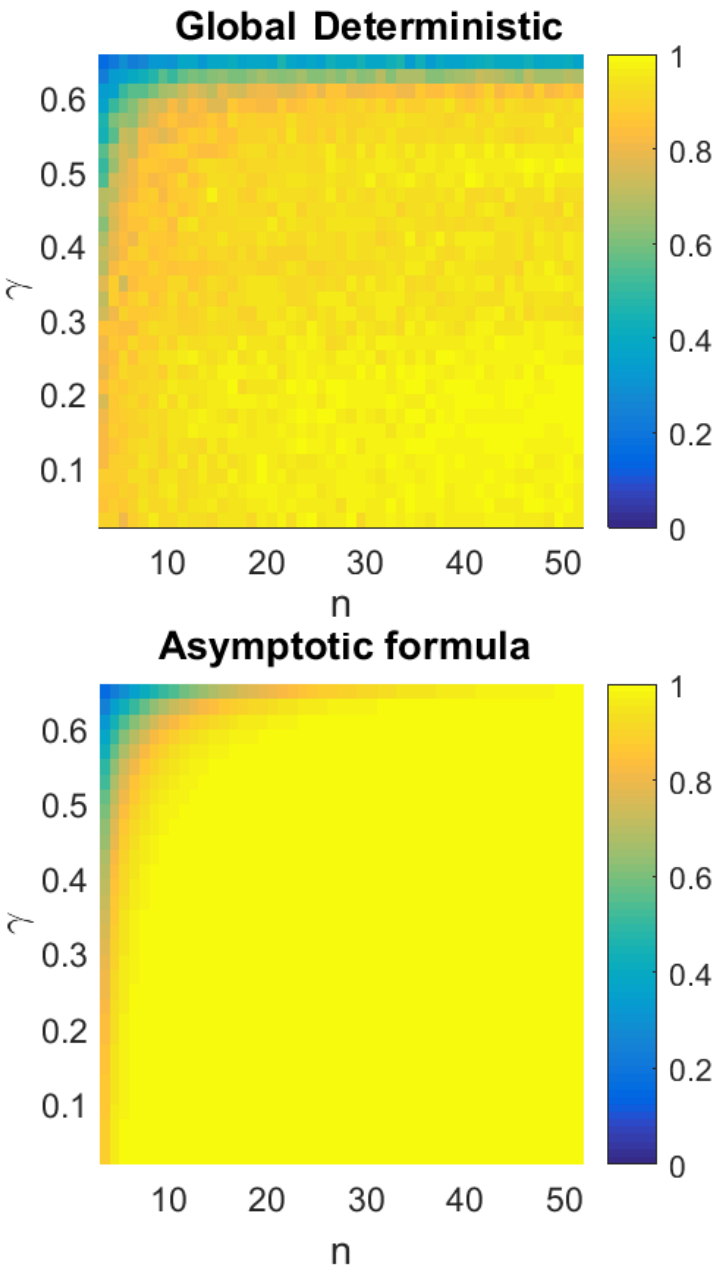

FIGURE 7. Extremal index θ as a function of the number of variables n and the coupling parameter γ. Top: global case ψ. Bottom: theoretical asymptotic formula.

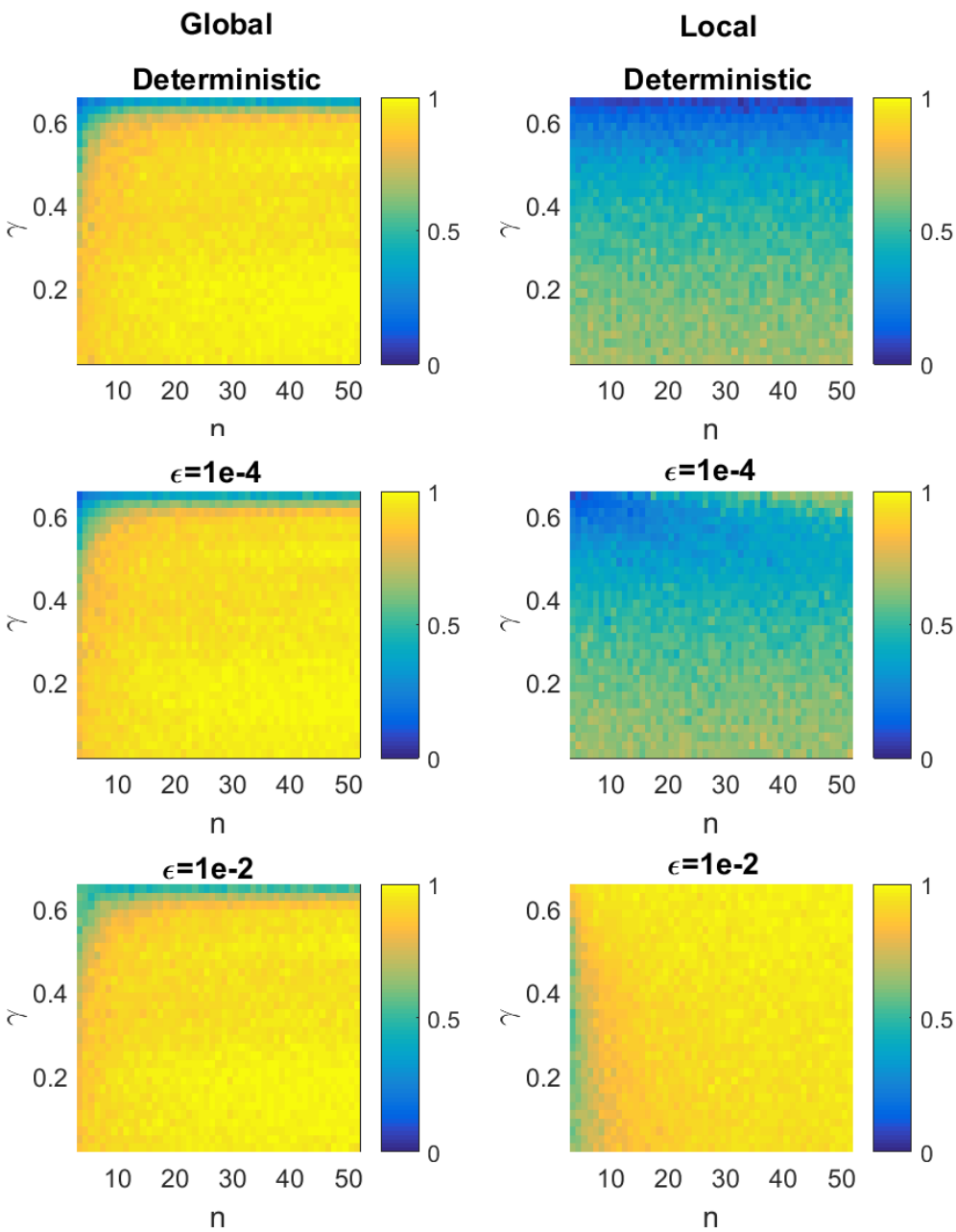

FIGURE 8. Extremal index θ as a function of the number of variables n and the coupling parameter γ. Left: global case ψ. Right: local case Θ. From top to bottom: Deterministic, additive noise with intensity $\varepsilon = 10^{-4}$, additive noise with intensity $\varepsilon = 10^{-2}$.

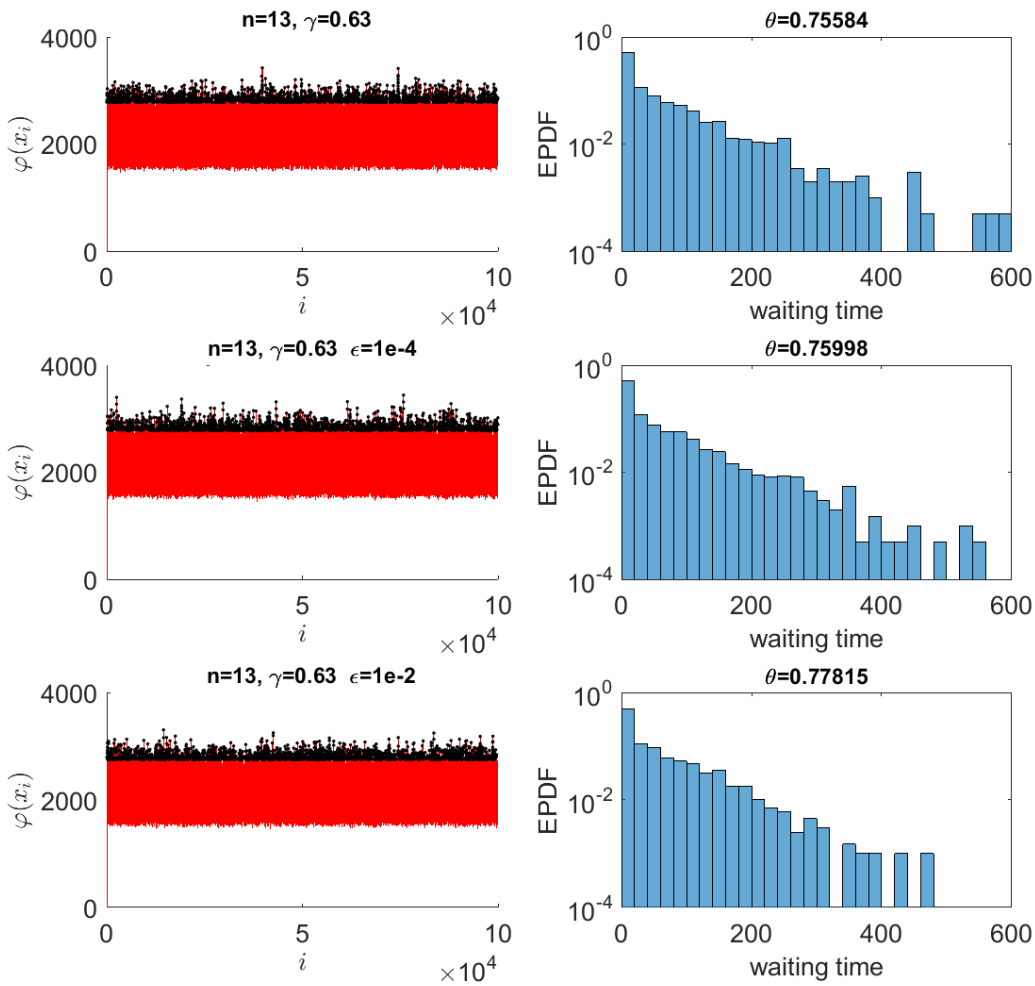

FIGURE 9. Example of global case. Right: ψ time series (red) and exceedances (black). Left: empirical probability distribution (EPDF) of waiting time in the clusters. From top to bottom: Deterministic, additive noise with intensity $\varepsilon = 10^{-4}$, additive noise with intensity $\varepsilon = 10^{-2}$.

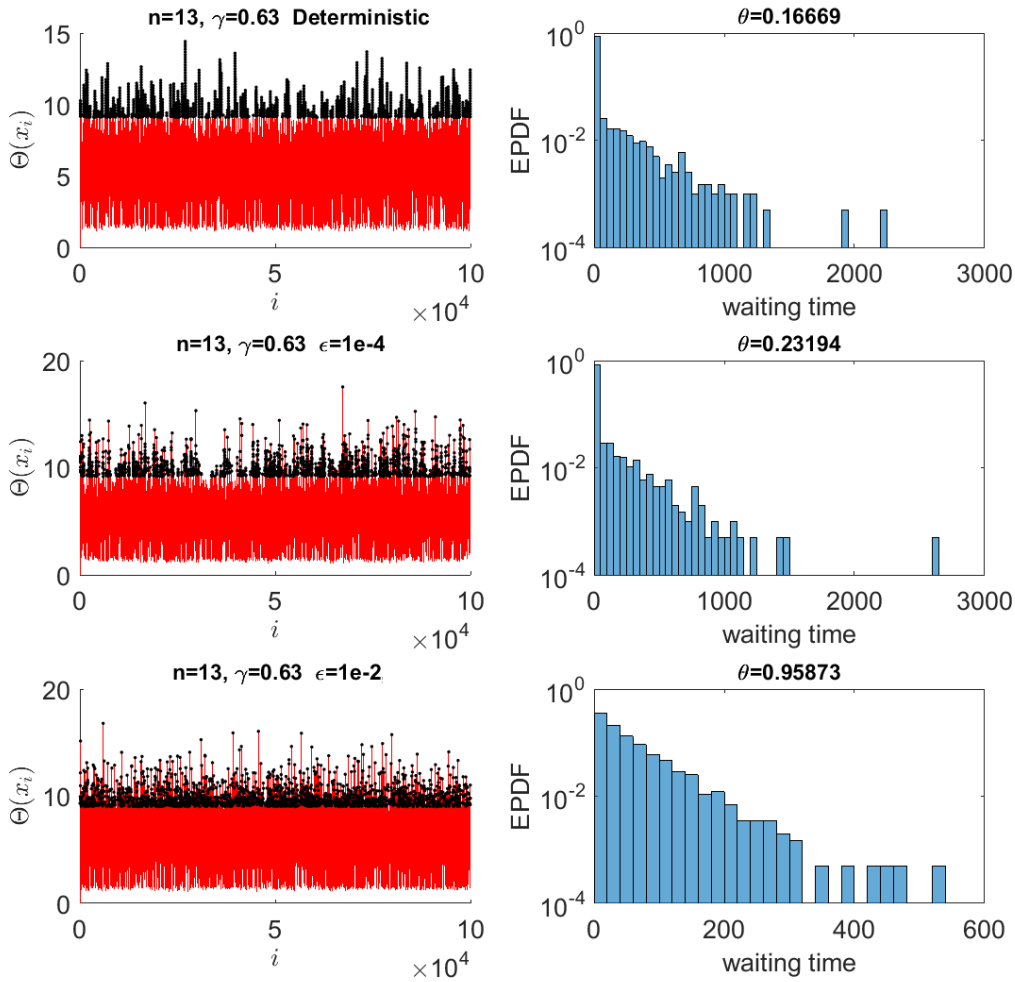

FIGURE 10. Example of local case. Right: Θ time series (red) and exceedances (black). Left: empirical probability distribution (EPDF) of waiting time in the clusters. From top to bottom: Deterministic, additive noise with intensity $\varepsilon = 10^{-4}$, additive noise with intensity $\varepsilon = 10^{-2}$.