HAL
open science

Extreme Value theory and Poisson statistics for discrete time samplings of stochastic differential equations

Franco Flandoli, Stefano Galatolo, Paolo Giulietti, Sandro Vaienti

To cite this version:

Franco Flandoli, Stefano Galatolo, Paolo Giulietti, Sandro Vaienti. Extreme Value theory and Poisson statistics for discrete time samplings of stochastic differential equations. 2023. hal-04389263

HAL Id: hal-04389263

https://hal.science/hal-04389263v1

Preprint submitted on 11 Jan 2024

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:2310.13972v1 [math.DS] 21 Oct 2023

Extreme Value theory and Poisson statistics for discrete time samplings of stochastic differential equations.

F. Flandoli* S. Galatolo† P. Giulietti‡ S. Vaienti§

October 24, 2023

Abstract

We investigate the distribution and multiple occurrences of extreme events stochastic processes constructed by sampling the solution of a Stochastic Differential Equation on $\mathbb{R}^n$. We do so by studying the action of an annealead transfer operators on ad-hoc spaces of probability densities. The spectral properties of such operators are obtained by employing a mixture of techniques coming from SDE theory and a functional analytic approach to dynamical systems.

Keywords Stochastic differential equations, extreme value theory, transfer operator, regularization by noise, perturbative spectral theory

MSC2010 MSC2020 Update : Primary/Secondary 37A50, 37H05, 60G70, 60H50

Contents

1 Introduction **2**
1.1 Literature Review 3

2 Main Results **4**
2.1 Organization of the paper 7

3 Transfer operators, Banach spaces and regularization lemmas. **7**
3.1 The Kolmogorov operator and the transfer operator 8
3.2 Perturbed operators 9

*Scuola Normale Superiore. Email:franco.flandoli@sns.it

†Dipartimento di Matematica - Università di Pisa. Email: stefano.galatolo@unipi.it

‡Dipartimento di Matematica - Università di Pisa. Email:paolo.giulietti@unipi.it orcid:0000-0001-9604-1699

§Aix Marseille Université, Université de Toulon, CNRS, CPT, 13009 Marseille, France. Email: vaienti@cpt.univ-mrs.fr.

3.3 Functional Spaces: quasi-Hölder space on $\mathbb{R}^d$ 11
3.4 Regularization properties for the transfer operator. 13
3.5 Regularization for the perturbed operators 18

4 Rare Events Via Transfer Operator **20**
4.0.1 Perturbative hypotheses 20
4.0.2 Sufficient conditions to check assumptions R1 and R2 . . 21
4.1 Verifying the perturbative assumptions in our case 22
4.2 Proof of Theorem 2 25

5 Poisson statistics **27**

1 Introduction

In the last decades modeling based on Stochastic Differential Equations has been used extensively in different scenarios whenever the statistical properties of extreme events is particularly important, ranging from economy (see e.g. [16]) to climate science (see e.g. [10]). In this case the SDE model becomes a forecasting tool: one can investigate the likelihood of extreme events, as well as the typical behaviour of such systems. Here we focus on the former.

We consider a system whose evolution is described by a stochastic differential equation of the following type

$$\begin{cases} dX_t = b\left(X_t\right) dt + dW_t \\ X_0 = x, \end{cases}$$

where the domain of definition is $\mathbb{R}^n$ and W_t is the Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. It is well known that under suitable assumptions (see Section 2 for the details of our setup) the problem above has a unique solution $(X_t^x)_{t\geq 0}$ and the resulting stochastic flow has a unique stationary measure μ.

We consider a specific unbounded observable $g_{x_0} : \mathbb{R}^d \to \mathbb{R}$, encoding the appearance of an extreme event, and sample the evolution of g_{x_0} along the trajectories of X_t at discrete times. This is done by choosing a time step $h \geq 0$, a sequence of times $t_n = hn$ and the process $g_{x_0}(X_{t_n}^x)$.

It is well known that pinpointing a single extreme event is still, if possible at all, beyond the reach of current techniques, yet there are various possibilities to investigate the distribution, either temporal or spatial, of such events also in the sense of their clusterization. Here we consider the distribution of the extreme events for the process $g_{x_0}(X_{t_n}^x)$; in particular we would like to find a sequence $\{u_n\}_{n\in\mathbb{N}}$ such that the following limit exists

$$\lim_{n\to\infty} \mathbb{P} \otimes \mu \left(\left\{ (\omega, x) \ : \ \max_{k=0,\ldots,n-1} g_{x_0}\left(X_{t_k}^x\right) \leq u_n \right\} \right) \in (0,1) \tag{1}$$

where the $\otimes$ represents the product of measures. The limit above encodes the notion of extreme event, as long as $u_n \to \infty$ for the process $g_{x_0}(X_{t_n}^x)$ where

the initial conditions x vary in the full set $\mathbb{R}^d$ and are weighted according to μ. Theorem 2 solves this problem i.e. it is possible to characterize such sequences and find an exponential distribution for the probability of not exceeding the thresholds u_n, as it is usually done in the extreme values theory.

In the literature, such limit is often referred to as Extreme Value Law (EVL); its existence and its properties are strictly connected to the distribution of the hitting time of the process to a sequence of (properly renormalized in measure) shrinking sets, see [8] for an extensive introduction to such relation.

Successively, we will consider the distribution of the occurrences of our event in the rare set and in a given time interval. More precisely we will count the number of visits of the time discretized solution of our stochastic differential equation to a decreasing sequence of balls B_n among the first n steps of length h. We will show that the distribution of the number of visits will converge to a standard Poisson distribution in the limit of large n. Our main results will be stated precisely in Section 2.

It is worth to remark that for both results our construction is rooted in the functional approach of [14, 15, 12]. The tools we use to get the results are related to the study of the functional analytic properties of the transfer operator L_h associated to the SDE system acting on suitable functional spaces. In particular we will relate the probability of occurrence of the rare events with the response of the dominating eigenvalue of L_h to the perturbation of the system done by adding a small hole corresponding to the rare event, transforming the initial system into an open one.

1.1 Literature Review

The study of extreme events has already been carried out in various framework and there is an extensive literature. Comprehensive surveys can be found in the book [18], with a particular focus on dynamical systems and in [19] with a focus on stochastic processes.

Let us comment on some works which are closer to the present one which allow us also to present the main novelties of our study.

In the paper [6], an Extreme Values Law in continuous time is found in the case of one dimensional stochastic differential equations, using the properties of the Ornstein-Uhlenbeck process (see also [19] for another point of view on these results).

For diffusion processes of gradient field type in $\mathbb{R}^d$, the Ph.D thesis results of Kuntz [17, Prop. 2.1, Theorem 2.4], prove exponential upper/lower bounds for $P(M_T \leq R)$, for large R, which can be turned into Extreme Values Law estimates.

In the paper [20] the hitting time distribution for random walks in the line is described. This kind of results are equivalent to EVL as proved in [8].

Most of the preceding works obtained an extreme value distributions for continuous time processes beginning with particular SDE, while in the present work we focus on discrete time but for a rather general class of Stochastic Differential Equations. In this case we show not only how the distribution of the first visit

in a rare event behave, but also the distribution of multiple visits (the Poisson statistics). We will comment more about the relation between continuous and discrete time at the end of this article by outlining a few perspectives.

We also like to reiterate the fact that our approach is directly based on the spectral properties of the transfer operators associated to the system and its response to suitable small perturbations representing the rare events. This technique, as already pointed out in [15], allows us to get EVL for deterministic dynamical systems whose transfer operator is quasi-compact (on a suitable Banach space), and therefore admits a spectral gap. The same technique has been generalized to random dynamical systems perturbed in an *annealed* way in [2, 4] and the recent paper [1] developed a spectral approach for a *quenched* extreme value theory that considers random dynamics and random observations.

2 Main Results

We consider a stochastic differential equation on $\mathbb{R}^d$ of the following type

$$\begin{cases} dX_t = b\left(X_t\right) dt + dW_t \\ X_0 = x, \end{cases} \tag{2}$$

Here W_t is the Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We make the following assumptions on b:

H There is K such that $\forall x, y \in \mathbb{R}^d$

$$|b(x) - b(y)| \leq K|x - y|.$$

AD (dissipativity) We assume that there exists constants $L_1, L_2 \in \mathbb{R}$ with $L_2 > 0$ such that for all $x \in \mathbb{R}^d$

$$\langle b(x), x\rangle \leq L_1 - L_2\|x\|^2. \tag{3}$$

Remark 1. *The previous assumptions include the fundamental case when*

$$b(x) = Ax + B(x)$$

where A is a matrix such that there exists $\nu > 0$ with the property that

$$\langle Ax, x\rangle \leq -\nu \left\|x\right\|^2 \tag{4}$$

for all $x \in \mathbb{R}^d$, and $B : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz continuous and fulfills

$$\langle B(x), x\rangle = 0 \tag{5}$$

or more generally

$$\langle B\left(x\right), x\rangle \leq C_1 + C_2 \left\|x\right\|^2 \qquad \textit{for all } x \in \mathbb{R}^d \tag{6}$$

*for some constants $C_1 \in \mathbb{R}$ and $C_2 < \nu$. Indeed assumption **H** is obviously satisfied and assumption **AD** holds because*

$$\langle Ax + B(x), x\rangle \leq C_1 - (\nu - C_2)\,\|x\|^2 .$$

Condition (4) is motivated for instance by the finite dimensional discretization of Partial Differential Equations (PDE henceforth) of parabolic type, where A is the discretization of the Laplacian or, more generally, of the second order strongly elliptic part. Condition (5) is motivated by the discretization of nonlinear operators like the inertial (convective) operator of the Navier-Stokes equations; we impose the Lipschitz continuity on B for simplicity. [1] *The generalization of Condition (6) may help to accommodate linear first order operators in the discretization of a PDE.*

Let $x_0 \in \mathbb{R}^d$ be a chosen point of the phase space. Let $g_{x_0} : \mathbb{R}^d \setminus \{x_0\} \to \mathbb{R}$ be

$$g_{x_0}(x) := -\log d(x, x_0) \tag{7}$$

where d is the euclidean distance.

The observable g_{x_0} hence measures how far we are from our chosen x_0 on a logarithmic scale. Let $h > 0$ and let

$$t_n := nh. \tag{8}$$

Given the assumptions **H** and **AD**, the system (2) has a unique invariant measure μ (see Section 4.1 for more details).

We now consider the problem (1), and about this problem for the reader convenience we state below the first main result of our work.

Theorem 2. *Let $h, \tau > 0$ and let X_t be the solution of (2) at time t. Let u_n be a real sequence such that*

$$n\,\mu(B_n) \to \tau \tag{9}$$

where B_n is the ball $B\left(x_0, e^{-u_n}\right)$. Let the sequence of discrete times in which we sample the process be denoted as $t_k := kh$. Then

$$\lim_{n\to\infty} \mathbb{P} \otimes \mu \left(\left\{ (\omega, x) \;:\; \max_{k=0,\dots,n-1} g_{x_0}\left(X^x_{t_k}\right) \leq u_n \right\} \right) \to e^{-\tau}. \tag{10}$$

Remark 3. *Analogous results holds if the system (2) is defined on $\mathbb{T}^d$. As many of the constructions presented here are either not necessary or can be simplified in such case, the explicit discussion of this case will appear somewhere else.*

The second main result we prove is a refinement of the first one and is about the distribution of multiple occurrences of the extreme events. Let B_n be again a ball centered at $z \in \mathbb{R}^d$ and of radius e^{-u_n}. We are now interested in studying the distribution of the number of visits to the set B_n in a prescribed time interval.

[1] This is not satisfied by quadratic operators, therefore an additional Lipschitz cut-off is necessary to get a finite dimensional operator as above.

We follow the strategy recently used in [12] for this kind of results. We argue that the exponential law given by the extreme value distribution describes the time between successive events in a Poisson process. We begin by introducing the following random variable

$$S_n := \sum_{i=0}^{n-1} 1_{B_n}(X_{ih}^x),$$

which counts the number of visits of the time sampled solution of our stochastic differential equation to the ball B_n among the first n iterations of the process. In order to get a limiting distribution when $n \to \infty$ we have to rescale time as we did in (9) and therefore we pose:

Definition 4. *Let us take $\tau > 0$ and $n \geq 1$. We define the sequence of discrete random variables*

$$S_{n,\tau} := \sum_{i=0}^{\lfloor \frac{\tau}{\mu(B_n)} \rfloor} 1_{B_n}(X_{ih}^x). \tag{11}$$

We say that $S_{n,\tau}$ converge in distribution to the discrete random variable W, possibly defined on a different probability space and with distribution ν_W, if we have for any $k \in \mathbb{N}$:

$$\lim_{n\to\infty} \mathbb{P} \otimes \mu \left(\sum_{i=0}^{\lfloor \frac{\tau}{\mu(B_n)} \rfloor} 1_{B_n}(X_{ih}^x) = k \right) = \nu_W(\{k\}). \tag{12}$$

The second main result of this work is the following result establishing the convergence toward a standard Poisson distribution:

Theorem 5. *Let $\tau > 0$, X_t be the solution of (2) on $\mathbb{R}^d$, let B_n the ball $B(x_0, \exp(-u_n))$ and let u_n such that $n\,\mu(B_n) \to \tau$. Then*

$$\lim_{n\to\infty} \mathbb{P} \otimes \mu \left(\sum_{i=0}^{\lfloor \frac{\tau}{\mu(B_n)} \rfloor} 1_{B_n}(X_{ih}^x) = k \right) = \frac{e^{-\tau}\tau^k}{k!}.$$

As mentioned before, the proof of Theorems 2 and 5 relies on the spectral perturbation results of [14] and [15]; we recall the main results we are going to use for completeness in Proposition 24. Let us say once more, at this stage, that the general strategy we follow is to relate the distribution of rare events we want to estimate to the behavior of the leading eigenvalues of a perturbed transfer operator associated to the system, where the (small) perturbations which are applied, encode the rare events we want to count (see Sections 4.2, 5).

To apply this strategy we consider the transfer operator L_h associated to the evolution of the system for a fixed time $h > 0$.

We find Banach spaces such that L_h, restricted to these spaces, satisfies a Lasota-Yorke type inequality which in turns, if the spaces also embeds compactly, implies a spectral gap. A key inequality will be obtained by exploiting

the regularizing effect of the noise in the stochastic differential equation and the presence of the dissipative assumption (AD). Since the ambient space is $\mathbb{R}^d$ i.e. a non compact space, we introduce BV_α: spaces of bounded variation densities decaying at infinity with a certain power law of exponent α.

The reason for considering spaces akin to bounded variation spaces is the representation of rare events, which are roughly modeled here as small holes in the space $\mathbb{R}^d$: the perturbed (by the hole) transfer operator will necessarily create discontinuities when acting on densities.

Once collected all the necessary estimates, the proofs of Theorem 2 and Theorem 5 are spelled out respectively in Section 4.2 and Section 5.

2.1 Organization of the paper

The plan of the paper is as follows: in Section 3 we set up the functional analytic framework necessary to study our problems and we perform the regularization estimates.

Section 4 recalls the main abstract tool we use: Proposition 24.

Section 4.2 shows how from Proposition 24 we can recover the extreme event laws stated in Theorem 2.

Section 5 shows how from Proposition 24 we can recover the Poisson statistics stated in Theorem 5.

3 Transfer operators, Banach spaces and regularization lemmas.

We first define and study the basic properties of the transfer operators associated to our system in sections 3.1 and 3.2. In section 3.3, we introduce a weighted L^1 space and bounded variation spaces which we will use to carry out the necessary estimates. In sections 3.4 and 3.5 we study the regularization properties of the transfer operators when applied to these suitable functional spaces.

Recall the SDE (2) on the euclidean space $\mathbb{R}^d$, assuming the regularity assumption (H) and the dissipative assumption (AD).

For this kind of stochastic differential equations it is known (see [9], [23]) that the equation has the properties of strong existence of the solutions and pathwise uniqueness.

Moreover, Menozzi, Pesce and Zhang [21] prove bounds on the transition probabilities for these systems (so called Aronson type estimate) in this setup. Such estimates imply that the transfer operator associated to the system has a regular kernel, and hence regularizing properties. Let θ_t, for $t \geq 0$ be the flow solving

$$\begin{cases} \dot{\theta}_t(x) = b(\theta_t(x)) \\ \theta_0(x) = x. \end{cases}$$

for a function b of the system considered. Let $\lambda \in (0,1], t > 0$ and g_λ be the

Gaussian distribution

$$g_\lambda(t,x) := t^{-\frac{d}{2}} e^{\frac{-\lambda|x|^2}{t}}.$$

Theorem 6 ([21], Theorem 1.2 and Remark 1.3.)**.** *Let us fix $T > 0$. For each $0 < t \leq T$ and $x \in \mathbb{R}^d$ let $X_t(x)$ be the unique solution of (2) starting from x at time t. Then $X_t(x)$ has a density which for each $y \in \mathbb{R}^d$ can be expressed as a function $S_t(x,y)$ which is continuous in both variables. Moreover S_t satisfies the following:*

1 *(Two sided density bounds) There exist $\lambda_0 \in (0,1], C_0 \geq 1$ depending on T, k, d such that for any $x, y \in \mathbb{R}^d,\ t < T$*

$$C_0^{-1} g_{\lambda_0^{-1}}(t, \theta_t(x) - y) \leq S_t(x,y) \leq C_0 g_{\lambda_0}(t, \theta_t(x) - y).$$

2 *(Gradient estimates) There exist $\lambda_1 \in (0,1],\ C_1 \geq 1$ depending on T, k, d such that for any $x, y \in \mathbb{R}^d,\ t < T$*

$$|\nabla_x S_t(x,y)| \leq C_1 t^{-\frac{1}{2}} g_{\lambda_1}(t, \theta_t(x) - y),$$

$$|\nabla_y S_t(x,y)| \leq C_1 t^{-\frac{1}{2}} g_{\lambda_1}(t, \theta_t(x) - y).$$

3.1 The Kolmogorov operator and the transfer operator

In this section we define the transfer operators associated to the evolution of a SDE and show some of the basic properties of these operators. The properties of the transition probabilities S_t inherited by [21] will be used to define a Kolmogorov operator (composition operator) associated to our system.

Definition 7. *The Kolmogorov operator $P_t : L^\infty(\mathbb{R}^d) \to C^0(\mathbb{R}^d)$ associated to the system (2) is defined as as follows. Let $\phi \in L^\infty(\mathbb{R}^d)$ then $\forall x \in \mathbb{R}^d$*

$$(P_t\phi)(x) := \mathbb{E}[\phi(X_t(x))].$$

In the literature this is also known as stochastic Koopman operator. Since in our case X_t^x has density $S_t(x,y)$ absolutely continuous w.r.t to Lebesgue, we let

$$(P_t\phi)(x) = \int_{\mathbb{R}^d} \phi(y) S_t(x,y) dy.$$

By this we see that

$$\|P_t\phi\|_\infty \leq \|\phi\|_\infty. \tag{13}$$

Now we define the transfer operator $L_t : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$ by duality. If ν is a Borel signed measure on $\mathbb{R}^d$

$$\int_{\mathbb{R}^d} (P_t\phi)(x) d\nu(x) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(y) S_t(x,y) dy d\nu(x)$$

supposing that ν has a density with respect to the Lebesgue measure $f \in L^1(\mathbb{R}^d)$ i.e $d\nu = f(x)dx$ we can thus write

$$\int_{\mathbb{R}^d} (P_t\phi)(x)d\nu(x) = \int_{\mathbb{R}^d} \phi(y) \left(\int_{\mathbb{R}^d} S_t(x,y)f(x)dx \right) dy. \tag{14}$$

We can then define the transfer operator associated to the system and to the evolution time t as

Definition 8 (transfer operator)**.** *Given $f \in L^1(\mathbb{R}^d)$ we define the measurable function $L_t f : \mathbb{R}^d \to \mathbb{R}^d$ as follows. For almost each $y \in \mathbb{R}^d$ let*

$$[L_t f](y) := \int S_t(x,y)f(x)dx. \tag{15}$$

By (14) we then get the duality relation between the Kolmogorov and the transfer operator

$$\int (P_t\phi)(x)d\nu(x) = \int \phi(y)[L_t f](y)dy. \tag{16}$$

Lemma 9. *The operator L_t preserves the integral and is a weak contraction with respect to the L^1 norm.*

Proof. The first statement directly follows from (16) setting $\phi = 1$. For the second, we can work similarly using (13) and (16) with $\phi = \frac{[L_t f]}{||[L_t f]||}$. □

Since clearly L_t is a positive operator, we also get that L_t is a Markov operator having kernel S_t.

In the following we will define suitable spaces where operators L_t have nice spectral properties, allowing the study of their leading eigenvalues.

3.2 Perturbed operators

To prove the extreme event laws shown in Theorem 2 we will use the construction outlined in [14, 15] (see Section 4) adapted to our case.

This is based on the idea of considering the target set B_m as a hole in the phase space and the system as an open system. We study the associated open system by the related transfer operators. We define these operators and study its basic properties in this section.

Definition 10. *Let $t \geq 0$, $x \in D$ and $u_n \to 0$ be a real sequence. We denote by B_n the ball $B(x, \exp(-u_n))$. We define the "perturbed" versions of the Kolmogorov and transfer operators by setting*

$$\begin{aligned} (P_{t,n}\phi)(x) &:= \mathbb{E}[1_{B_n^c}(X_t^x)\phi(X_t^x)], \ \ \phi \in L^\infty(\mathbb{R}^d) \\ (L_{t,n}f)(x) &:= 1_{B_n^c}(x)(L_t f)(x), \ \ f \in L^1(\mathbb{R}^d). \end{aligned} \tag{17}$$

Analogously to what is done in Section 3.1 one can prove that the perturbed operators enjoy the following duality relation:

$$\int (P_{t,n}\phi)(x)f(x)dx = \int\int \phi(y)1_{B_n^c}(y)S_t(x,y)f(x)dydx \tag{18}$$

$$= \int \phi(y)1_{B_n^c}(y)\int S_t(x,y)f(x)dxdy \tag{19}$$

$$= \int \phi(y)\left[1_{B_n^c}(y)(L_tf)(y)\right]dy = \int \phi(y)(L_{t,n}f)(y)dy. \tag{20}$$

The iterates of the perturbed operator inherits its properties from the following lemma.

Lemma 11. *For every $t, s \geq 0$, $\phi \in L^\infty\left(\mathbb{R}^d\right)$*

$$P_{t,n}\left(P_{s,n}\left(\phi\right)\right)(x) = \mathbb{E}\left[1_{B_n^c}\left(X_t^x\right)1_{B_n^c}\left(X_{t+s}^x\right)\phi\left(X_{t+s}^x\right)\right].$$

Proof.

$$\begin{aligned} P_{t,n}\left(P_{s,n}\left(\phi\right)\right)(x) &= \mathbb{E}\left[1_{B_n^c}\left(X_t^x\right)P_{s,n}\left(\phi\right)\left(X_t^x\right)\right] \\ &= \mathbb{E}\left[1_{B_n^c}\left(X_t^x\right)\mathbb{E}\left[1_{B_n^c}\left(X_s^y\right)\phi\left(X_s^y\right)\right]_{y=X_t^x}\right] \\ &= \mathbb{E}\left[1_{B_n^c}\left(X_t^x\right)\mathbb{E}\left[1_{B_n^c}\left(X_{t+s}^x\right)\phi\left(X_{t+s}^x\right)|\mathcal{F}_t\right]\right] \end{aligned}$$

by Markov property,

$$\begin{aligned} &= \mathbb{E}\left[\mathbb{E}\left[1_{B_n^c}\left(X_t^x\right)1_{B_n^c}\left(X_{t+s}^x\right)\phi\left(X_{t+s}^x\right)|\mathcal{F}_t\right]\right] \\ &= \mathbb{E}\left[1_{B_n^c}\left(X_t^x\right)1_{B_n^c}\left(X_{t+s}^x\right)\phi\left(X_{t+s}^x\right)\right] \end{aligned}$$

by the basic properties of the conditional expectation (see [24, Section 9.7]). □

Remark 12. *In particular, given generic $s, t \in \mathbb{R}$*

$$P_{t+s,n}\left(\phi\right)(x) \neq P_{t,n}\left(P_{s,n}\left(\phi\right)\right)(x)$$

namely $P_{t,n}$ is not a semigroup. However notice that if $t_1 < t_2$, then (we take $t = t_1$ and $t + s = t_2$ above)

$$P_{t_1,n}\left(P_{t_2-t_1,n}\left(\phi\right)\right)(x) = \mathbb{E}\left[1_{B_n^c}\left(X_{t_1}^x\right)1_{B_n^c}\left(X_{t_2}^x\right)\phi\left(X_{t_2}^x\right)\right].$$

Thus, for $P_t^{(n)}$ we have the following

Corollary 13. *For every $x \in \mathbb{R}^d$, $0 = t_0 < t_1 < ... < t_n$, $\phi \in L^\infty\left(\mathbb{R}^d\right)$, we have*

$$\left(P_{t_0,n} \circ P_{t_1-t_0,n} \circ \cdots \circ P_{t_n-t_{n-1},n}\right)(\phi)(x) = \mathbb{E}\left[1_{B_n^c}\left(X_{t_0}^x\right)\cdots 1_{B_n^c}\left(X_{t_n}^x\right)\phi\left(X_{t_n}^x\right)\right].$$

3.3 Functional Spaces: quasi-Hölder space on $\mathbb{R}^d$

We now define suitable functional spaces on which the transfer operators introduced in the previous sections have a regularizing behavior and nice spectral properties.

We construct spaces which are suitable for our noncompact environment, imposing a controlled behavior far away from the origin by using weight functions growing at infinity. These spaces will be denoted as BV_α and L^1_α. Let $\alpha > 0$ and define the weight function

$$\rho_\alpha(|x|) = \left(1 + |x|^2\right)^{\alpha/2}. \tag{21}$$

Let $L^1_\alpha\left(\mathbb{R}^d\right)$ be the space of Lebesgue measurable $f : \mathbb{R}^d \to \mathbb{R}$ such that

$$\|f\|_{L^1_\alpha(\mathbb{R}^d)} := \int_{\mathbb{R}^d} \rho_\alpha(|x|)\,|f(x)|\,dx < \infty.$$

Note that, $L^1_\alpha \subset L^1$ and if $f \in L^1_\alpha$ then $\|f\|_{L^1} \leq \|f\|_{L^1_\alpha}$. Moreover for $\alpha = 0$, $L^1_0 = L^1$.

We now adapt the Bounded Variation spaces used in [3] to the cases at hand.[2] For a Borel subset $S \subseteq \mathbb{R}^d$ let us define

$$osc(f, S) = \operatorname*{ess\,sup}_{x\in S} f - \operatorname*{ess\,inf}_{x\in S} f.$$

We now define the seminorm:

$$\|f\|_{osc(\mathbb{R}^d)} = \sup_{\epsilon\leq\epsilon_0} \epsilon^{-1} \int_{\mathbb{R}^d} osc(f, B_\epsilon(x))d\psi(x),$$

where ϵ_0 is a positive number which is not restrictive to choose equal to 1 in our case [3]

Here the measure ψ is a Radon probability measure on $\mathbb{R}^d$ and we require that:

($A_\psi 1$) ψ is absolutely continuous with respect to the Lebesgue measure, having a continuous bounded density ψ' such that $\psi' > 0$ everywhere.

We can define a $\|\cdot\|_{BV_\alpha}$ norm by setting

$$\|f\|_{BV_\alpha} := \|f\|_{L^1_\alpha(\mathbb{R}^d)} + \sup_{\epsilon\in(0,1]} \epsilon^{-1} \int_{\mathbb{R}^d} osc\left(f, B_\epsilon(x)\right) d\psi(x). \tag{22}$$

[2] These kind of spaces have also been used already in the context of extreme value theory (see [7, 5]).

[3] By inspecting carefully the proof of next Lemma 15 and Proposition 21 we see that our choice of $\epsilon_0 = 1$ in the definition of oscillation does not impose any limitation. Since we are considering a localization scheme through $B_n(x_0)$ it sufficient to choose ϵ large enough to include sets $B_n(x_0)$. Moreover by Lemma 18 we see that, beside the discontinuities created by our 1_{B_n}, the oscillation of $L_t f$ can be easily controlled, as $L_t f$ is regularized to C^1. This setup is different from those originally of Keller and Saussol, where ϵ_0 was tied to the size of the partitions and the expansions of the considered maps, and had to be chosen accordingly.

It is not difficult to show that $\|\cdot\|_{BV_\alpha}$ indeed defines a norm and the set of L^1_α functions for which this norm is finite is a Banach space which we denote by $BV_\alpha\left(\mathbb{R}^d\right)$.[4]

We prove that our weighted bounded variation space is compactly immersed in L^1.

Theorem 14. *$BV_\alpha\left(\mathbb{R}^d\right) \hookrightarrow L^1\left(\mathbb{R}^d\right)$ is a compact embedding.*

Proof. We prove that given a sequence $g_n \in L^1$ such that $||g_n||_{BV_\alpha} \leq M$ for some M there is a subsequence g_{n_k} and $g \in L^1$ such that $||g_{n_k} - g||_{L^1} \to 0$.

Let us consider a sequence $B_m = B_m(0)$ of balls centered in the origin with radius m, eventually covering $\mathbb{R}^d$. Let us fix m. By the fact that on a compact domain the usual BV topology is equivalent to BV_α and the space $BV(B_m)$ has a compact immersion in $L^1(B_m)$ there is a subsequence $g_{n_{m,k}}$ and a function $f_m : B_m \to \mathbb{R}$ such that $g_{n_{m,k}}$ restricted to B_m converges to f_m in the L^1 topology.

Let us define the extension $\overline{f_m}$ of f_m to $\mathbb{R}^d$ by

$$\overline{f_m}(x) = \begin{cases} f_m(x) \ if \ x \in B_m \\ 0 \ if \ x \notin B_m. \end{cases}$$

Since $\|g_{n_{m,k}}\|_{L^1} \leq M$ we also have $\|\overline{f_m}(x)\|_{L^1} \leq M$. Once found $g_{n_{m,k}}$ and f_m, now consider B_{m+1} and from the sequence $g_{n_{m,k}}$ let us draw a subsequence $g_{n_{m+1,k}}$ converging on B_{m+1} to some f_{m+1}. Being a subsequence of the previously extracted sequence, $g_{n_{m+1,k}}$ will converge to f_m on B_m and then $f_m = f_{m+1}$ on B_m. We can then continue inductively and define for each $m \geq 0$ a sub sequence $g_{n_{m,k}}$ and a function f_m with $g_{n_{m,k}} \to f_m$ on B_m. Furthermore we will also have an extension $\overline{f_m}$ on $\mathbb{R}^d$ for each $m \geq 0$. What we have seen shows that the sequence $m \to \overline{f_m}$ converges pointwise to some function $f : \mathbb{R}^d \to \mathbb{R}$. The sequence $\left|\overline{f_m}(x)\right|$ is an increasing sequence and then by the monotone convergence theorem $f \in L^1$ and $||f||_{L^1} \leq M$.

Now for each m consider k_m such that

$$\int_{B_m} |g_{n_{m,k_m}} - f_m| dx \leq \frac{1}{m}.$$

Since $\|g_{n_{m,k_m}}\|_{L^1_\alpha}$ is uniformly bounded we have that there is some M_2 independent of m such that $\int_{\mathbb{R}^d \setminus B_m} |g_{n_{m,k_m}}| dm \leq \frac{M_2}{\rho_\alpha(m)}$. We have that

$$||g_{n_{m,k_m}} - \overline{f_m}||_{L^1} \leq \frac{1}{m} + \frac{M_2}{\rho_\alpha(m)}$$

and thus $g_{n_{m,k_m}} \to f$ in the L^1 topology. □

[4]The proof can be obtained by adapting Propositions B.4 and B.5 in B. Saussol PhD thesis ([22]). First of all one notices that $L^1_\alpha(\mathbb{R}^d)$ is complete. Then, if f_n is a Cauchy sequence in BV_α, it is also Cauchy in $L^1_\alpha(\mathbb{R}^d)$ and therefore in $L^1(\psi)$, since $\psi' \in L^\infty(\mathbb{R}^d)$. Then one finally applies Propositions B.4 and B.5 in Saussol's thesis which explicitly uses $L^1(\psi)$ for the oscillatory part.

Let us suppose now that $f \in BV_\alpha$ and let $\mathcal{K}$ be a compact set in $\mathbb{R}^d$. We will need later on an estimate of the $L^\infty(\psi)$ norm of f on $\mathcal{K}$, designated as $||f||_{L^\infty(\psi,\mathcal{K})}$.

Proposition 15. *For any compact set $\mathcal{K} \subset \mathbb{R}^d$ we have*

$$||f||_{L^\infty(\psi,\mathcal{K})} \leq \frac{\max\left(||\psi'||_{L^\infty(\mathbb{R}^d)}, 1\right)}{d_{\mathcal{K}}} ||f||_{BV_\alpha}. \tag{23}$$

Where ψ' denotes the density of ψ with respect to the Lebesgue measure and $d_{\mathcal{K}} = essinf_{x\in\mathcal{K}}\psi(B_1(x))$. (We remark that by $A_\psi 1$ we have $d_{\mathcal{K}} > 0$.)

Following Proposition B.6 in [22] we can write for any $x \in \mathcal{K}$ and $y \in B_1(x)$: $|f(x)| \leq |f(y)| + \mathrm{osc}(f, B_1(x))$. By integrating in y over $B_1(x)$ we get

$$|||f(\cdot)|\psi(B_1(\cdot))||_{L^\infty(\psi,\mathcal{K})} \leq ||f||_{L^1(\psi)} + \epsilon||f||_{osc(\mathbb{R}^d)} \leq ||f||_{L^1(\psi)} + ||f||_{osc(\mathbb{R}^d)}. \tag{24}$$

By using the fact that $||f||_{L^1(\psi)} \leq ||\psi'||_{L^\infty(\mathbb{R}^d)}||f||_{L^1_2(\mathbb{R}^d)}$, we finally get the statement.

Remark 16. *From now on we will denote with $||f||_{L^\infty(B)}$ the L^∞ norm with respect to the Lebesgue measure restricted to the compact set B. By the assumption ($A_\psi 1$), $||f||_{L^\infty(B)}$ differs from $L^\infty(\psi, B)$ by a multiplicative constant depending only upon B.*

3.4 Regularization properties for the transfer operator.

In this section we see how the properties of the SDE (2) we consider, such those derived from the Brownian motion, have a regularizing effect at the level of the associated transfer operators.

In the following lemma the notation L^1_2 stands for the space L^1_α when $\alpha = 2$.

Lemma 17. *Given $h > 0$, there exist constants $A, B > 0$ and $\lambda \in (0, 1)$ (also depending on h) such that*

$$\|L^n_h f\|_{L^1_2} \leq A\lambda^n \|f\|_{L^1_2} + B \|f\|_{L^1} \tag{25}$$

for every $f \in L^1_2$ and every $n \in \mathbb{N}$.

Proof. **Step 1**. Call L^1_{dens} the set of all probability density functions, namely the elements $p \in L^1$ such that $p \geq 0$ a.s. and $\|p\|_{L^1} = 1$.

The statement of the lemma is equivalent to prove there exist constants $C, D > 0$ and $\lambda \in (0, 1)$ such that inequality

$$\|L^n_h p\|_{L^1_2} \leq C\lambda^n \|p\|_{L^1_2} + D \tag{26}$$

holds true for every $p \in L^1_{\mathrm{dens}} \cap L^1_2$. That the statement of the lemma implies this new one is obvious (with the same constants). Let us prove the converse. Take $f \in L^1_2$ and call $f^+ = f \vee 0$, $f^- = (-f) \vee 0$ (thus $f = f^+ - f^-$). Since

$|f^{\pm}(x)| \leq |f(x)|$, we have $f^{\pm} \in L^1_2 \subset L^1$. Call $Z_{\pm} = \|f^{\pm}\|_{L^1}$ and assume $Z_{\pm} > 0$ (the case when one or both are zero is easier). Call $p^{\pm} = Z_{\pm}^{-1} f^{\pm}$, elements of $L^1_{\text{dens}} \cap L^1_2$. By assumption,

$$\left\|L_h^n p^{\pm}\right\|_{L^1_2} \leq C\lambda^n \left\|p^{\pm}\right\|_{L^1_2} + D$$

hence, by linearity of L_h^n and homogeneity of the norms,

$$\left\|L_h^n f^{\pm}\right\|_{L^1_2} \leq C\lambda^n \left\|f^{\pm}\right\|_{L^1_2} + Z_{\pm} D.$$

Again by linearity of L_h^n,

$$\begin{aligned}
\|L_h^n f\|_{L^1_2} &= \left\|L_h^n \left(f^+ - f^-\right)\right\|_{L^1_2} \leq \left\|L_h^n f^+\right\|_{L^1_2} + \left\|L_h^n f^-\right\|_{L^1_2} \\
&\leq C\lambda^n \left(\left\|f^+\right\|_{L^1_2} + \left\|f^-\right\|_{L^1_2}\right) + \left(\left\|f^+\right\|_{L^1} + \left\|f^-\right\|_{L^1}\right) D \\
&\leq 2C\lambda^n \|f\|_{L^1_2} + 2D \|f\|_{L^1}
\end{aligned}$$

where in the last step we have used the fact that $|f^{\pm}(x)| \leq |f(x)|$, and the definition of the norms in L^1_2 and L^1. Therefore (25) holds with $A = 2C$ and $B = 2D$.

Step 2. Let us prove (26). Given $p \in L^1_{\text{dens}} \cap L^1_2$, on a suitable probability space choose a random initial condition X_0 with density p, independent of the Brownian motion. Call X_t the solution of the Cauchy problem with initial condition X_0. Recall that the Kolmogorov operator P_t is defined by means of the solutions $X_t(x)$ of the same Cauchy problem but with deterministic initial condition x, $(P_t\phi)(x) = \mathbb{E}[\phi(X_t(x))]$. A simple disintegration argument proves that

$$\int_{\mathbb{R}^d} (P_t\phi)(x)\, p(x)\, dx = \mathbb{E}[\phi(X_t)] \tag{27}$$

where X_t, as defined above, is the solution with initial condition X_0. We shall use now this relation together with the duality relation between the Kolmogorov operator and the transfer operator. We have (recall that $L_h^n = L_{hn}$)

$$\begin{aligned}
\|L_h^n p\|_{L^1_2} &= \int_{\mathbb{R}^d} \left(1 + |x|^2\right) |(L_{hn}p)(x)|\, dx \\
&= \int_{\mathbb{R}^d} \left(1 + |x|^2\right) (L_{hn}p)(x)\, dx
\end{aligned}$$

because p is a probability density and $L_{hn}p$ is thus non-negative,

$$= \int_{\mathbb{R}^d} (P_{hn}\phi)(x)\, p(x)\, dx$$

by the duality relation mentioned above, where $\phi(x) = 1 + |x|^2$,

$$= \mathbb{E}[\phi(X_{hn})]$$

by (27). We have thus proved that

$$\|L_h^n p\|_{L_2^1} = \mathbb{E}\left[1 + |X_{nh}|^2\right].$$

Hence inequality (26) reduces to prove

$$\mathbb{E}\left[1 + |X_{nh}|^2\right] \leq C\lambda^n \mathbb{E}\left[1 + |X_0|^2\right] + D. \tag{28}$$

Step 3. In this step we prove the inequality

$$\mathbb{E}\left[|X_t|^2\right] \leq e^{-2L_2}\mathbb{E}\left[|X_0|^2\right] + \frac{2L_1 + d}{2L_2} \tag{29}$$

where L_1, L_2 are the constants in the assumption on b and d is the space dimension. It is straightforward to see that (29) implies (28), completing the proof of the lemma.

It is well known that, when $\mathbb{E}\left[|X_0|^2\right] < \infty$, we have

$$\sup_{t\in[0,T]} \mathbb{E}\left[|X_t|^2\right] < \infty$$

for every $T > 0$ (also with the supremum inside the expectation). For completeness we give the proof of this statement in Step 4 below. Here we shall use this fact.

By Itô formula,

$$|X_t|^2 = |X_0|^2 + \int_0^t 2\left\langle X_s, b\left(X_s\right)\right\rangle ds + \int_0^t 2\left\langle X_s, dW_s\right\rangle + Tr\left(I\right)t.$$

Assume $\mathbb{E}\left[|X_0|^2\right] < \infty$. Then $\mathbb{E}\int_0^T |X_t|^2\, dt < \infty$ for every $T > 0$ and thus, by the properties of Itô integrals, $\mathbb{E}\int_0^t 2\left\langle X_s, dW_s\right\rangle = 0$. Then

$$\mathbb{E}\left[|X_t|^2\right] = \mathbb{E}\left[|X_0|^2\right] + \int_0^t 2\mathbb{E}\left\langle X_s, b\left(X_s\right)\right\rangle ds + Tr\left(I\right)t.$$

This identity and the fact that the function $s \mapsto \mathbb{E}\left\langle X_s, b\left(X_s\right)\right\rangle$ is continuous, imply that the function $t \mapsto \mathbb{E}\left[|X_t|^2\right]$ is of class C^1 and

$$\frac{d}{dt}\mathbb{E}\left[|X_t|^2\right] = 2\mathbb{E}\left\langle X_t, b\left(X_t\right)\right\rangle + d.$$

From the assumptions on b,

$$\frac{d}{dt}\mathbb{E}\left[|X_t|^2\right] \leq -2L_2\mathbb{E}\left[|X_t|^2\right] + 2L_1 + d.$$

This implies

$$\begin{aligned}\mathbb{E}\left[|X_t|^2\right] &\leq e^{-2L_2 t}\mathbb{E}\left[|X_0|^2\right] + \int_0^t e^{-2L_2(t-s)}\left(2L_1+d\right)ds \\ &\leq e^{-2L_2 t}\mathbb{E}\left[|X_0|^2\right] + \frac{2L_1+d}{2L_2}.\end{aligned}$$

Step 4. Given $R>0$, let τ_R be the first time $|X_t|$ exceeds R, infinity if this never happens. We have

$$\begin{aligned}|X_{t\wedge\tau_R}|^2 &= |X_0|^2 + \int_0^{t\wedge\tau_R} 2\left\langle X_s, b\left(X_s\right)\right\rangle ds + \int_0^{t\wedge\tau_R} 2\left\langle X_s, dW_s\right\rangle + Tr\left(I\right)t\wedge\tau_R \\ &= |X_0|^2 + \int_0^t 1_{\{s\leq\tau_R\}} 2\left\langle X_s, b\left(X_s\right)\right\rangle ds + \int_0^t 1_{\{s\leq\tau_R\}} 2\left\langle X_s, dW_s\right\rangle + Tr\left(I\right)t\wedge\tau_R.\end{aligned}$$

Since $\mathbb{E}\int_0^t 1_{\{s\leq\tau_R\}} 2\left\langle X_s, dW_s\right\rangle = 0$ we get

$$\begin{aligned}\mathbb{E}\left[|X_{t\wedge\tau_R}|^2\right] &\leq \mathbb{E}\left[|X_0|^2\right] + \mathbb{E}\int_0^t 1_{\{s\leq\tau_R\}} 2\left\langle X_s, b\left(X_s\right)\right\rangle ds + Tr\left(I\right)t \\ &\leq \mathbb{E}\left[|X_0|^2\right] + \mathbb{E}\int_0^t 1_{\{s\leq\tau_R\}}\left(2L_2\left|X_s\right|^2 + 2C_1\right)ds + Tr\left(I\right)t \\ &\leq \mathbb{E}\left[|X_0|^2\right] + \mathbb{E}\int_0^t \left(2L_2\left|X_{s\wedge\tau_R}\right|^2 + 2C_1\right)ds + Tr\left(I\right)t.\end{aligned}$$

By Gronwall lemma this implies, given any $T>0$,

$$\mathbb{E}\left[|X_{t\wedge\tau_R}|^2\right] \leq \left(\mathbb{E}\left[|X_0|^2\right] + 2TC_1 + Td\right)e^{2L_2 T} =: C$$

for every $t\in[0,T]$. By Fatou lemma,

$$\mathbb{E}\left[\lim_{R\to\infty}|X_{t\wedge\tau_R}|^2\right] \leq C.$$

Now, a.s., $\lim_{R\to\infty}\tau_R = +\infty$, because the solution X_t exists globally. Hence $\lim_{R\to\infty}|X_{t\wedge\tau_R}|^2 = |X_t|^2$ and the proof is complete. □

In the following Lemma we see how the presence of the noise, and then the possibility to see the transfer operator as a kernel operator, also provides a form of regularization.

Lemma 18. *For every $t>0$, L_t is bounded linear from $L^1\left(\mathbb{R}^d\right)$ to $C^1\left(\mathbb{R}^d\right)$; in particular there exists $C_t>0$ such that*

$$\|L_t f\|_{C^1(\mathbb{R}^d)} \leq C_t\|f\|_{L^1(\mathbb{R}^d)}. \tag{30}$$

Moreover if $f\in L^1_2$ then

$$\|L_t f\|_{BV_2} \leq C_{t,2}\|f\|_{L^1_2(\mathbb{R}^d)}. \tag{31}$$

Proof. The estimate (30) follows by the definition of the transfer operator (15) and the estimate on its derivatives at item 2 of Theorem 6. The estimate (31) follows from (22) considering that $\|L_t(f)\|_{L^1_2}$ is estimated by Lemma 17 and for $f \in C^1$ it holds

$$\begin{aligned} \|f\|_{osc(\mathbb{R}^d)} &= \sup_{\epsilon \leq 1} \epsilon^{-1} \int_{\mathbb{R}^d} osc(f, B_\epsilon(x)) d\psi(x) \\ &\leq 2\|f\|_{C^1} \int_{\mathbb{R}^d} d\psi(x) \leq C\|f\|_{C^1} \end{aligned} \tag{32}$$

since ψ is a probability measure. Finally by (30) we can bound $\|f\|_{osc(\mathbb{R}^d)}$ by the L^1 norm and then by the L^1_2 norm of f. □

By Lemma 17 and Lemma 18 we get

Lemma 19. *Given $h > 0$, there exist constants $A, B > 0$ and $\lambda \in (0,1)$ (also depending on h) such that*

$$\|L_h^n f\|_{BV_2} \leq A\lambda^n \|f\|_{BV_2} + B \|f\|_{L^1} \tag{33}$$

for every $f \in BV_2$ and every $n \in \mathbb{N}$.

Proof. By Lemma 18

$$\|L_h^n f\|_{BV_2} \leq C||L_h^{n-1} f||_{L^1_2} \tag{34}$$

and by Lemma 17

$$||L_h^{n-1} f||_{L^1_2} \leq A\lambda^{n-1} \|f\|_{L^1_2} + B \|f\|_{L^1} . \tag{35}$$

Putting together these two inequalities we get the statement. □

Remark 20. *As it is well known (see e.g [11]) the Lasota-Yorke-Doeblin-Fortet inequality proved in Lemma 19 together with the compact embedding proved in Theorem 14 implies that the operator L_h acting on BV_2 is quasicompact (see Section 4 for more details). The use of (31) together with Lemma 17, and Theorem 14 also implies that L_t is quasicompact when acting on L^1_2.*

It is worth to remark that since the phase space is $\mathbb{R}^d$ (and thus not bounded) L_t is not in general a compact operator when acting on L^1_2 and hence we cannot rely on a simple spectral perturbation theory for such operators.

For simplicity let us illustrate this in the particular case where $b(x) = -x$ and h small enough. We will find a bounded sequence $f_n \in L^1_2$ such that $L_t f_n$ has no converging subsequences. If h is small enough for each $x_0 \in \mathbb{R}^d$ by Theorem 6 one can find a radius $r_0 < 1$ and a point y_0 such that if a density f_0 with $||f_0||_{BV_2} = 1$ is supported on $B(x_0, r_0)$ then $L_h(f_0)$ is such that $\int_{B(y_0,1)} L_h(f_0)dx \geq 0.9 \int_{\mathbb{R}^d} f_0 dx$ (think about r_0 being very small and f_0 concentrated around a certain point x_0, then $L_h(f_0)$ will be concentrated in a certain neighborhood, of $y_0 = \theta_h(x_0)$ and the distribution will be dominated by gaussians depending on h). Still by Theorem 6 and the fact that $b(x) = -x$ we can get

that if h is small enough we can choose y_0 such that $|y_0| \geq 0.7|x_0|$. Thus, if $|x_0| \geq 10$

$$\int_{B(y_0,1)} L_h(f_0)[|x|^2+1]dx \geq \frac{1}{2}||f_0||_{L^1_2} = \frac{1}{2}. \tag{36}$$

Now we can find f_1 with support far enough from the support of f_0 such that: $||f_1||_{L^1_2} = 1$, the support of f_1 is near to a point x_1 with $|x_1| \geq 10$ as before, there is a point y_1, such that $\int_{B(y_1,1)} L_h(f_1)dx \geq 0.9 \int_{\mathbb{R}^d} f_1 dx$ and $B(y_1,1) \cap B(y_0,1) = \emptyset$. This mean that $||L_h(f_1) - L_h(f_0)||_{L^2_1} \geq \frac{1}{2}$. One can then find similarly f_2 with support far enough from f_0, f_1 such that $||L_h(f_j) - L_h(f_i)||_{L^2_1} \geq \frac{1}{2}$ for $i \neq j \in \{0,1,2\}$ and so on, finding a sequence f_i on the unit ball such that $L_t f_i$ has no converging subsequences.

3.5 Regularization for the perturbed operators

In this section we prove a uniform Lasota Yorke inequality for the perturbed operators with strong and weak spaces BV_2, L^1_2.

Proposition 21. *There is $0 < \lambda < 1$ and two positive constants A', B' such that for any $f \in BV_2(\mathbb{R}^d)$, $n \geq 1$ and $m \geq 1$, we have*

$$||L^m_{t,n} f||_{BV_2} \leq A'\lambda^m ||f||_{BV_2} + B'||f||_{L^1(\mathbb{R}^d)}.$$

The proof of the Proposition is based on the following two preliminary results.

Proposition 22. *There are $\lambda < 1$, $A, B \geq 0$ such that for each $f \in L^1_2$, $m, n \geq 1$*

$$||L^m_{t,n} f||_{L^1_2} \leq A\lambda^m ||f||_{L^1_2} + B||f||_{L^1}. \tag{37}$$

We first consider the simplified case in which $f \geq 0$.

Lemma 23. *There are $\lambda < 1$, $B \geq 0$ such that for each $f \in L^1_2$ with $f \geq 0$, $m, n \geq 1$*

$$||L^m_{t,n} f||_{L^1_2} \leq A\lambda^m ||f||_{L^1_2} + B||f||_{L^1}. \tag{38}$$

Proof. Consider $f \geq 0$. We have that $L_{t,n} f = 1_{B_n^c} L_t(f)$ and $L_t f = L_{t,n} f + 1_{B_n} L_t(f)$. We define the sequence g_n in the following way: $g_m = L^m_t f - L^m_{t,n} f$.

Since $f \geq 0$ and clearly also $g_n \geq 0$ we have that

$$\|L^m_{t,n} f\|_{L^1_2} = \int |L^m_{t,n} f| \rho_2 dx \leq \int (L^m_{t,n} f + g_m)\rho_2 dx = ||L^m_t f||_{L^1_2},$$

by Lemma 17 we have then (38). □

Proof of Proposition 22. Let $f = f^+ - f^-$ be decomposed into its positive and negative part (where $f^+, f^- \geq 0$). It holds

$$||f||_{L^1_2} = \int |(f^+ - f^-)|\rho_2 dx = ||f^+||_{L^1_2} + ||f^-||_{L^1_2}$$

and by Lemma 23 we get

$$\begin{aligned} ||L^m_{t,n}f||_{L^1_2} &\leq 2A\lambda^n(||f^+||_{L^1_2}+||f^-||_{L^1_2})+2B(||f^+||_{L^1}+||f^-||_{L^1}) \\ &\leq 2A\lambda^n||f||_{L^1_2}+2B||f||_{L^1}. \end{aligned}$$

□

Proof of Proposition 21. The BV_2 norm is the sum of the oscillation part and of the L^1_2 part. For the latter we use Lemma 23. For the oscillation we integrate [3, Proposition 3.2 (ii)] with respect to our norm. Namely

$$\|1_{B_n^c}L_tf\|_{osc} \leq \sup_{0<\eta\leq 1}\frac{1}{\eta}\int Osc(L_tf, B_n^c\cap B_\eta(x))\mathbf{1}_{B_n^c}(x)d\psi(x)+ \tag{39}$$

$$\sup_{0<\eta\leq 1}\frac{1}{\eta}\int 2\left[\sup_{B_\eta(x)\cap B_n^c}|L_tf|\right]\mathbf{1}_{B_\eta(B_n)\cap B_\eta(B_n^c)}(x)d\psi(x), \tag{40}$$

where $B_\eta(x)$ is a ball centered at x and with radius η and given a set A, $B_\eta(A)=\{x;\mathrm{dist}(x,A)\leq\eta\}$. There are now two cases. We suppose first than $\eta<e^{-u_n}$; in this case the only points x contributing to the rightmost integral in the previous inequality, are those belonging to a 2η-closed neighborhood of the boundary of the ball B_n; we call S_n such an annulus. The Lebesgue measure of S_n will be bounded by a constant $\tilde{C}$ (depending on d) times η.[5]. In the second case, $\eta\geq e^{-u_n}$ only the points belonging to $B_{2\eta}(B_n)$ will contribute to the integral and the measure of these points is the volume of the hyper sphere of radius η which is $O(\eta^d)$. The term $\sup_{B_\eta(x)\cap B_n^c}|L_tf|$ will be bounded using Lemma 15 and 18, therefore By calling $\mathcal{K}_n$ the closed ball of radius $2e^{-u_n}$ and by using Proposition 15, we continue to bound the quantity in (40) as

$$\sup_{0<\eta\leq 1}\frac{\tilde{C}\eta}{\eta}||L_tf||_{L^\infty(\psi,S_n)}\psi(S_n)\leq 2\hat{C}\tilde{C}||f||_{L^1_2(\mathbb{R}^d)}||\psi'||_{L^\infty(\mathbb{R}^d)}, \tag{41}$$

where we used the bound of order η given by the first case, which includes also the second case of higher order η^{d-1}. The constant $\hat{C}$ maximizes the constant on the right hand side of (23), depending on ψ and on its strictly positive infimum over S_n, and the constant entering formula (31).

The right-hand side of (39) is bounded by $\|L_tf\|_{osc}$ and the latter is again bounded as in (31). Therefore we get:

$$||1_{B_n^c}L_tf||_{osc}\leq(\hat{C}+2\hat{C}\tilde{C}||\psi'||_{L^\infty(\mathbb{R}^d)})||f||_{L^1_2(\mathbb{R}^d)}.$$

If we now iterate this one we have, calling $c^*=\hat{C}+2\hat{C}\tilde{C}||\psi'||_{L^\infty(\mathbb{R}^d)}$, it holds

$$||L^m_{t,n}f||_{osc}\leq||L_{t,n}L^{m-1}_{t,n}f||_{osc}\leq c^*||L^{m-1}_{t,n}f||_{L^1_2(\mathbb{R}^d)}$$

[5]The volume of S_n is bounded by the volume of the corresponding annulus of size 2η around an hypersphere of radius 1. In this case the lowest order of such a volume is η times a constant depending on d.

and using Lemma 23 we continue as

$$||L^m_{t,n} f||_{osc} \leq c^* A \lambda^{m-1} ||f||_{L^1_2(\mathbb{R}^d)} + c^* B ||f||_{L^1(\mathbb{R}^d)}$$

If we now define $A' = \max(c^*/\lambda, A), B' = \max(c^*, B)$, we finally have

$$||L^m_{t,n} f||_{BV_2} \leq A' \lambda^m ||f||_{L^1_2(\mathbb{R}^d)} + B' ||f||_{L^1(\mathbb{R}^d)}, \tag{42}$$

which implies the desired result.

□

4 Rare Events Via Transfer Operator

For completeness, we remind here the statement of a result due to Keller and Liverani [14] which is fundamental to our general construction. We consider a Banach spaces $(\mathcal{B}, \|\cdot\|)$ and we denote with $\mathcal{B}^*$ its dual. Then let $L_\epsilon : \mathcal{B} \to \mathcal{B}$ be a family of uniformly bounded linear operators, where $\epsilon \in E$, and E is the interval $E = (0, \overline{\epsilon}]$ for some $\overline{\epsilon} > 0$.[6].

4.0.1 Perturbative hypotheses

R1 The operators L_ϵ, $\epsilon \in E$, must satisfy the spectral decomposition

$$\lambda_\epsilon^{-1} L_\epsilon = \varphi_\epsilon \otimes \nu_\epsilon + Q_\epsilon$$

where $\lambda_\epsilon \in \mathbb{C}$, $\varphi_\epsilon \in \mathcal{B}, \nu_\epsilon \in \mathcal{B}^*$, $Q_\epsilon : \mathcal{B} \to \mathcal{B}$ is a linear operator verifying

$$\sum_{n=0}^{\infty} \sup_{\epsilon \in E} \|Q^n_\epsilon\| < \infty. \tag{43}$$

Moreover, $L_\epsilon \varphi_\epsilon = \lambda_\epsilon \varphi_\epsilon$, $\nu_\epsilon L_\epsilon = \lambda_\epsilon \nu_\epsilon$, $\nu_\epsilon(\varphi_\epsilon) = 1$, $\nu_\epsilon Q_\epsilon = 0$, $Q_\epsilon(\varphi_\epsilon) = 0$. We also require that $\nu_0(\phi_\epsilon) = 1$ and

$$\sup_{\epsilon \in E} \|\phi_\epsilon\| < \infty. \tag{44}$$

R2 When ε is small, L_ε is a small perturbation of L_0, in the following sense:

$$\pi_\epsilon := \sup_{f \in \mathcal{B}, \|f\| \leq 1} |\nu_0((L_0 - L_\epsilon)(f))| \to 0, \ \epsilon \to 0.$$

[6] In general the set E could be taken as a closed set of parameters with $\epsilon = 0$ as an accumulation point. Moreover, as Keller wrote in [15] "It is assumed that E is a closed subset of $\mathbb{R}$, but again the parameter ϵ enters the estimates only via the derived quantities π_ϵ and Δ_ϵ (*see our conditions R2 and R3 below*), so that also this result is valid for more general sets of parameters."

R3 We now set

$$\Delta_\epsilon := \nu_0((L_0 - L_\epsilon)(\phi_0)).$$

Then we require that

$$\pi_\epsilon \|(L_0 - L_\epsilon)\varphi_0\| \leq \text{ constant } |\Delta_\epsilon|.$$

where the constant is independent of ϵ.

R4 Let us consider the following quantities

$$q_{k,\varepsilon} := \frac{\nu_0((L_0 - L_\varepsilon)L_\varepsilon^k(L_0 - L_\varepsilon)(\varphi_0))}{\Delta_\epsilon}.$$

We will assume that for each $k \geq 0$ the following limit exists

$$\lim_{\epsilon \to 0} q_{k,\varepsilon} = q_k,$$

and we pose

$$\theta = 1 - \sum_{k=0}^{\infty} q_k.$$

Proposition 24 (Proposition[14])**.** *If $R1 - R4$ are satisfied then*

$$\lambda_\epsilon = 1 - \theta\Delta_\epsilon + o(\Delta_\epsilon). \tag{45}$$

4.0.2 Sufficient conditions to check assumptions R1 and R2

We now give sufficient conditions to check R1 and we will show how to rewrite R2 in a form adapted to our current setting. We begin to introduce a second Banach space $(\mathcal{B}_w, \|\cdot\|_w)$ which we qualify as *weak* when compared with the *strong* Banach space $\mathcal{B}$. We will first suppose

R0' For all $f \in \mathcal{B}$, $\|f\|_w \leq \|f\|$ i.e. the weak norm is bounded by the strong norm and the unit ball of $\mathcal{B}$ is compact in $\mathcal{B}_w$. Moreover $\exists G \geq 0$ s.t.

$$\forall \epsilon \in E \ \forall f \in \mathcal{B}, \ \forall n \in \mathbb{N} : \|L_\epsilon^n f\|_w \leqslant G \, \|f\|_w.$$

We then require that R1 be satisfied for $\epsilon = 0$. To achieve the same result for $\epsilon > 0$ we need two ingredients which will allow us to apply another perturbative result by Keller and Liverani, [13]. The first ingredient asks that:

R1' The operators L_ϵ satisfy a uniform, with respect to $\epsilon \in E$, Lasota-Yorke (or Doeblin-Fortet) inequality: there exists $\alpha \in (0,1), D > 0$ such that

$$\forall \epsilon \in E \ \forall f \in \mathcal{B} \ \forall n \in \mathbb{N} : \|L_\epsilon^n f\| \leqslant D\alpha^n \|f\| + D\|f\|_w \tag{46}$$

Finally we ask the closeness of the operators in the so-called triple norm $|||\cdot|||$, namely there exists an upper semi-continuous function $u_\epsilon : [0,\infty) \to [0,\infty)$, $u_\epsilon > 0$, such that

R2'

$$|||L_0 - L_\epsilon||| := \sup_{f\in\mathcal{B},\|f\|\leq 1} \|(L_0 - L_\epsilon)(f)\|_w \leq u_\epsilon \to 0,\ \epsilon \to 0. \tag{47}$$

By assuming the previous conditions, it now follows from [13], see also [15], that the *quasi-compactness condition* (4.0.1) holds for any $\epsilon \in E, \epsilon > 0$. Moreover there will be $0 < \rho < 1$ such that for any $\epsilon \in E$, the spectral radius of Q_ϵ is bounded by ρ, which implies (43). Finally (44) follows immediately from (46).

Condition R2 could be easily worked out when the ν_0 will be a measure because in this case we could write

$$\pi_\epsilon \leq \sup_{f\in\mathcal{B},\|f\|\leq 1} \int |(L - L_\epsilon)(f)| d\nu_0. \tag{48}$$

If the weak norm is strong enough to bound this integral, condition R2' implies condition R2. More precisely, if we have that there is some C such that

$$\int |(L - L_\epsilon)(f)| d\nu_0 \leq C\|(L_0 - L_\epsilon)(f)\|_w$$

then R2' implies R2. We will see that in our case this holds.

4.1 Verifying the perturbative assumptions in our case

In this section we verify the perturbative assumptions needed to apply Proposition 24. In order to do this, we will get information on the spectral picture of L_t when applied to BV_2 and show that the perturbed operators $L_{t,n}$ (see (17)) can be seen as a small perturbations in some sense of the operator L_t (see (15)).

We denote again by B_n the ball $B(x_0, \exp(-u_n))$ and we remind that

$$(L_{t,n}f)(x) = 1_{B_n^c}(x)(L_t f)(x).$$

We now check that the perturbative assumptions listed in Sections 4.0.1 and 4.0.2 apply to these operators choosing L^1 as the weak norm and BV_2 as the strong one.

Assumption R0' comes directly from the definitions of the norms given in Section 3.3, from the compact embedding result proved at Theorem 14 and finally from Lemma 9

Assumption R1 . We first prove it for $\epsilon = 0$, then we extend it to $\epsilon \in E, \epsilon > 0$ once R1' will be proved, see below. Assumption R1 for $\epsilon = 0$ is verified by classical results on the spectral picture of regularizing linear operators. The compact immersion proved at Theorem 14 and the Lasota Yorke inequality proved at Lemma 19 for L_0, allows to apply the Ionescu-Tulcea-Marinescu theorem (see for instance [11]), and therefore the operator L_t has the following spectral decomposition $L_t = \sum_i \upsilon_i \Pi_i + Q$, where all υ_i are eigenvalues of L of modulus 1, Π_i are finite-rank projectors onto the associated eigenspaces, Q is a bounded operator with a spectral radius strictly less than 1. They satisfy

$\Pi_i\Pi_j = \delta_{\delta_{ij}}\Pi_i, Q\Pi_i = \Pi_i Q = 0$. By this theorem we also get that 1 is an eigenvalue and therefore the transfer operator will admit finitely many absolutely continuous stationary measures. Furthermore, the peripheral spectrum is completely cyclic. Then we need to show that 1 is a simple eigenvalue of L_t and that there is no other peripheral eigenvalue. This is achieved by the strict positivity of the Markov kernel S_t (see (15)) provided by Theorem 6. [7].Notice that the projector Π_1 will be the linear functional ν_0 in the assumption R1.
Assumption R1' verified in Proposition 21.
Assumption R2' is verified in the following proposition

Proposition 25. *The operator $L_{t,n}$ is a small perturbation of L_t, in the following sense: there is a monotone sequence $\pi_n \to 0$such that*

$$||L_{t,n} - L_t||_{BV_2 \to L^1} \leq \pi_n.$$

Proof. We have $||L_{t,n} - L_t||_{L^1} = \|1_{B_n}(L_t f)\|_{L^1} \leq \mathrm{Leb}(B_n)||L_t f||_\infty$. Notice that by Lemma 18 the $C^1(\mathbb{R})$ norm of $L_t f$, and therefore its L^∞ norm is bounded by $C_t\|f\|_{L^1(\mathbb{R}^d)}$. We could finally take $\pi_n = C_t\mathrm{Leb}(B_n)$.

□

Since in our case ν_0 is nothing but the integral with respect to the Lebesgue measure, and the weak norm is the L^1 norm, R2' implies R2 as remarked at the end of Section 4.0.2.
Assumption R3 in our setting can be stated in the following way:

Proposition 26. *Let f_0 denote the invariant density of the invariant measure μ; then there exists a constant C' such that*

$$\pi_n\|(L_t - L_{t,n})f_0\|_{BV_2} \leq C'\mu(B_n).$$

We saw above that π_n could be taken as $C_t\mathrm{Leb}(B_n)$; moreover by Proposition 21 the quantity $\|(L_t - L_{t,n})f_0\|_{BV_2} = \|1_{B_n}L_t f_0\|_{BV_2}$ is bounded by a constant $\tilde{K}$. Therefore $\pi_n\|(L_t - L_{t,n})f_0\|_{BV_2} \leq C_t\tilde{K}\mathrm{Leb}(B_n) \leq C_t\tilde{K}\frac{\mu(B_n)}{\inf_{B_n} f_0}$, which gives the desired result since the density f_0 is strictly positive on any finite domain of $\mathbb{R}^d$, as it is proved in the next lemma.

[7]Here is the argument. Consider first the operator $Lf(y) = \int S_h(x,y)f(x)dx$, where $S_h(\cdot,\cdot)$ is the Markov kernel defined on $\mathbb{R}^2$ and is bounded from below on any compact set $K \in \mathbb{R}^2$. We want to show that there exists only one fixed point for L. Suppose there are two, say h_1, h_2. Define $\hat{h} = \min(h_1, h_2)$. Notice that $h_1 - \hat{h}$ is nonnegative on a set Ω of positive Lebesgue measure; moreover from a classical trick, we have easily that $L(h_1 - \hat{h}) = h_1 - \hat{h}$. Take now a sequence of monotonically increasing compact sets K_n such that the first verifies $K_1 \subset \Omega$, $\mathrm{Leb}(K_1 \cap \Omega) > 0$ and $\bigcup_n K_n = \mathbb{R}^2$ (use the regularity of the Lebesgue measure). Then take $y \in K_n$ and we get $h_1(y) - \hat{h}(y) \geq \int_{x \in K_n \cap \Omega} S_h(x,y)(h_1(x) - \hat{h}(x))dx$. Notice that the right hand side is strictly positive. So we have that $h_1(y) > h_2(y),\ y \in K_n \ \Rightarrow \int_{K_n} h_1 dx > \int_{K_n} h_2 dx$. By passing to the limit we have $\int h_1 dx > \int h_2 dx$, which is impossible. Call h the unique fixed point of L. Consider now the peripheral spectrum of L which consists of a finite union of finite cyclic groups. For each of those eigenvalues call g one of the corresponding eigenvector. There exists $k \geq 1$ such that 1 is the unique peripheral eigenvalue of $L^k : L^k g = g$. But $L^k h = h$, and by repeating the proof above we see that $g = h$.

Lemma 27. *Let f_0 be the density of the unique invariant measure. Then*

$$\inf_{x\in B} f_0(x) > 0,$$

where B is any bounded subset of $\mathbb{R}^d$.

Proof. For $t > 0$, part 1 of Theorem 6 gives, for $(x, y) \in \mathbb{R}^d$:

$$D(x,y) := C_0^{-1} g_{\lambda_0^{-1}}(t, \theta_t(x) - y) \leq S_t(x,y).$$

The invariant measure $\mu = f_0 dx$ satisfies $f_0(x) = \int_{\mathbb{R}^d} S_t(x,y) f_0(y)\, dy$. Then take a compact $\hat{K}$ such that $\int_{\hat{K}} f_0 dy > 0.5$, which is possible by the regularity of the Lebesgue measure. Therefore, for $x \in B$

$$f_0(x) \geq \min_{x\in B, y\in \hat{K}} D(x,y) \int_{\hat{K}} f_0(y) dy > 0.5 \min_{x\in B, y\in \hat{K}} D(x,y)$$

and the minimum on D is strictly positive by the smoothness of $g_{\lambda_0}^{-1}$.

□

Assumption R4 needs some more work. The next proposition will show that all the quantities q_k defined in the Assumption R7 are equal to 0. In the present setting they are defined as the limit for $n \to \infty$ of the following quantities:

$$q_{k,n} = \frac{\int (L_t - L_{t,n}) L_n^k (L_t - L_{t,n})(f_0) dm}{\mu(B_n)}$$

where m is the Lebesgue measure on $\mathbb{R}^d$.

Proposition 28. *For each $k \geq 0$ we have*

$$\lim_{n\to\infty} q_{k,n} = 0.$$

Proof. Let us introduce the function $f_{k,n} := \frac{L_n^k (L_t - L_{t,n})(f_0)}{\mu(B_n)}$; we have

$$\lim_{n\to\infty} q_{k,n} = \lim_{n\to\infty} \int (L_t - L_{t,n}) f_{k,n} dm = \lim_{n\to\infty} \int 1_{B_n} L_t f_{k,n} dm \leq \lim_{n\to\infty} m(B_n) \|L_t f_{k,n}\|_\infty.$$

As in the proof of Proposition 25, by Lemma 18, the $C^1(\mathbb{R}^d)$ norm of $L_t f_{k,n}$, and therefore its infinity norm, is bounded by $C_t \|f_{k,n}\|_{L^1(\mathbb{R}^d)}$. On the other hand $C_t \|f_{k,n}\|_{L^1(\mathbb{R}^d)} \leq \frac{C_t Leb(B_n)}{\mu(B_n)}$ by proposition 25. The ratio $\frac{Leb(B_n)}{\mu(B_n)}$ is bounded by a constant, as shown in the proof of Proposition 26. Then we get

$$\lim_{n\to\infty} q_{k,n} \leq \mathrm{Leb}(B_n) \frac{C_t Leb(B_n)}{\mu(B_n)} \to 0.$$

□

4.2 Proof of Theorem 2

In this section we can collect all the previous estimates and finally prove the main result of the paper.

Proof of Theorem 2. We recall some notations. We consider a given point in the phase space x_0 and a sequence u_n going to 0; then we denote by B_n the ball $B\left(x_0, \exp\left(-u_n\right)\right)$. To compact notation we will write g for $g_{x_0} := -\log(d(x, x_0))$. We will denote with L_t the unperturbed transfer operator at time t (see (15)) and with $L_{t,n}$ the perturbed transfer operator (see (17)). Let us also denote as $h > 0$ the time discretization step introduced in (8).

Now, we rewrite (1), using the notation above and remembering that $t_k = kh$, as

$$\begin{aligned}
&\mathbb{P} \otimes \mu \left(\left\| X_{t_k}^x - x_0 \right\| \leq \exp\left(-u_n\right) \text{ for every } k = 0, \ldots, n-1 \right) \\
&= \mathbb{P} \otimes \mu \left(X_{t_k}^x \in B_n^c \text{ for every } k = 0, \ldots, n-1 \right) \\
&= \int_{\Omega \times \mathbb{R}^d} 1_{B_n^c}\left(X_{t_0}^x(\omega)\right) \cdots 1_{B_n^c}\left(X_{t_{n-1}}^x(\omega)\right) d\omega d\mu(x) \\
&= \int_{\mathbb{R}^d} \mathbb{E}\left[1_{B_n^c}\left(X_{t_0}^x\right) \cdots 1_{B_n^c}\left(X_{t_{n-1}}^x\right)\right] d\mu(x).
\end{aligned} \tag{49}$$

Thus we can reformulate (1) by identifying sequences $\{u_n\}_{n \in \mathbb{N}}$ such that

$$\lim_{n \to \infty} \int_{\mathbb{R}^d} \mathbb{E}\left[1_{B_n^c}\left(X_{t_0}^x\right) \cdots 1_{B_n^c}\left(X_{t_{n-1}}^x\right)\right] d\mu(x) \in (0, 1). \tag{50}$$

By using Corollary 13 together with (18), after recalling that the invariant measure μ is absolutely continuous w.r.t to Lebesgue with density by f_0, we are finally able to write the distribution of the maxima in an operator-like way as

$$\begin{aligned}
&\mathbb{P} \otimes \mu \left(\left(\max_{k=0,\ldots,n-1} g_{x_0}\left(X_{t_k}^x\right) \right) \leq u_n \right) \\
&= \int_D \mathbb{E}\left[1_{B_n^c}(X_{t_0}^x) \cdots 1_{B_n^c}(X_{t_{n-1}}^x) 1(X_{t_{n-1}}^x)\right] f_0(x) dx \\
&= \int \left(P_{t_0,n} \circ P_{t_1 - t_0,n} \circ \cdots \circ P_{t_n - t_{n-1},n}\right)(1)(x) f_0(x) dx = \int L_{h,n}^n f_0 dx
\end{aligned} \tag{51}$$

where $L_{h,n}^n$ denotes the n-th power of the operator $L_{h,n}$.

Now we apply Proposition 24 to the transfer operator L_h and to its perturbations $L_{h,n}$. We consider as a strong space the space $\mathcal{B}(\mathbb{R}^d) = BV_2$ and $L^1(\mathbb{R}^d)$ as a weak space. The assumptions of section 4.0.1 are verified by L_h and by $L_{h,n}$ thanks to the sufficient conditions quoted in section 4.1; therefore we have the following spectral decomposition for the perturbed operator $L_{h,n}$:

$$\lambda_{h,n}^{-1} L_{h,n} = f_{h,n} \otimes \mu_{h,n} + Q_{h,n} \tag{52}$$

where $f_{h,n} \in BV_2$, $\mu_{h,n} \in (BV_2)'$ and $Q_{h,n} : BV_2 \to BV_2$ is a bounded operator with spectral radius uniformly bounded in n by some $\rho < 1$. Moreover $L_{h,n} f_{h,n} = \lambda_{h,n} f_{h,n}$ and $\mu_{h,n} L_{h,n} = \lambda_{h,n} \mu_{h,n}$.

By denoting with $\langle \mu_{h,n}, g \rangle$ the action of the linear functional $\mu_{h,n}$ over $g \in BV_2$, we have

$$L_n g = \lambda_{h,n} f_{h,n} \langle \mu_{h,n}, g \rangle + \lambda_{h,n} Q_{h,n}(g); \tag{53}$$

moreover we use the normalization $\int f_{h,n} dx = 1$ and $\langle \mu_{h,n}, f_{h,n} \rangle = 1$. Thus it is sufficient to control $\int (L_n^{n+1} f_0)(x) dx$ by plugging (53) in it. Since we have a direct sum decomposition of our operator, we can iterate and get

$$\int (L_n^n f_0)(x)\ dx = \lambda_{h,n}^n \langle \mu_{h,n}, f_0 \rangle + \lambda_{h,n}^n \int (Q_{h,n}^n f_0)(x)\ dx. \tag{54}$$

Now, 1 is the largest unique eigenvalue of the unperturbed operator L_h and by proposition 24 and proposition 28 we have $\theta = 1$, therefore:

$$\lambda_n = 1 - \mu(B_n) + o(\mu(B_n)). \tag{55}$$

Then by substituting (55) in (54) we have

$$\begin{aligned} \int (L_n^n f_0)(x)\ dx &= e^{n \log(1 - \mu(B_n) + o(\mu(B_n))} \left[\langle \mu_{h,n}, f_0 \rangle + \int Q_{h,n}^n f_0\ dx \right] \\ &= e^{-n\mu_0(B_n) + no(\mu_0(B_n))} \left[\langle \mu_n, f_0 \rangle + \int Q_{h,n}^n f_0\ dx \right]. \end{aligned} \tag{56}$$

Let us now recall that by [14, Lemma 6.1] we have $\langle \mu_n, f_0 \rangle \to 1$. Moreover by (43) we a uniform exponential convergence to zero of

$$\int Q_{h,n}^n f_0\ dx \le \|Q_{h,n}^n f_0\|_{L^1} \le \|Q_{h,n}^n f_0\|_{BV_2} \le \text{Const}\ \rho^n.$$

Now, as assumed in the statement of Proposition 2, we choose the sequence $\{u_n\}_{n \in \mathbb{N}}$ and a $\tau \in \mathbb{R}, \tau > 0$ such that

$$n\, \mu(B_n) \to \tau. \tag{57}$$

Thus

$$\int L_n^n f_0\ dx \to e^{-\tau} \tag{58}$$

proving (10).

□

5 Poisson statistics

Proof of Theorem 5. For this proof, similarly to the proof of Theorem 2 we will apply 24. We start by computing the characteristic function of the random variable $S_n = \sum_{i=0}^{n-1} 1_{B_n}(X_{ih}^x)$:

$$\Phi_n(s) = \int e^{isS_n} d\mathbb{P} d\mu = \int \mathbb{E}(e^{isS_n}) f_0 dx, \tag{59}$$

where f_0 is the density of μ. We then introduce the perturbed operators for $f \in BV_2(\mathbb{R}^d)$:[8]

$$\mathcal{L}_{t,n} f(x) = e^{is1_{B_n}(x)} L_t f(x), \tag{60}$$

$$\mathcal{P}_{t,n} f(x) = \mathbb{E}(e^{is1_{B_n}(X_h^x)} f(X_h^x)) \tag{61}$$

By using (15) we get, for $f \in BV_2(\mathbb{R}^d), g \in L^\infty$:

$$\int \mathbb{E}(e^{is1_{B_n}(X_h^x)} g(X_h^x)) f(x) dx = \int \mathcal{P}_{t,n} g(x) f(x) dx = \int \mathcal{L}_{t,n} f(x) g(x) dx. \tag{62}$$

We then observe that Lemma 11 holds for the new operator $\mathcal{P}_{t,n}$ just by replacing the characteristic function 1_{B_n} with $e^{is1_{B_n}}$.

Lemma 29. *For every $t, s \geq 0$, $\phi \in L^\infty\left(\mathbb{R}^d\right)$*

$$\mathcal{P}_{t,n}\left(\mathcal{P}_{s,n}\left(\phi\right)\right)(x) = \mathbb{E}\left[e^{is1_{B_n}(X_t^x)} e^{is1_{B_n}\left(X_{t+s}^x\right)} \phi\left(X_{t+s}^x\right)\right].$$

Using this and the equalities (62), we have

$$\Phi_n(s) = \int e^{isS_n} d\mathbb{P} d\mu = \int \mathcal{L}_{t,n}^n f_0(x) dx, \tag{63}$$

where f_0 is the density of the stationary measure. We now apply Theorem 24 to the operators $\mathcal{L}_{t,n}$, using BV_2 and L^1 as a strong and weak space . In order to do this, we check the assumptions (R1) to (R4), in particular (R1'), (R2'), (R3), (R4). We will omit some detail, as the proofs are similar to what is done in Section 4.1.

- (R1') Lasota-Yorke inequality. We follow Propositions 21 and Proposition 22. Let first compute the norm $\|\mathcal{L}_{t,n}^m f\|_{L_2^1}$, $m \geq 1$. We have

$$\|\mathcal{L}_{t,n}^m f\|_{L_2^1} = \|\mathcal{L}_{t,n} \mathcal{L}_{t,n}^{m-1} f\|_{L_2^1} = \|L_t \mathcal{L}_{t,n}^{m-1} f\|_{L_2^1} \leq$$

$$\|L_t |\mathcal{L}_{t,n}^{m-1} f|\|_{L_2^1|} \leq \|L_t^2 |\mathcal{L}_{t,n}^{m-2} f|\|_{L_2^1} \leq \cdots \leq \|L_t^m |f|\|_{L_2^1}.$$

[8] We now require that our functions have valued in $\mathbb{C}$; in this case the oscillation is defined as $\text{osc}(f, S) = \text{esssup}|f(x) - f(y)|, x, y \in S$. With this definition the function spaces and the norms considered extend straightforward to the complex case. We will keep denoting these spaces as $L_\alpha^1(\mathbb{R}^d)$ and $BV_\alpha(\mathbb{R}^d)$ also in the complex valued case.

We finally apply lemma 17 to the rightmost quantity to get the equivalent of Proposition 22. We now need the equivalent of Proposition 21, in particular we have to estimate the oscillation seminorm of $e^{is1_{B_n}} L_t f$. Using the formula[9], we get

$$\|e^{is1_{B_n}} L_t f\|_{\mathrm{osc}(\mathbb{R}^d)} = \sup_{0<\eta\leq 1} \frac{1}{\eta} \int \mathrm{osc}(L_t f, B_\eta(x)) d\psi +$$

$$\sup_{0<\eta\leq 1} \frac{1}{\eta} \int \mathrm{osc}(e^{is1_{B_n}}, B_\eta(x)) \sup |Lf| d\psi := (I) + (II)$$

The first piece (I) on the right hand side is $\|L_t f\|_{\mathrm{osc}(\mathbb{R}^d)}$ and is bounded as in (31). For the second one, we now distinguish two cases. We suppose first than $\eta < e^{-u_n}$; the oscillation in the integral will contribute only when the balls $B_\eta(x)$ will cross at the same time the ball B_n and its complement. In this case the oscillation will be constant and equal to $|e^{is} - 1|$ on the 2η-neighborhood of the ball B_n. If we call S_n such a neighborhood as in the proof of Proposition 21, the Lebesgue measure of S_n will be bounded by a constant $\tilde{C}$ (depending on d) times η. In the second case, $\eta \geq e^{-u_n}$ only the points belonging to $B_{2\eta}(B_n)$ will contribute to the integral and the measure of these points is the volume of the hyper sphere of radius η which is $O(\eta^d)$. Then we get as in (41), where we also introduced the constant $\hat{C}$:

$$(II) \leq \sup_{0<\eta\leq 1} \frac{\tilde{C}\eta}{\eta} ||L_t f||_{L^\infty(\psi, S_n)} \psi(S_n) \leq 2\hat{C}\tilde{C} ||f||_{L^1_2(\mathbb{R}^d)} ||\psi'||_{L^\infty(\mathbb{R}^d)}$$

Therefore we get

$$\|e^{is1_{B_n}} L_t f\|_{\mathrm{osc}(\mathbb{R}^d)} \leq C_{t,2} \|L_t f\|_{\mathrm{osc}(\mathbb{R}^d)} + \overline{C} ||f||_{L^1_2(\mathbb{R}^d)} \leq [C_{t,2} + \overline{C}] \, \|f\|_{L^1_2(\mathbb{R}^d)} ,$$

where $\overline{C} = 2\hat{C}\tilde{C}||\psi'||_{L^\infty(\mathbb{R}^d)}$. We then put $C^{@} := C_{t,2} + \overline{C}$ and we continue as

$$\|\mathcal{L}^m_{t,n} f\|_{\mathrm{osc}(\mathbb{R}^d)} = \|\mathcal{L}_{t,n} \mathcal{L}^{m-1}_{t,n} f\|_{\mathrm{osc}(\mathbb{R}^d)} = \|e^{is1_{B_n}} L_t(\mathcal{L}^{m-1}_{t,n} f)\|_{\mathrm{osc}(\mathbb{R}^d)} \leq$$

$$C^{@} \|\mathcal{L}^m_{t,n} f\|_{L^1_2(\mathbb{R}^d)} \leq C^{@} \|L^{m-1}_t |f|\|_{L^1_2}.$$

We therefore bound the right hand side of the previous inequality with lemma 17 and this will finally give us the Lasota-Yorke inequality, since $\|\mathcal{L}^m_{t,n} f\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)}$.

- (R2') We have to bound the quantity $||L_{t,n} - L_t||_{BV_2 \to L^1} \leq \pi_n$, where the quantity π_n was defined at item R2 in section 4.0.1.

  We have for $f \in BV_2$ and using Lemma 18:

[9] if $u, v \in BV_2$ and B a measurable set, then $\mathrm{osc}(uv, B) \leq \mathrm{osc}(u, B)\mathrm{esssup}_B v + \mathrm{osc}(v, B)\mathrm{essinf}_B |u|$.

$$||(\mathcal{L}_{t,n} - L_t)(f)||_{L^1} = \int_{B_n} |e^{is} - 1||L_t f| dx \leq 2\text{Leb}(B_n)\|L_t f\|_{C^1(\mathbb{R}^d)} \leq$$

$$2\text{Leb}(B_n)C_t\|f\|_{L^1(\mathbb{R}^d)} \leq 2\text{Leb}(B_n)C_t\|f\|_{BV_2},$$

with $\pi_n = 2\text{Leb}(B_n)C_t$.

- (R3') The closeness of the two operators is also quantified by

$$\Delta_n = \int (L_t - \mathcal{L}_{t,n}) f_0 dx = (1 - e^{is})\mu(B_n).$$

We have now to show that

$$\pi_n \|(\mathcal{L}_{t,n} - L_t)(f_0)\|_{BV_2} \leq \text{constant } |\Delta_n|.$$

Since the density f_0 is locally bounded away from zero, see Lemma 27, it will be enough to show that the quantity $\|(e^{is1_{B_n}} - 1)L_t(f_0)\|_{BV_2}$ is bounded by a constant independent of n. This follows by exactly the same arguments which allowed us to bound the BV_2 norm of $e^{is1_{B_n}} L_t$ in item (R1').

- (R4) The quantities q_k[10]. associated to this perturbation will therefore have the form

$$q_k = \lim_{n\to\infty} \frac{1}{\Delta_n} \int (L_t - \mathcal{L}_{t,n})\mathcal{L}_n^k (L_t - \mathcal{L}_{t,n})(f_0) dx, \tag{64}$$

provided the limits exists. We now show that all these quantities are zero, just by repeating the proof of Proposition 28. Let us consider the function $g_{k,n} := \frac{\mathcal{L}_{t,n}^k (L_t - \mathcal{L}_{t,n})(f_0)}{(1-e^{is})\mu(B_n)}$. We have

$$q_k = \lim_{n\to\infty} \int (L_t - \mathcal{L}_{t,n}) g_{k,n} dx = \lim_{n\to\infty} \int_{B_n} (1 - e^{is}) L_t g_{k,n} dx \leq$$

$$\lim_{n\to\infty} |1 - e^{is}| \text{Leb}(B_n) \|L_t g_{k,n}\|_\infty.$$

By using lemma 18, we have that $\|L_t g_{k,n}\|_\infty \leq C_t \|g_{k,n}\|_{L^1(\mathbb{R}^d)}$, and it is immediate to check that $\|g_{k,n}\|_{L^1(\mathbb{R}^d)} \leq 1$. In conclusion the limit defining q_k will tend to 0 when $n \to \infty$.

As in the proof of the Gumbel's law, we need that, given the number τ, the measure of the set B_n scales like $n\mu(B_n) \to \tau, n \to \infty$.

Then we get for the top eigenvalue ι_n of $\mathcal{L}_{t,n}$, see Proposition 24

$$\iota_n = 1 - (1 - \sum_{k=0}^{\infty} q_k)\Delta_n + o(\Delta_n) = 1 - (1 - e^{is})\mu(B_n) + o(\mu(B_n)).$$

[10] We use the same symbol as for the q_k introduced in item R4 in section 4.0.1

Since the assumptions of section 4.0.1 are verified by L_t and by $\mathcal{L}_{t,n}$ thanks to the sufficient conditions quoted in section 4.1, we could repeat the steps from eq. (53) to (56) showing that the leading term in the growing of $\Psi_n(s)$ is just given by the n-th power of ι_n. Therefore

$$\lim_{n\to\infty} \Phi_n(s) = \lim_{n\to\infty} \int e^{isS_n} d\mathbb{P} d\mu = \lim_{n\to\infty} \iota_n^n = e^{-(1-e^{is})\tau} := \Phi(s).$$

Notice that this is just the pointwise limit of the characteristic function of the random variable

$$S_{n,\tau} := \sum_{i=0}^{\lfloor \frac{\tau}{\mu(B_n)} \rfloor} 1_{B_n}(X_{ih}^x).$$

Since $\Phi(s)$ is continuous in $s = 0$, it is the characteristic function of a random variable W to which $S_{n,\tau}$ converges in distribution. But such a limiting variable W has the Poisson distribution

$$\nu_W(\{k\}) = \frac{e^{-\tau}\tau^k}{k!}.$$

□

Acknowledgement

P.G acknowledges the support of the Centro di Ricerca Matematica Ennio de Giorgi and of UniCredit Bank R&D group for financial support through the "Dynamics and Information Theory Institute" at Scuola Normale Superiore. The research of SV was supported by the project *Dynamics and Information Research Institute* within the agreement between UniCredit Bank and Scuola Normale Superiore di Pisa and by the Laboratoire International Associé LIA LYSM, of the French CNRS and INdAM (Italy); P.G acknowledges partial funding from PRIN 2022NTKXCX. SV thanks the Mathematical Research Institute MATRIX, the Sydney Mathematical Research Institute (SMRI), the University of New South Wales, and the University of Queensland for their support and hospitality and where part of this research was performed. SV was also supported by the project MATHAmSud TOMCAT 22-Math-10, N. 49958WH, du french CNRS and MEAE. The research of S.G. was partially supported by the research project "Stochastic properties of dynamical systems" (PRIN 2022NTKXCX) of the Italian Ministry of Education and Research. S.G. also thanks Aix-Marseille Université for hospitality during the research.

References

[1] J. Atnip, G. Froyland, C. González-Tokman, and S. Vaienti. "Thermodynamic Formalism and Perturbation Formulae for Quenched Random Open Dynamical Systems". In: *submitted, arxiv.org/pdf/2307.00774.pdf* (2023).

[2] Hale Aytaç, Jorge Milhazes Freitas, and Sandro Vaienti. "Laws of rare events for deterministic and random dynamical systems." English. In: *Trans. Am. Math. Soc.* 367.11 (2015), pp. 8229–8278. ISSN: 0002-9947; 1088-6850/e.

[3] B.Saussol. "Absolutely continuous invariant measures for multidimensional expanding maps." In: *Israel Journal of Mathematics* 116 (2000), pp. 223–248.

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

[5] Theophile Caby, Davide Faranda, Sandro Vaienti, and P. Yiou. "On the Computation of the Extremal Index for Time Series". In: *Journal of Statistical Physics* (Nov. 2019). DOI: 10.1007/s10955-019-02423-z.

[6] R.A. Davis. "Maximum and minimum of one-dimensional diffusions." In: *Stoch. Proc. Appl.* 13 (1982), pp. 1–9.

[7] D. Faranda, H. Ghoudi, P. Guiraud, and S. Vaienti. "Extreme value theory for synchronization of coupled map lattices." English. In: *Nonlinearity* 31.7 (2018), pp. 3326–3358. ISSN: 0951-7715; 1361-6544/e.

[8] Todd M Freitas A.C.M. Freitas J.M. "Hitting time statistics and extreme value theory". In: *Probab. Theory Relat. Fields* 147 (2010), pp. 675–710. DOI: https://doi.org/10.1007/s00440-009-0221-y.

[9] Avner Friedman. "Stochastic Differential Equations and Applications". In: *Stochastic Differential Equations*. Ed. by Jaures Cecconi. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 75–148. ISBN: 978-3-642-11079-5. DOI: 10.1007/978-3-642-11079-5_2. URL: https://doi.org/10.1007/978-3-642-11079-5_2.

[10] Michael Ghil and Valerio Lucarini. "The physics of climate variability and climate change". In: *Rev. Mod. Phys.* 92 (2020).

[11] L. Hervé H. Hennion. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasicompactness,* Lect. Notes in Math., 1766, Springer-Verlag, 2001.

[12] S. Vaienti J. Atnip G Froyland C. Gonzalez-Tokman. "Poisson statistics for quenched random dynamics". In: *In preparation* (0).

[13] G. Keller and C. Liverani. "Stability of the spectrum for transfer operators." In: *Annali Scuola Normale Superiore Pisa* XXVIII (1999), pp. 141–152.

[14] G. Keller and C. Liverani. "Rare events, escape rates and quasistationarity: Some exact formulae". In: *Journal of Statistical Physics* 135 (2009), pp. 519–534.

[15] Gerhard Keller. "Rare events, exponential hitting times and extremal indices via spectral perturbation". In: *Dynamical Systems-an International Journal - DYN SYST* 27 (Feb. 2012). DOI: 10.1080/14689367.2011.653329.

[16] FC Klebaner. *Introduction to stochastic calculus with applications.* English. Imperial College Press, 2012.

[17] A. Kuntz. "On Extremes of Multidimensional Stationary Diffusion Processes in Euclidean Norm." In: *https://mediatum.ub.tum.de/doc/1097695/1097695.pdf* (2002).

[18] Valerio Lucarini, Davide Faranda, Ana Cristina Gomes Monteiro Moreira de Freitas, Jorge Miguel Milhazes de Freitas, Mark Holland, Tobias Kuna, Matthew Nicol, Mike Todd, and Sandro Vaienti. *Extremes and recurrence in dynamical systems.* English. Hoboken, NJ: John Wiley & Sons, 2016, pp. xi + 296.

[19] H Rootzen MR Leadbetter. "Extremal theory for stochastic processes." In: *Ann. Prob.* 16 (1988), pp. 431–478.

[20] Françoise Pène, Benoît Saussol, and Roland Zweimüller. "RECURRENCE RATES AND HITTING-TIME DISTRIBUTIONS FOR RANDOM WALKS ON THE LINE". In: *The Annals of Probability* 41.2 (2013), pp. 619–635. URL: http://www.jstor.org/stable/23469427 (visited on 06/14/2023).

[21] X. Zhang S. Menozzi A. Pesce. "Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift". In: *J. Diff. Eq.* 272 (2021), pp. 330–369.

[22] B. Saussol. "Absolutely continuous invariant measures for multidimensional expanding maps. Etude statistique de systèmes dynamiques dilatants". In: *Université de Toulon et du Var, http://www.math.univ-brest.fr/perso/benoit.saussol/art/these.p (1998)* ().

[23] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes.* Classics in Mathematics. Reprint of the 1997 edition. Springer-Verlag, Berlin, 2006, pp. xii+338.

[24] D. Williams. *Probability with Martingales.* English. Cambridge Math. Textbooks, University Press, Cambridge, 1991.