Commun. Math. Phys. (2025) 406:231
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-025-05385-4

Communications in
Mathematical Physics

Extreme Value Theory and Poisson Statistics for Discrete Time Samplings of Stochastic Differential Equations

F. Flandoli[1], **S. Galatolo**[2], **P. Giulietti**[2], **S. Vaienti**[3]

[1] Scuola Normale Superiore, Pisa, Italy. E-mail: franco.flandoli@sns.it
[2] Dipartimento di Matematica, Università di Pisa, Pisa, Italy. E-mails: paolo.giulietti@unipi.it, stefano.galatolo@unipi.it
[3] Université de Toulon, Aix Marseille Université, CNRS, CPT, 13009 Marseille, France. E-mail: vaienti@cpt.univ-mrs.fr

Received: 19 March 2024 / Accepted: 17 June 2025
Published online: 1 September 2025 – © The Author(s) 2025

Abstract: We investigate the distribution and clustering of extreme events of 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 annealed transfer operators on suitable spaces of 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.

Contents

1. Introduction

In the last decades modeling based on Stochastic Differential Equations has been used extensively in different scenarios, in particular whenever the statistical properties of extreme events are 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 SDE of the 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 on the function b (see Sect. 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 for a given point x_0 of the phase space, and sample the evolution of g_{x_0} along the trajectories of X_t^x 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 clustering. 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 $\{u_n\}_{n\in\mathbb{N}}$, $u_n \in \mathbb{R}^+$, such that the following limit exists

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

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 called 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) shrinking sets (see [24] for an introduction to such relation).

Subsequently, we will consider the distribution of extreme occurrences in a given time interval. More precisely we will count the number of visits of the time-discretized solution of our SDE to a shrinking sequence of balls B_n among the first n steps of length h, suitably normalized by the measure of the balls. We will show that the distribution of the number of visits will converge to a standard Poisson distribution in the limit of large n, see Theorem 6.

We remark that, for both results, our construction is rooted in the functional approach of [2,14,15]. The tools we use to get the results are related to the study of the functional analytic properties of a transfer operator $\mathcal{L}_h$ 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 $\mathcal{L}_h$ to the perturbation of the system constructed by adding a small hole corresponding to the rare event, transforming the initial system into an open one (see [3–5] and references therein).

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 [20] with a focus on stochastic processes.

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

In [6], an EVL in continuous time is found in the case of one dimensional stochastic differential equations, using the properties of the Ornstein-Uhlenbeck process (see also [20] 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. Here $M_T = \max_{0\leq t\leq T} \|X_t\|$, where $\{X_t\}_{t\geq 0}$ is a $\mathbb{R}^n$-valued stationary reversible diffusion process and $\|\cdot\|$ the euclidean norm in $\mathbb{R}^n$. About these continuous time results, we remark that, as we show in "Appendix A", our explicit relation between the time T and the thresholds R (see Eq. (8) in Theorem 2) used for the discrete time result is not the right one in the continuous case. The time/threshold rescaling used in the mentioned continuous time results seems to be quite implicit and not easy to be determined in concrete examples.

In the absence of the drift term b, the discretization of $dX_t = dW_t$ is akin to a random walk. In [19] the hitting time distribution for random walks on the line is described; these results are equivalent to EVL as proved in [24].

Most of the preceding works obtained an extreme value distributions for continuous time processes beginning with specific SDEs, while in the present work we focus on discrete time but for a large class of SDEs. The latter should verify assumptions **H** and **AD** in Sect. 2, and also the conditions stated in Theorem 7. In this case we describe not only the distribution of the "first" rare event, but also the distribution of multiple extremes i.e. the Poisson statistics.

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 (Lipschitz drift) 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 $R_1, R_2 \in \mathbb{R}$ with $R_2 > 0$ such that for all $x \in \mathbb{R}^d$

$$\langle b(x), x\rangle \leq R_1 - R_2\|x\|^2. \tag{3}$$

Given the assumptions **H** and **AD**, the system (2) has a unique invariant measure μ (see Sect. 4.3 for more details).

Remark 1. The dissipativity assumption implies that the flow associated to the vector field b transports measures towards a compact region of the phase space. The previous assumptions include the fundamental case

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

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

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

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 $\langle B(x), x\rangle \leq C_1 + C_2 \|x\|^2$ for all $x \in \mathbb{R}^d$ 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 on B may help to accommodate linear first order operators in the discretization of a PDE in order to treat a larger class of SDE.

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

$$g_{x_0}(x) := -\log d(x, x_0) \tag{6}$$

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 the sequence of discrete times in which we sample the process be denoted as

$$t_k := kh. \tag{7}$$

We now consider the problem (1), and 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*

$$n\,\mu(B_n) \to \tau \tag{8}$$

where B_n is the ball $B\left(x_0, e^{-u_n}\right)$. Let $t_k := kh$. Then

$$\lim_{n\to\infty} \mathbb{P} \otimes \mu \left(\left\{(\omega, x) \quad \max_{k=0,\ldots,n-1} g_{x_0}\left(X^x_{t_k}\right) \leq u_n\right\}\right) = e^{-\tau}. \tag{9}$$

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

Remark 3. An analogous result 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.

Remark 4. The right hand side of the asymptotic law found in (9) does not depend on h. This is, in our opinion, due to the presence of a uniform noise, analogously to similar results for dynamical systems perturbed by annealed noise (see [4]). See "Appendix A" for a thorough discussion of the matter.

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 balls as in the previous theorem. We are now interested in studying the distribution of the number of visits to the set B_n in a prescribed time interval. In order to get a limiting distribution when $n \to \infty$ we have to rescale time as we did in (8).

Definition 5. Let us take $\tau > 0$ and $n \geq 1$. We define a sequence of discrete random variables counting the number of visits of the time sampled solution to the ball B_n among the first $\lfloor \frac{\tau}{\mu(B_n)} \rfloor$ iterations of the process

$$S_{n,\tau} := \sum_{i=0}^{\lfloor \frac{\tau}{\mu(B_n)} \rfloor} 1_{B_n}(X^x_{ih}). \tag{10}$$

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 $k \in \mathbb{N}$:

$$\lim_{n\to\infty} \mathbb{P} \otimes \mu \left(S_{n,\tau} = k\right) = \nu_W(\{k\}). \tag{11}$$

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

Theorem 6. *Let $\tau > 0$, X_t be the solution of* (2) *on $\mathbb{R}^d$, let B_n the ball $B\left(x_0, \exp\left(-u_n\right)\right)$ and let u_n such that $n\,\mu(B_n) \to \tau$. Then*

$$\lim_{n\to\infty} \mathbb{P} \otimes \mu \left(S_{n,\tau} = k\right) = \frac{e^{-\tau}\tau^k}{k!}.$$

As mentioned before, the proofs of Theorems 2 and 6 rely on the spectral perturbation results of [14,15]. For the proof of Theorem 6 we follow the strategy recently used in [2]. We consider the transfer operator $\mathcal{L}_h$ associated to the evolution of the system for a fixed time $h > 0$. We find Banach spaces where $\mathcal{L}_h$ satisfies a Lasota-Yorke type of inequality which in turn, if the spaces also embed compactly, implies a spectral gap, which is an important ingredient in order to apply the above cited spectral perturbation results. Since the ambient space is $\mathbb{R}^d$ and therefore non compact and we use singular perturbations generating discontinuous densities, we consider the space BV_α. This is a space of bounded variation densities, decaying at infinity with a certain power law of exponent α (hence allowing discontinuities). 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).

We remark that the use of bounded variation spaces, rather than the more regular Sobolev Space, has already been very fruitful in the study of SDE (see [1,9] to begin with).

2.1. Organization of the paper. The plan of the paper is as follows: in Sect. 3 we set up the functional analytic framework necessary to study our problems and we perform some regularization estimates which will be used in the sequel.

Section 4 recalls the main abstract tool we use, regarding the asymptotic behavior of the largest eigenvalue after perturbation (Proposition 23). Section 4.4 shows how from Proposition 23 we can recover the laws of extreme events stated in Theorem 2.

Section 5 shows how from Proposition 23 we can recover the Poisson statistics stated in Theorem 6.

"Appendix A" contains remarks on the choice of the time discretization step h and how, sending $h \to 0$, does not give further information on the EVL.

Notations. We will denote by L^p, $p \geq 1$ the space of measurable functions which are p-summable with respect to the Lebesgue measure on $\mathbb{R}^n$. If we consider another absolutely continuous measure with density ψ, we will write $L^p(\psi)$. If we will restrict to a measurable subset B of $\mathbb{R}^n$, we will set $L^p(B)$ or $L^p(\psi, B)$. For a particular density indexed by α, see Sect. 3.3, we put L^p_α. We will write C^r for $C^r(\mathbb{R}^d)$.

3. Transfer Operators, Banach Spaces and Regularization Lemmas

In this section, starting by a general result on the transition probabilities of a system like (2) we define and study the basic properties of the transfer operators associated to such system. In Sect. 3.3, we then introduce a certain weighted, bounded variation space. We will see in Sects. 3.4 and 3.5 that the transfer operator when applied to these functional spaces has useful regularization properties.

For the SDE considered in (2), assuming the regularity assumption (H) and the dissipative assumption (AD); it is known (see [8,23]) that the equation has the properties of strong existence of the solutions and pathwise uniqueness.

Moreover, Menozzi, Pesce and Zhang [26] 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

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

where b is the same as in (2). 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 7 ([26], 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 for each $y \in \mathbb{R}^d$. It 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 Theorem 7 will be used to define a Kolmogorov operator (composition operator) associated to our system.

Definition 8. The Kolmogorov operator $P_t : L^\infty \to C^0$ associated to the system (2) is defined as follows. Let $\phi \in L^\infty$. 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. Given the transition probabilities S_t it holds

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

By this we see that

$$\|P_t\phi\|_\infty \leq \|\phi\|_\infty. \tag{12}$$

Now we define the transfer operator $\mathcal{L}_t : L^1 \to L^1$ associated to the system and to the evolution time t 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 $f \in L^1$ with respect to the Lebesgue measure i.e $d\nu = f(x)dx$, we can thus write

$$\int_{\mathbb{R}^d} (P_t\phi)(x) d\nu(x) = \int_{\mathbb{R}^d} \phi(y) \left(\int_{\mathbb{R}^d} S_t(x, y) f(x) dx \right) dy. \tag{13}$$

We can then define the transfer operator $\mathcal{L}_t$ as

Definition 9 *(transfer operator).* Given $f \in L^1$ we define $\mathcal{L}_t f : \mathbb{R}^d \to \mathbb{R}^d$ as follows. For almost each $y \in \mathbb{R}^d$ let

$$(\mathcal{L}_t f)(y) := \int S_t(x, y) f(x) dx. \tag{14}$$

By (13) we recover the duality relation between the Kolmogorov and the transfer operator whenever $d\nu = f(x)dx$ i.e.

$$\int (P_t\phi)(x) d\nu(x) = \int \phi(y)(\mathcal{L}_t f)(y) dy. \tag{15}$$

Next Lemma is a well known property of the transfer operator, following directly from (12) and (15).

Lemma 10. *The operator $\mathcal{L}_t$ preserves the integral and is a weak contraction with respect to the L^1 norm.*

Since clearly $\mathcal{L}_t$ is a positive operator, we also get that $\mathcal{L}_t$ is a Markov operator having kernel S_t.

Remark 11. In the following we will define suitable spaces where the operators $\mathcal{L}_t$ have nice spectral properties. As a byproduct of this analysis we will give a proof of the existence and uniqueness of the stationary measure μ (see Sect. 4.3). We remind that the stationary measure satisfies

$$\int \phi(x) d\mu(x) = \int (P_t\phi)(x) d\mu(x).$$

Whenever it is absolutely continuous, its density f_0 is the fixed point of the transfer operator: $\mathcal{L}_t f_0 = f_0$.

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

This is based on the idea of considering the target set B_n as a hole in the phase space and consider it as an open system. Then we will study it by means of the related transfer operators and their properties.

Definition 12. Let $t > 0$, $x_0 \in \mathbb{R}^d$ and $u_n \to 0$ be a real sequence. We denote by B_n the ball $B\left(x_0, \exp\left(-u_n\right)\right)$. We define the "perturbed" versions of the Kolmogorov and transfer operators by setting

$$\begin{aligned}(P_{t,n}\phi)(x) &:= \mathbb{E}[1_{B_n^c}(X_t^x)\phi(X_t^x)],\ \phi \in L^\infty \\ (\mathcal{L}_{t,n} f)(x) &:= 1_{B_n^c}(x)(\mathcal{L}_t f)(x),\ f \in L^1.\end{aligned} \tag{16}$$

Analogously to what is done in Sect. 3.1, using the definition above, one can prove that the perturbed operators enjoy the following duality relation:

$$\int (P_{t,n}\phi)(x) f(x) dx = \int \phi(y)(\mathcal{L}_{t,n} f)(y) dy. \tag{17}$$

The iterates of the perturbed operator inherit their properties from the following lemma.

Lemma 13. *For every $t, s \geq 0$, $\phi \in L^\infty$*

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

Proof.

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

(by Markov property and where $\mathcal{F}_t$ is the filtration adapted to W_t)

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

by the basic properties of the conditional expectation (see [25, Sect. 9.7]). □

Remark 14. In particular, given generic $s, t \in \mathbb{R}^+$

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

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)

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

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

Corollary 15. *For every $x \in \mathbb{R}^d$, $0 = t_0 < t_1 < \cdots < t_n$, $\phi \in L^\infty$, 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 non-compact environment, imposing a controlled behavior far away from the origin by using weight functions growing at infinity. These spaces will be denoted by BV_α and L^1_α. Let $\alpha > 0$ and define the weight function

$$\rho_\alpha(|x|) = \left(1 + |x|^2\right)^{\alpha/2}. \tag{18}$$

Let L^1_α be the space of Lebesgue measurable $f : \mathbb{R}^d \to \mathbb{R}$ such that

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

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 [22] (see also [5,7,12], to the setup at hand. 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[2]:

$$\|f\|_{osc(\mathbb{R}^d)} = \sup_{0<\epsilon\leq 1} \epsilon^{-1} \int_{\mathbb{R}^d} osc(f, B_\epsilon(x))\,d\psi(x),$$

which we will simply denote by $\|\cdot\|_{osc}$. 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.

[2] In the definitions of [12,22] we have that $\epsilon < \epsilon_0$, for a suitable $\epsilon_0 > 0$ depending from the size of specific partitions and the expansions of the considered maps. However in our case such issues are not present and it is not restrictive to choose $\epsilon_0 = 1$.

Notice that by the assumption $(A_\psi 1)$, and given a measurable subset B, we have that $||f||_{L^\infty(B)} = ||f||_{L^\infty(\psi,B)}$.

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

$$\|f\|_{BV_\alpha} := \|f\|_{L^1_\alpha} + \sup_{\epsilon\in(0,1]} \epsilon^{-1} \int_{\mathbb{R}^d} osc\,(f, B_\epsilon\,(x))\, d\psi(x). \tag{19}$$

It is not difficult to show that $\|\cdot\|_{BV_\alpha}$ 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_α.[3]

We prove that BV_α is compactly immersed in L^1.

Theorem 16. *$BV_\alpha \hookrightarrow L^1$ is a compact embedding.*

Proof. We prove that given a sequence $g_n \in L^1$ such that $\|g_n\|_{BV_\alpha} \le M$ 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 $\mathfrak{B}_m = B(0, m)$ 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 $BV(\mathfrak{B}_m)$ has a compact immersion in $L^1(\mathfrak{B}_m)$ there are a subsequence $g_{n_{m,k}}$ and a function $f_m : \mathfrak{B}_m \to \mathbb{R}$ such that $g_{n_{m,k}}$ restricted to $\mathfrak{B}_m$ 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

$$\overline{f_m}(x) = \begin{cases} f_m(x) & \text{if } x \in \mathfrak{B}_m \\ 0 & \text{if } x \notin \mathfrak{B}_m. \end{cases}$$

Since $\|g_{n_{m,k}}\|_{L^1} \le M$ we also have $\|\overline{f_m}(x)\|_{L^1} \le M$. Once found $g_{n_{m,k}}$ and f_m, we then consider $\mathfrak{B}_{m+1}$ and from the sequence $g_{n_{m,k}}$ let us draw a subsequence $g_{n_{m+1,k}}$ converging on $\mathfrak{B}_{m+1}$ to some f_{m+1}. Being a subsequence of the previously extracted sequence, $g_{n_{m+1,k}}$ will converge to f_m on $\mathfrak{B}_m$ and then $f_m = f_{m+1}$ on $\mathfrak{B}_m$. We can then continue inductively and define for each $m \ge 0$ a subsequence $g_{n_{m,k}}$ and a function f_m with $g_{n_{m,k}} \to f_m$ on $\mathfrak{B}_m$. Furthermore we will also have an extension $\overline{f_m}$ on $\mathbb{R}^d$ for each $m \ge 0$. Thus 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 $||f||_{L^1} \le M$.

Now for each m consider k_m such that

$$\int_{\mathfrak{B}_m} |g_{n_{m,k_m}} - f_m|dx \le \frac{1}{m}.$$

Since $\|g_{n_{m,k_m}}\|_{L^1_\alpha}$ is uniformly bounded we have that there is some $M_2 > 0$ independent of m such that $\int_{\mathbb{R}^d \setminus \mathfrak{B}_m} |g_{n_{m,k_m}}|dm \le \frac{M_2}{\rho_\alpha(m)}$. We have that

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

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

[3] The proof can be obtained by adapting Propositions B.4 and B.5 in B. Saussol Ph.D. Thesis [21]. First of all one notices that L^1_α is complete. Then, if f_n is a Cauchy sequence in BV_α, it is also Cauchy in L^1_α and therefore in $L^1(\psi)$, since $\psi' \in L^\infty$ 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^∞ norm of f on $\mathcal{K}$, $||f||_{L^\infty(\mathcal{K})}$.

Proposition 17. *For any compact set $\mathcal{K} \subset \mathbb{R}^d$ we have*

$$\|f\|_{L^\infty(\mathcal{K})} \le \frac{\max\left(||\psi'||_{L^\infty}, 1\right)}{d_{\mathcal{K}}} ||f||_{BV_\alpha}, \tag{20}$$

where again 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 $d_{\mathcal{K}} > 0$.)

Proof. The proof follows closely that of Proposition 3.2, part (iv) in [22], with the difference that we use the measure ψ instead of Lebesgue. Since ψ is equivalent to Lebesgue on $\mathcal{K}$ with density $\psi,'$ this explains the presence of such a density in the numerator, if compared with the analogous formula in [22]. The latter has the Lebesgue measure of the ball $B_1(x)$ in the denominator, which we replaced with $d_{\mathcal{K}}$. □

3.4. Regularization properties for the transfer operator. In this section we see how the presence of the dissipativity assumption and of noise in the SDE (2) have a regularizing effect at the level of the associated transfer operators when these are applied to densities belonging to the spaces defined in the previous section. Having in mind the main goal of the paper which is the study of the extreme events of the process $X^x_{t_n}$ where $t_n = hn$ for a certain fixed time step $h > 0$, we now consider the iterates $\mathcal{L}^n_h$ of the transfer operators corresponding to a time time evolution $t = h$.

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

Lemma 18. *Given $h > 0$, there exist constants $A, B > 0$ and $\lambda \in (0, 1)$ (also depending on h) such that*

$$\left\|\mathcal{L}^n_h f\right\|_{L^1_2} \le A\lambda^n \|f\|_{L^1_2} + B\|f\|_{L^1} \tag{21}$$

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 $p \ge 0$ a.s. and $\|p\|_{L^1} = 1$.

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

$$\left\|\mathcal{L}^n_h p\right\|_{L^1_2} \le C\lambda^n \|p\|_{L^1_2} + D \tag{22}$$

holds true for every $p \in L^1_{\text{dens}} \cap L^1_2$. That the statement of the lemma implies this new one is obvious (with the same constants). For the converse, taking $f \in L^1_2$ and calling $f^+ = f \vee 0$, $f^- = (-f) \vee 0$ (thus $f = f^+ - f^-$), it is sufficient to notice that f^+ and f^- have disjoint supports and the L^1-norms are additive for functions with disjoint supports.

Step 2. We have to prove (22). In this step we show that it reduces to prove

$$\mathbb{E}\left[1 + |X_{nh}|^2\right] \le C\lambda^n \mathbb{E}\left[1 + |X_0|^2\right] + D \tag{23}$$

where X_0 is a random initial condition with density $p \in L^1_{\text{dens}} \cap L^1_2$, defined on a suitable probability space, and X_t the solution of the Cauchy problem with initial condition X_0. Indeed,

$$\begin{aligned}\left\|\mathcal{L}_h^n p\right\|_{L_2^1} &= \int_{\mathbb{R}^d}\left(1+|x|^2\right)|(\mathcal{L}_{hn} p)(x)|\,dx = \int_{\mathbb{R}^d}\left(1+|x|^2\right)(\mathcal{L}_{hn} p)(x)\,dx \\ &= \int_{\mathbb{R}^d}(P_{hn}\phi)(x)\,p(x)\,dx = \mathbb{E}\left[\phi\left(X_{hn}\right)\right] = \mathbb{E}\left[1+|X_{nh}|^2\right]\end{aligned}$$

where $\phi(x) = 1+|x|^2$, and $\mathbb{E}\left[1+|X_0|^2\right]$ is equal to $\|p\|_{L_2^1}$. The second identity above holds since p is positive and $\mathcal{L}_{hn}$ preserves positivity; the third identity is the duality relation between the Kolmogorov operator and the transfer operator; the fourth identity is due to a classical disintegration argument.

Step 3. In this step we prove the inequality

$$\mathbb{E}\left[|X_t|^2\right] \le e^{-2tR_2}\mathbb{E}\left[|X_0|^2\right] + \frac{2R_1+d}{2R_2} \tag{24}$$

where R_1, R_2 are the constants in the assumption on b, d is the space dimension and t is an arbitrary time. It is straightforward to see that (24) implies (23), 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,

$$|X_t|^2 = |X_0|^2 + \int_0^t 2\langle X_s, b(X_s)\rangle\,ds + \int_0^t 2\langle X_s, dW_s\rangle + \mathrm{Tr}(I)\,t,$$

where $Tr(I) = d$ is the trace of the identity matrix, the quadratic variation in the Ito formula.

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, $\mathbb{E}\int_0^t 2\langle X_s, dW_s\rangle = 0$. Then

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

This identity and the fact that the function $s \mapsto \mathbb{E}\langle X_s, b(X_s)\rangle$ is continuous, imply that the function $t \mapsto \mathbb{E}\left[|X_t|^2\right]$ is in C^1; moreover, recalling the dissipativity condition (3), we have

$$\frac{d}{dt}\mathbb{E}\left[|X_t|^2\right] = 2\mathbb{E}\langle X_t, b(X_t)\rangle + d \le -2R_2\mathbb{E}\left[|X_t|^2\right] + 2R_1 + d.$$

This implies

$$\begin{aligned}\mathbb{E}\left[|X_t|^2\right] &\le e^{-2R_2t}\mathbb{E}\left[|X_0|^2\right] + \int_0^t e^{-2R_2(t-s)}\left(2R_1+d\right)ds \\ &\le e^{-2R_2t}\mathbb{E}\left[|X_0|^2\right] + \frac{2R_1+d}{2R_2}.\end{aligned}$$

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

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

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

$$\mathbb{E}\left[\left|X_{t\wedge\tau_R}\right|^2\right] \le \mathbb{E}\left[|X_0|^2\right] + \mathbb{E}\int_0^t 1_{\{s\le\tau_R\}} 2\langle X_s, b(X_s)\rangle\, ds + \mathrm{Tr}(I)\, t$$
$$\le \mathbb{E}\left[|X_0|^2\right] + \mathbb{E}\int_0^t 1_{\{s\le\tau_R\}} \left(2R_2|X_s|^2 + 2C_1\right) ds + \mathrm{Tr}(I)\, t$$
$$\le \mathbb{E}\left[|X_0|^2\right] + \mathbb{E}\int_0^t \left(2R_2\left|X_{s\wedge\tau_R}\right|^2 + 2C_1\right) ds + \mathrm{Tr}(I)\, t.$$

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

$$\mathbb{E}\left[\left|X_{t\wedge\tau_R}\right|^2\right] \le \left(\mathbb{E}\left[|X_0|^2\right] + 2TC_1 + Td\right) e^{2R_2T} =: C$$

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

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

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

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 19. *For every $h > 0$, $\mathcal{L}_h$ is bounded linear from $L^1 \to C^1$; in particular there exists $C, C_2 > 0$ (depending on h) such that*

$$\|\mathcal{L}_h f\|_{C^1} \le C\,\|f\|_{L^1}. \tag{25}$$

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

$$\|\mathcal{L}_h f\|_{BV_2} \le C_2\,\|f\|_{L^1_2}. \tag{26}$$

Proof. The estimate (25) follows by the definition of the transfer operator (14) and the estimate on its derivatives by the results of Menotti–Pesce–Zhang i.e. item 2 of Theorem 7. The estimate (26) follows from (19) considering that $\|\mathcal{L}_h(f)\|_{L^1_2}$ is estimated by Lemma 18 and for $f \in C^1$ it holds

$$\|f\|_{osc} = \sup_{\epsilon\le 1}\epsilon^{-1}\int_{\mathbb{R}^d} osc(f, B_\epsilon(x))\,d\psi(x) \le 2\|f\|_{C^1}\int_{\mathbb{R}^d} d\psi(x) \le 2\|f\|_{C^1} \tag{27}$$

since ψ is a probability measure. Finally by (25) we can bound $\|f\|_{osc(\mathbb{R}^d)}$ by the L^1 norm and then by the L^1_2 norm of f. □

By Lemmas 18 and 19 we get

Lemma 20. *Given $h > 0$, there exist constants $A, B > 0$ and $\lambda \in (0, 1)$ (also depending on h) such that*

$$\left\| \mathcal{L}_h^n f \right\|_{BV_2} \leq A\lambda^n \left\| f \right\|_{BV_2} + B \left\| f \right\|_{L^1} \tag{28}$$

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

Proof. By Lemma 19

$$\left\| \mathcal{L}_h^n f \right\|_{BV_2} \leq C_2 ||\mathcal{L}_h^{n-1} f||_{L^1_2} \tag{29}$$

and by Lemma 18

$$||\mathcal{L}_h^{n-1} f||_{L^1_2} \leq A\lambda^{n-1} \left\| f \right\|_{L^1_2} + B \left\| f \right\|_{L^1} . \tag{30}$$

Putting together these two inequalities we get the statement.[4] □

3.5. Regularization for the perturbed operators. In this section we continue showing the regularization properties of the trasfer operators. We prove a uniform Lasota Yorke inequality for the perturbed operators $\mathcal{L}_{h,n}$ acting on strong and weak spaces BV_2, L^1_2. Again the time evolution for each operator is set to the fixed $h > 0$ also considered in Sect. 3.4.

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

$$||\mathcal{L}_{h,n}^m f||_{BV_2} \leq A'\lambda^m ||f||_{BV_2} + B'||f||_{L^1}.$$

The proof of the proposition above is based on the following preliminary result, which we spell out also for a later usage

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

$$||\mathcal{L}_{h,n}^m f||_{L^1_2} \leq A\lambda^m ||f||_{L^1_2} + B||f||_{L^1}. \tag{31}$$

Proof. Let us begin by restricting ourselves to $f \geq 0$. We let $\mathcal{L}_{h,n} f = 1_{B_n^c} \mathcal{L}_h(f)$ so that, by construction, $\mathcal{L}_h f = \mathcal{L}_{h,n} f + 1_{B_n} \mathcal{L}_h(f)$.

Fix n and let $g_m = \mathcal{L}_h^m f - \mathcal{L}_{h,n}^m f$. Since $f \geq 0$ and clearly also $g_m \geq 0$ we have that

$$\|\mathcal{L}_{h,n}^m f\|_{L^1_2} = \int |\mathcal{L}_{h,n}^m f| \rho_2 dx \leq \int (\mathcal{L}_{h,n}^m f + g_m) \rho_2 dx = ||\mathcal{L}_h^m f||_{L^1_2},$$

by Lemma 18 we have then

$$||\mathcal{L}_{h,n}^m f||_{L^1_2} \leq A\lambda^m ||f||_{L^1_2} + B||f||_{L^1}. \tag{32}$$

[4] Note that, while proving Lemmas 19 and 18, we obtained that ρ_2 is, essentially, a Lyapunov function in L^1_2; thus some of the previous proofs and of the following ones could be rewritten following this viewpoint. We thank the anonymous referee for such remark.

Now let us remove the restriction $f \geq 0$ and decompose $f = f^+ - f^-$ into its positive and negative part. By homogeneity of the norms, addivity of the integral, applying (32) twice gives

$$||\mathcal{L}_{h,n}^m f||_{L_2^1} \leq 2A\lambda^m ||f||_{L_2^1} + 2B||f||_{L^1}.$$

□

Proof of Proposition 21. The BV_2 norm is the sum of the oscillation part and of the L_2^1 part. For the latter we use Proposition 22. For the oscillation we integrate [22, Proposition 3.2 (ii)] with respect to our norm[5] Namely

$$\|1_{B_n^c}\mathcal{L}_h f\|_{osc} \leq \sup_{0<\eta\leq 1} \frac{1}{\eta}\int osc(\mathcal{L}_h f, B_n^c \cap B_\eta(x))\mathbf{1}_{B_n^c}(x)d\psi(x) + \tag{33}$$

$$\sup_{0<\eta\leq 1} \frac{1}{\eta}\int 2\left[\sup_{B_\eta(x)\cap B_n^c}|\mathcal{L}_h f|\right]\mathbf{1}_{B_\eta(B_n)\cap B_\eta(B_n^c)}(x)d\psi(x), \tag{34}$$

where $B_\eta(x)$ is a ball centered at x and with radius η and given a set S, we denote $B_\eta(S) = \{x; \text{dist}(x, S) \leq \eta\}$. There are now two cases. We suppose first than $\eta < e^{-u_n}$; in this case the only points x contributing to the second integral in the previous inequality, are those belonging to a 2η-closed neighborhood of the boundary of the ball B_n; we call $S_{n,1}$ such an annulus. The Lebesgue measure of $S_{n,1}$ will be bounded by a constant $\tilde{C}$ (depending on d) times η.[6] In the second case, $\eta \geq e^{-u_n}$ only the points belonging to $S_{n,2} := B_{2\eta}(x_0)$ 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 $\sup_{B_\eta(x)\cap B_n^c}|\mathcal{L}_h f|$ will be bounded using Propositions 17 and Lemma 19; therefore by calling $\tilde{\mathcal{K}}$ the closed ball of radius 2 centered at the point x_0, which contains both $S_{n,1}$ and $S_{n,2}$ and by using Proposition 17, we continue to bound the quantity in (34) as

$$\sup_{0<\eta\leq 1} \frac{2}{\eta}||\mathcal{L}_h f||_{L^\infty(\tilde{\mathcal{K}})} \sup_{i=1,2} \psi(S_{n,i}) \leq 2\hat{C}\tilde{C}||f||_{L_2^1}||\psi'||_{L^\infty}, \tag{35}$$

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 (20), depending on ψ and on its strictly positive infimum over $\tilde{\mathcal{K}}$, and the constant entering formula (26).

The right-hand side of (33) is bounded by $\|\mathcal{L}_h f\|_{osc}$ and the latter is again bounded as in (26). Therefore we get:

$$||1_{B_n^c}\mathcal{L}_h f||_{osc} \leq (\hat{C} + 2\hat{C}\tilde{C}||\psi'||_{L^\infty})||f||_{L_2^1}.$$

---

5 Proposition 3.2 (ii) in [22] says that if $f \in L^\infty$ is a real function, $a > 0$ and S is a Borel subset, then

$$\text{osc}(f\mathbf{1}_S, B_a(\cdot)) \leq \text{osc}(f, S \cap B_a(\cdot))\mathbf{1}_S(\cdot) + 2\left[\text{esssup}_{B_a(\cdot)\cap S}|f|\mathbf{1}_{B_a(S)\cap B_a(S^c)}(\cdot)\right]\ (*).$$

Although f is required to be in L^∞, a quick inspection at the proof shows that f could be locally L^∞, which is our case. By integrating $(*)$ with respect to ψ we obtain (34).

6 The volume of $S_{n,1}$ 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.

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

$$||\mathcal{L}_{h,n}^m f||_{osc} \le ||\mathcal{L}_{h,n}\mathcal{L}_{h,n}^{m-1} f||_{osc} \le c^*||\mathcal{L}_{h,n}^{m-1} f||_{L^1_2}$$

and using Proposition 22 we continue as

$$||\mathcal{L}_{h,n}^m f||_{osc} \le c^* A\lambda^{m-1}||f||_{L^1_2} + c^* B||f||_{L^1}$$

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

$$||\mathcal{L}_{h,n}^m f||_{BV_2} \le A'\lambda^m||f||_{L^1_2} + B'||f||_{L^1}, \tag{36}$$

which implies the desired result. □

4. Rare Events Via Transfer Operator

For completeness, we recall the statement of a result due to Keller and Liverani [14] which is fundamental to our construction. Let $(\mathcal{B}, \|\cdot\|)$ be a Banach space and $\mathcal{B}^*$ its dual. Let $\mathcal{L}_\epsilon : \mathcal{B} \to \mathcal{B}$ be a family of uniformly bounded linear operators, where $\epsilon \in E$, and E is the interval $E = (0, \overline{\epsilon}]$ for some $\overline{\epsilon} > 0$.

4.1. Perturbative hypotheses.

R1 The operators $\mathcal{L}_\epsilon$, $\epsilon \in E$, must satisfy the spectral decomposition

$$\lambda_\epsilon^{-1}\mathcal{L}_\epsilon = \varphi_\epsilon \otimes \nu_\epsilon + Q_\epsilon \tag{37}$$

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

$$\sum_{n=0}^{\infty} \sup_{\epsilon \in E} \|Q_\epsilon^n\| < \infty. \tag{38}$$

Moreover, $\mathcal{L}_\epsilon\varphi_\epsilon = \lambda_\epsilon\varphi_\epsilon$, $\nu_\epsilon\mathcal{L}_\epsilon = \lambda_\epsilon\nu_\epsilon$, $\nu_\epsilon(\varphi_\epsilon) = 1$, $\nu_\epsilon Q_\epsilon = 0$, $Q_\epsilon(\varphi_\epsilon) = 0$. We also require that $\nu_0(\varphi_\epsilon) = 1$ and

$$\sup_{\epsilon \in E} \|\varphi_\epsilon\| < \infty. \tag{39}$$

R2 When ε is small, $\mathcal{L}_\varepsilon$ is a small perturbation of $\mathcal{L}_0$, in the following sense:

$$\pi_\epsilon := \sup_{f \in \mathcal{B}, \|f\| \le 1} |\nu_0((\mathcal{L}_0 - \mathcal{L}_\epsilon)(f))| \to 0, \ \epsilon \to 0.$$

R3 We now set $\Delta_\epsilon := \nu_0((\mathcal{L}_0 - \mathcal{L}_\epsilon)(\varphi_0))$. Then we require that there is $C \ge 0$ such that for each $\epsilon \in E$

$$\pi_\epsilon \|(\mathcal{L}_0 - \mathcal{L}_\epsilon)\varphi_0\| \le C\,|\Delta_\epsilon|.$$

R4 Let us consider the following quantities

$$q_{k,\varepsilon} := \frac{\nu_0((\mathcal{L}_0 - \mathcal{L}_\varepsilon)\mathcal{L}_\varepsilon^k(\mathcal{L}_0 - \mathcal{L}_\varepsilon)(\varphi_0))}{\Delta_\epsilon}.$$

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

$$\lim_{\epsilon \to 0} q_{k,\varepsilon} = q_k, \tag{40}$$

and we define the *extremal index of the system* as

$$\theta = 1 - \sum_{k=0}^{\infty} q_k. \tag{41}$$

Under these assumptions, the main result of [14, Sect. 2] establishes

Proposition 23. *If $R1 - R4$ are satisfied then*

$$\lambda_\epsilon = 1 - \theta\Delta_\epsilon + o(\Delta_\epsilon). \tag{42}$$

4.2. 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 call as the *weak* when compared with the *strong* Banach space $\mathcal{B}$. To ease our path, we state a sequence of four assumptions, which are easier to verify, implying R1:

$R1^{(i)}$ The unitary ball of $\mathcal{B}$ is compact in $\mathcal{B}_w$. Moreover, $\forall f \in \mathcal{B}$, $\|f\|_w \leq \|f\|$ i.e. the weak norm is bounded by the strong norm. Last $\exists G \geq 0$ s.t.

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

$R1^{(ii)}$ We require that the spectral decomposition (37) holds for $\epsilon = 0$.

$R1^{(iii)}$ The operators $\mathcal{L}_\epsilon$ 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

$$\forall \epsilon \in E\ \forall f \in \mathcal{B}\ \forall n \in \mathbb{N} : \|\mathcal{L}_\epsilon^n f\| \leqslant D\alpha^n\|f\| + D\|f\|_w. \tag{43}$$

$R1^{(iv)}$ There exists an upper semi-continuous function $u_\epsilon : [0, \infty) \to [0, \infty)$, $u_\epsilon > 0$, such that

$$|||\mathcal{L}_0 - \mathcal{L}_\epsilon||| := \sup_{f \in \mathcal{B}, \|f\| \leq 1} \|(\mathcal{L}_0 - \mathcal{L}_\epsilon)(f)\|_w \leq u_\epsilon \to 0,\ \epsilon \to 0. \tag{44}$$

usually referred to as “closeness of the operators in triple norm".

By assuming the previous conditions, it follows from the main result of [13,15] that the *quasi-compactness condition* R1 holds for any $\epsilon \in E, \epsilon > 0$. Moreover, by Corollary (2), part 2 in [13], there will be $0 < \rho < 1$ such that for all ϵ and n we have

$$\|Q_\epsilon^n\| \leq K_*\rho^n, \tag{45}$$

where ρ is chosen in such a way the spectrum of $\mathcal{L}_0$ does not intersect the circle $|z| = \rho$, and the constant K_* depends on ρ and on the spectral gap of $\mathcal{L}_0$. Notice that (45) immediately implies (38). Finally (39) follows immediately from (43).

Condition R2 can be easily worked out when ν_0 is a measure because, in this case, we can write

$$\pi_\epsilon \leq \sup_{f \in \mathcal{B}, \|f\| \leq 1} \int |(\mathcal{L} - \mathcal{L}_\epsilon)(f)| d\nu_0. \tag{46}$$

If the weak norm is strong enough to bound this integral, condition $R1^{(iv)}$ implies condition R2. More precisely, if we have that there is some C such that

$$\int |(\mathcal{L} - \mathcal{L}_\epsilon)(f)| d\nu_0 \leq C \|(\mathcal{L}_0 - \mathcal{L}_\epsilon)(f)\|_w$$

then $R1^{(iv)}$ implies R2. We will see that in our case this holds.

4.3. Verifying the perturbative assumptions in our case. In this section we verify the perturbative assumptions needed to apply Proposition 23. We now check that the perturbative assumptions listed in Sects. 4.1 and 4.2 apply, for the family of perturbed operators $\mathcal{L}_\epsilon$ for $0 < \epsilon \leq 1$, defined with a slight abuse of notation by $\mathcal{L}_0 := \mathcal{L}_h$ and $\mathcal{L}_\epsilon := \mathcal{L}_{h, \lfloor \frac{1}{\epsilon} \rfloor}$ for $0 < \epsilon \leq 1$, acting on BV_2 as a strong space and L^1 as a weak space. Here $\mathcal{L}_h$ and $\mathcal{L}_{h,n}$ are the unperturbed and perturbed operators defined in Sect. 3. We recall that

$$(\mathcal{L}_{h,n} f)(x) = 1_{B_n^c}(x)(\mathcal{L}_h f)(x).$$

where B_n is the ball $B(x_0, \exp(-u_n))$.
Assumption $R1^{(i)}$ comes directly from the definitions of the norms given in Sect. 3.3, from the compact embedding result proved at Theorem 16 and finally from Lemma 10.
Assumption $R1^{(ii)}$. The compact immersion proved in Theorem 16 and the Lasota Yorke inequality proved in Lemma 20 for $\mathcal{L}_h$, allows us to apply the Ionescu-Tulcea-Marinescu theorem (see for instance [11]), and therefore the operator $\mathcal{L}_h$ has the following spectral decomposition $\mathcal{L}_h = \sum_i \upsilon_i \Pi_i + Q$, where all υ_i are eigenvalues of $\mathcal{L}_h$ 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 $\mathcal{L}_h$ and that there is no other peripheral eigenvalue. This is achieved by the strict positivity of the Markov kernel S_h (see (14)) provided by Theorem 7.[7]Notice that the projector Π_1 will be the linear functional ν_0 in the assumption R1.
Assumption $R1^{(iii)}$ is now verified in Proposition 21.
Assumption $R1^{(iv)}$ is verified in the following proposition

[7] Here is the argument. Consider the operator $\mathcal{L}_h f(y) = \int S_h(x, y) f(x) dx$, 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 $\mathcal{L}_h$. 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 $\mathcal{L}_h(h_1 - \hat{h}) = h_1 - \hat{h}$. Take now a sequence of monotonically increasing compact sets K_n such that the

Proposition 24. *The operator $\mathcal{L}_{h,n}$ is a small perturbation of $\mathcal{L}_h$, in the following sense: there is a monotone sequence $\pi_n \to 0$ such that*

$$||\mathcal{L}_{h,n} - \mathcal{L}_h||_{BV_2 \to L^1} \leq \pi_n.$$

Proof. We have $||(\mathcal{L}_{h,n} - \mathcal{L}_h) f||_{L^1} = \|1_{B_n}(\mathcal{L}_h f)\|_{L^1} \leq \text{Leb}(B_n)||\mathcal{L}_h f||_\infty$. Notice that by Lemma 19 the C^1 norm of $\mathcal{L}_h f$, and therefore its L^∞ norm is bounded by $C_h\|f\|_{L^1}$. We could finally take $\pi_n = C_h\text{Leb}(B_n)$, since $\|f\|_{L^1} \leq \|f\|_{L^1_2} \leq \|f\|_{BV_2}$. □

Since in our case ν_0 is the integral with respect to the Lebesgue measure, and the weak norm is the L^1 norm, $R1^{(iv)}$ implies R2 as remarked at the end of Sect. 4.2.
Assumption R3 in our setting can be stated in the following way:

Proposition 25. *Let f_0 denote the invariant density of the invariant measure μ; then there exists a constant C' such that*

$$\pi_n\|(\mathcal{L}_h - \mathcal{L}_{h,n})f_0\|_{BV_2} \leq C'\mu(B_n).$$

We saw above that π_n could be taken as $C_h\text{Leb}(B_n)$; moreover by Proposition 21 the quantity $\|(\mathcal{L}_h - \mathcal{L}_{h,n})f_0\|_{BV_2} = \|1_{B_n}\mathcal{L}_h f_0\|_{BV_2}$ is bounded by a constant $\tilde{K}$. Therefore $\pi_n\|(\mathcal{L}_h - \mathcal{L}_{h,n})f_0\|_{BV_2} \leq C_h\tilde{K}\text{Leb}(B_n) \leq C_h\tilde{K}\frac{\mu(B_n)}{\inf_{B_n} f_0}$, which gives the desired result since the density f_0 is strictly positive on any bounded domain.[8]
Assumption R4 needs some more work. The next proposition will show that all the quantities q_k defined in the Assumption R4 are equal to 0. In the present setting they are defined as the limit for $n \to \infty$ of the following quantities:

$$q_{k,n} = \frac{\int(\mathcal{L}_h - \mathcal{L}_{h,n})\mathcal{L}^k_{h,n}(\mathcal{L}_h - \mathcal{L}_{h,n})(f_0)dm}{\mu(B_n)} \tag{47}$$

where m is the Lebesgue measure on $\mathbb{R}^d$. Using the duality between $\mathcal{L}_t$ and P_t, Lemma 12 and Corollary 15, it is not difficult to show that

$$\sum_{k=0}^{\infty} q_k \leq 1,$$

however, in our case, it will be sufficient to prove the following

Footnote 7 continued
first verifies $K_1 \subset \Omega$, $\text{Leb}(K_1 \cap \Omega) > 0$ and $\bigcup_n K_n = \mathbb{R}^2$ (use the regularity of the Lebesgue measure). Then take $y \in K_n$; we get

$$h_1(y) - \hat{h}(y) \geq \int_{x \in K_n \cap \Omega} S_h(x, y)(h_1(x) - \hat{h}(x))dx \geq \int_{x \in K_1 \cap \Omega} S_h(x, y)(h_1(x) - \hat{h}(x))dx > 0.$$

Notice that the right hand side is strictly positive and independent of n. Then we get $\int h_1 dx > \int h_2 dx$, which is impossible. Call h^* the unique fixed point of $\mathcal{L}_h$. Consider now the peripheral spectrum of $\mathcal{L}_h$ which consists of a finite union of finite cyclic groups. For each of those eigenvalues call g one of the corresponding eigenvectors. There exists $k \geq 1$ such that 1 is the unique peripheral eigenvalue of $\mathcal{L}^k_h : \mathcal{L}^k_h g = g$. But $\mathcal{L}^k_h h^* = h^*$, and by repeating the proof above we see that $g = h^*$.

8 More precisely: $\inf_{x \in B} f_0(x) > 0$ where B is any bounded subset of $\mathbb{R}^d$. This directly follows from the fact that f_0 is invariant for the transfer operator, having a strictly positive kernel as described in Theorem 7.

Proposition 26. *For each $k \geq 0$ we have*

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

Proof. Let us introduce the function $f_{k,n} := \frac{\mathcal{L}^k_{h,n}(\mathcal{L}_h - \mathcal{L}_{h,n})(f_0)}{\mu(B_n)}$; we have

$$\begin{aligned}\lim_{n\to\infty} q_{k,n} &= \lim_{n\to\infty} \int (\mathcal{L}_h - \mathcal{L}_{h,n}) f_{k,n} dm = \lim_{n\to\infty} \int 1_{B_n} \mathcal{L}_h f_{k,n} dm \\ &\leq \lim_{n\to\infty} m(B_n) \|\mathcal{L}_h f_{k,n}\|_\infty.\end{aligned}$$

As in the proof of Proposition 24, by Lemma 19, the C^1 norm of $\mathcal{L}_h f_{k,n}$, and therefore its infinity norm, is bounded by $C\|f_{k,n}\|_{L^1}$ for some $C \geq 0$ depending on h. On the other hand $\|f_{k,n}\|_{L^1} = \int f_{k,n} dm \leq \frac{1}{\mu(B_n)} \int_{B_n} f_0 dm = 1$. Then we get $\lim_{n\to\infty} q_{k,n} = 0$. ◻

4.4. 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; we denote by B_n the ball $B\left(x_0, \exp\left(-u_n\right)\right)$. We write g for $g_{x_0} := -\log(d(x, x_0))$. We denote by $\mathcal{L}_h$ the unperturbed transfer operator and by $\mathcal{L}_{h,n}$ the perturbed one.

Rewriting (1), using the notation above, remembering that $t_k = kh$, we get

$$\begin{aligned}&\mathbb{P} \otimes \mu \left(\left\|X^x_{t_k} - x_0\right\| \geq \exp\left(-u_n\right) \text{ for every } k = 0, \ldots, n-1\right) \\ &\quad= \mathbb{P} \otimes \mu \left(X^x_{t_k} \in B^c_n \text{ for every } k = 0, \ldots, n-1\right) \\ &\quad= \int_{\Omega\times\mathbb{R}^d} 1_{B^c_n}\left(X^x_{t_0}(\omega)\right) \cdots 1_{B^c_n}\left(X^x_{t_{n-1}}(\omega)\right) d\omega d\mu(x) \\ &\quad= \int_{\mathbb{R}^d} \mathbb{E}\left[1_{B^c_n}\left(X^x_{t_0}\right) \cdots 1_{B^c_n}\left(X^x_{t_{n-1}}\right)\right] d\mu(x).\end{aligned} \tag{48}$$

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

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \mathbb{E}\left[1_{B^c_n}\left(X^x_{t_0}\right) \cdots 1_{B^c_n}\left(X^x_{t_{n-1}}\right)\right] d\mu(x) \in (0, 1). \tag{49}$$

By using Corollary 15 together with (17), after recalling that the invariant measure μ is absolutely continuous w.r.t to Lebesgue with density f_0, we are finally able to write the distribution of the maxima in an operator-like way as

$$\begin{aligned}&\mathbb{P} \otimes \mu \left(\left(\max_{k=0,\ldots,n-1} g_{x_0}\left(X^x_{t_k}\right)\right) \leq u_n\right) \\ &\quad= \int \mathbb{E}\left[1_{B^c_n}(X^x_{t_0}) \cdots 1_{B^c_n}(X^x_{t_{n-1}}) 1(X^x_{t_{n-1}})\right] f_0(x) dx \\ &\quad= \int \left(P_{t_0,n} \circ P_{t_1-t_0,n} \circ \cdots \circ P_{t_n-t_{n-1},n}\right)(1)(x) f_0(x) dx = \int \mathcal{L}^n_{h,n} f_0 dx\end{aligned} \tag{50}$$

where $\mathcal{L}^n_{h,n}$ denotes the n-th power of the operator $\mathcal{L}_{h,n}$.

Now we apply Proposition 23 to the transfer operator $\mathcal{L}_h$ and to its perturbations $\mathcal{L}_{h,n}$. We consider as a strong space the space $\mathcal{B}(\mathbb{R}^d) = BV_2$ and L^1 as a weak space. The assumptions of Sect. 4.1 are verified by $\mathcal{L}_h$ and by $\mathcal{L}_{h,n}$ thanks to the sufficient conditions quoted in Sect. 4.3; therefore we have the following spectral decomposition for the perturbed operator $\mathcal{L}_{h,n}$:

$$\lambda_{h,n}^{-1}\mathcal{L}_{h,n} = f_{h,n} \otimes \mu_{h,n} + Q_{h,n} \tag{51}$$

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 strictly less than 1. Moreover $\mathcal{L}_{h,n} f_{h,n} = \lambda_{h,n} f_{h,n}$ and $\mu_{h,n}\mathcal{L}_{h,n} = \lambda_{h,n}\mu_{h,n}$.

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

$$\mathcal{L}_{h,n} g = \lambda_{h,n} f_{h,n} \langle \mu_{h,n}, g\rangle + \lambda_{h,n} Q_{h,n}(g); \tag{52}$$

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 $\int (\mathcal{L}_{h,n}^n f_0)(x)dx$ by plugging (52) in it. Since we have a direct sum decomposition of our operator, we can iterate and get

$$\int (\mathcal{L}_{h,n}^n f_0)(x)\, dx = \lambda_{h,n}^n \langle \mu_{h,n}, f_0\rangle + \lambda_{h,n}^n \int (Q_{h,n}^n f_0)(x)\, dx. \tag{53}$$

Now, 1 is the largest unique eigenvalue of the unperturbed operator $\mathcal{L}_h$ and by proposition 23 and proposition 26 we have $\theta = 1$, therefore:

$$\lambda_n = 1 - \mu(B_n) + o(\mu(B_n)). \tag{54}$$

Then by substituting (54) in (53) we have

$$\begin{aligned}\int (\mathcal{L}_{h,n}^n f_0)(x)\, dx &= e^{n \log(1-\mu(B_n)+o(\mu(B_n)))} \left[\langle \mu_{h,n}, f_0\rangle + \int Q_{h,n}^n f_0\, dx\right] \\ &= e^{-n\mu(B_n)+no(\mu(B_n))} \left[\langle \mu_{h,n}, f_0\rangle + \int Q_{h,n}^n f_0\, dx\right].\end{aligned} \tag{55}$$

Let us now recall that by [14, Lemma 6.1] we have $\langle \mu_n, f_0\rangle \to 1$. Moreover by (45) we get a uniform exponential convergence to zero of

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

where the constant on the right hand side is given in terms of the spectral data of $\mathcal{L}_h$, see discussion after the inequality (45). 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

$$n\, \mu(B_n) \to \tau. \tag{56}$$

Thus

$$\int \mathcal{L}_{h,n}^n f_0\, dx \to e^{-\tau} \tag{57}$$

proving (9). □

5. Poisson Statistics

Proof of Theorem 6. For this proof, similarly to the proof of Theorem 2 we will apply Proposition 23. We start by computing the characteristic function of the random variable $S_n = \sum_{i=0}^{n-1} 1_{B_n}(X_{ih}^x)$:

$$\Phi_n(s) = \int e^{isS_n} d\mathbb{P} d\mu = \int \mathbb{E}(e^{isS_n}) f_0 dx, \tag{58}$$

where f_0 is the density of μ. We then introduce the perturbed operators for $f \in BV_2$:[9]

$$\mathcal{L}_{h,n} f(x) = e^{is1_{B_n}(x)} \mathcal{L}_h f(x), \tag{59}$$

$$\mathcal{P}_{h,n} f(x) = \mathbb{E}(e^{is1_{B_n}(X_h^x)} f(X_h^x)) \tag{60}$$

By using (14) we get, for $f \in BV_2(\mathbb{R}^d)$, $g \in L^\infty$:

$$\int \mathbb{E}(e^{is1_{B_n}(X_h^x)} g(X_h^x)) f(x) dx = \int \mathcal{P}_{h,n} g(x) f(x) dx = \int \mathcal{L}_{h,n} f(x) g(x) dx. \tag{61}$$

We then observe that Lemma 13 holds for the new operator $\mathcal{P}_{h,n}$ just by replacing the characteristic function 1_{B_n} with $e^{is1_{B_n}}$.

Using this and the equalities (61), we have

$$\Phi_n(s) = \int e^{isS_n} d\mathbb{P} d\mu = \int \mathcal{L}_{h,n}^n f_0(x) dx. \tag{62}$$

We now apply Theorem 23 to the operators $\mathcal{L}_{h,n}$, using BV_2 and L^1 as a strong and weak space. In order to do this, we check the assumptions $(R1)^{(i)}, \ldots, (R1)^{(iv)}$, (R3), (R4) as done for the proof of Theorem 2. We will omit some detail, as the proofs are similar to what is done in Sect. 4.3.

- $(R1)^{(i)}, \ldots, (R1)^{(iii)}$ can be verified similarly as done for the proof of Theorem 2. Some more work in necessary to verify the Lasota–Yorke inequality, which follows by a straightforward application of Propositions 21 and 22. To estimate the oscillation seminorm of $e^{is1_{B_n}} \mathcal{L}_t f$ in Proposition 21, by using formula[10], we get

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

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

The first piece (I) on the right hand side is $\|\mathcal{L}_h f\|_{\mathrm{osc}(\mathbb{R}^d)}$ and is bounded as in (26). For the second one, we have to proceed as in the proof of Proposition 21 by splitting it into two cases; similar arguments gives the bound:

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

This and Lemma 18 will finally give us the desired Lasota-Yorke inequality.

[9] We now require that our functions take values in $\mathbb{C}$; in this case the oscillation is defined as $\mathrm{osc}(f, S) = \mathrm{esssup}|f(x) - f(y)|, x, y \in S$. With this definition the function spaces and the norms considered extend straightforwardly to the complex case. We will keep denoting these spaces as L^1_α and BV_α also in the complex valued case.

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

- $(R1)^{(iv)}$ Using Lemma 19 we have for $f \in BV_2$

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

$$2\text{Leb}(B_n)C_h\|f\|_{L^1} \leq 2\text{Leb}(B_n)C_h\|f\|_{BV_2},$$

with $\pi_n = 2\text{Leb}(B_n)C_h$.
- (R3) The closeness of the two operators is also quantified by

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

Since the density f_0 is locally bounded away from zero (see Footnote 8) and by using the same arguments which allowed us to bound the BV_2 norm of $e^{is1_{B_n}}\mathcal{L}_h$ in $(R1)^{(iii)}$, we get that

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

- (R4) The quantities q_k[11] associated to this perturbation will therefore have the form

$$q_k = \lim_{n\to\infty} \frac{1}{\Delta_n} \int (\mathcal{L}_h - \mathcal{L}_{h,n})\mathcal{L}_n^k(\mathcal{L}_h - \mathcal{L}_{h,n})(f_0)dx, \quad (63)$$

provided the limits exists. By repeating verbatim the proof of Proposition 26 it is easy to show that the limit defining q_k exists and it tends to 0 when $n \to \infty$.

We recall that by assumption, 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 $\mathcal{L}_{h,n}$, see Proposition 23

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

Since the assumptions of Sect. 4.1 are verified by $\mathcal{L}_h$ and by $\mathcal{L}_{h,n}$ thanks to the sufficient conditions quoted in Sect. 4.3, we could repeat the steps from Eqs. (52)–(55) showing that the leading term in the growth 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

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

□

[11] We use the same symbol as for the q_k introduced in item R4 in Sect. 4.1

Acknowledgement P.G. and S.V. acknowledge the support of the SNS-CRM Ennio de Giorgi and of UniCredit Bank R&D group for financial support through the "Dynamics and Information Theory Institute" at Scuola Normale Superiore. P.G. was partially supported by INDAM - GNFM project "Deterministic and stochastic dynamical systems for climate studies" and by PRIN 2022NTKXCX. The research of S.V. was supported by the Laboratoire International Associé LIA LYSM, of the French CNRS and INdAM (Italy). S.V. 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, where part of this research was performed. S.V. was also supported by the project MATHAmSud TOMCAT 22-Math-10, N. 49958WH, CNRS and MEAE. S.V. also thanks the CNRS International Research Laboratory IRL2032 FAMSI, France Australia, at the Australian National University in Canberra, where this work was completed. 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. This research of F.F. has been funded by the European Union, ERC NoisyFluid, No. 101053472. We thank the anonymous referees for the patience and the care they reserved to our paper.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A Dependence on h of the Limit Law

Our main result, Theorem 2, is formulated assuming the time discretization of step h by considering X_{nh}. However, note that the right hand side limiting law of the statement, Eq. (9), does not depend on h. Let us notice that the limit for $h \to 0$, of the above results does not provide more information and should not be thought as a way to get a continuous time version of our result: this does not come as a surprise as the passage from discrete to continuous time might require another not obvious rescaling (see also on the matter [20, Sect. 3.6]).

The next proposition illustrates the need for a different scaling in the continuous time case by showing that if the general assumptions and the sequence of thresholds u_n is taken as in Theorem 2, but the time is allowed to range continuously in the interval $[0, hn]$, then the values of the observable g considered along typical trajectories will overcome the threshold sequence almost certainly. Thus, in continuous time a different sequence of thresholds needs to be considered if one wants to characterize the typical way the extremal values of g change along the dynamics.

Proposition 27. *Let the datum be as in Theorem 2. It holds*

$$\lim_{n\to\infty} \mathbb{P}\otimes\mu\left(\left\{(\omega, x) : \max_{t\in[0,nh]} g\left(X_t^x\right) \le u_n\right\}\right) = 0.$$

Note, with respect to Theorem 2 that here $t \in [0, nh]$.

Proof. Let $M > 0$ be an integer. In the rest of the proof, we shorten $\mathbb{P} \otimes \mu(\{(\omega, x) : \ldots\}) := \mathbb{P} \otimes \mu(\ldots)$, characterizing the sets considered by the inequalities that define them. We have

$$\mathbb{P} \otimes \mu \left(\max_{t \in [0,nh]} g\left(X_t^x\right) \leq u_n \right) \leq \mathbb{P} \otimes \mu \left(\max_{k=0,\ldots,nM} g\left(X_{k\frac{h}{M}}^x\right) \leq u_n \right)$$

hence

$$\limsup_{n\to\infty} \mathbb{P} \otimes \mu \left(\max_{t \in [0,nh]} g\left(X_t^x\right) \leq u_n \right) \leq \lim_{n\to\infty} \mathbb{P} \otimes \mu \left(\max_{k=0,\ldots,nM} g\left(X_{k\frac{h}{M}}^x\right) \leq u_n \right).$$

Now, if $n\mu_0\left(B\left(x_0, e^{-u_n}\right)\right) \to \tau$, then $nM\mu_0\left(B\left(x_0, e^{-u_n}\right)\right) \to \tau M$. Consider a sequence $\{v_j\}$ such that $v_{nM} = u_n$ for every n. We have

$$nM\mu_0\left(B\left(x_0, e^{-v_{nM}}\right)\right) \to \tau M.$$

One can define $\{v_j\}$ in such a way that $j\mu_0\left(B\left(x_0, e^{-v_j}\right)\right) \to \tau M$ as $j \to \infty$, not only along the subsequence $j_n = nM$. Thus, by the theorem (applied with $\frac{h}{M}$ in place of h)

$$\lim_{j\to\infty} \mathbb{P} \otimes \mu \left(\max_{k=0,\ldots,j} g\left(X_{k\frac{h}{M}}\right) \leq v_j \right) = e^{-\tau M}$$

and thus, for the subsequence $j_n = nM$,

$$\lim_{n\to\infty} \mathbb{P} \otimes \mu \left(\max_{k=0,\ldots,nM} g\left(X_{k\frac{h}{M}}\right) \leq v_{nM} \right) = e^{-\tau M}.$$

But $v_{nM} = u_n$. Hence

$$\limsup_{n\to\infty} \mathbb{P} \otimes \mu \left(\max_{t \in [0,nh]} g\left(X_t\right) \leq u_n \right) \leq e^{-\tau M}.$$

Since this holds true for every M, we complete the proof. □

References

1. Ambrosio, L., Miranda, M., Maniglia, S., Pallara, D.: Towards a theory of BV functions in abstract Wiener spaces. Phys. D: Nonlinear Phenom. **239**(15), 1458–1469 (2010). https://doi.org/10.1016/j.physd.2009.03.007
2. Atnip, J., Froyland, G., González-Tokman, C., Vaienti, S.: Compound Poisson statistics for dynamical systems via spectral perturbation (2023). http://arxiv.org/abs/2308.10798pdf
3. Atnip, J., Froyland, G., González-Tokman, C., Vaienti, S.: Thermodynamic formalism and perturbation formulae for quenched random open dynamical systems (2023). http://arxiv.org/abs/2307.00774pdf. To appear in dissertationes-mathematicae
4. Aytaç, H., Freitas, J.M., Vaienti, S.: Laws of rare events for deterministic and random dynamical systems. Trans. Am. Math. Soc. **367**(11), 8229–8278 (2015)
5. Caby, T., Faranda, D., Vaienti, S., Yiou, P.: On the computation of the extremal index for time series. J. Stat. Phys. **179**, 1666–1697 (2019). https://doi.org/10.1007/s10955-019-02423-z
6. Davis, R.A.: Maximum and minimum of one-dimensional diffusions. Stoch. Proc. Appl. **13**, 1–9 (1982)
7. Faranda, D., Ghoudi, H., Guiraud, P., Vaienti, S.: Extreme value theory for synchronization of coupled map lattices. Nonlinearity **31**(7), 3326–3358 (2018)
8. Friedman, A.: Stochastic differential equations and applications. In: Cecconi, J. (ed.) Stochastic Differential Equations, pp. 75–148. Springer, Berlin (2011)

9. Fukushima, M.: BV functions and distorted Ornstein Uhlenbeck processes over the abstract Wiener space. J. Funct. Anal. **174**(1), 227–249 (2000)
10. Ghil, M., Lucarini, V.: The physics of climate variability and climate change. Rev. Mod. Phys. **92**, 035002 (2020)
11. Hennion, H., Hervé, L.: Limit Theorems for Markov Chains and Stochastic Properties of Dynamical systems by quasicompactness. Lecture Notes in Mathematics, vol. 1766. Springer, Berlin (2001)
12. Keller, G.: Generalized bounded variation and applications to piecewise monotonic transformations. Z. Wahr. verw. Geb **69**, 461–478 (1985)
13. Keller, G., Liverani, C.: Stability of the spectrum for transfer operators. Ann. Sc. Norm. Super. Pisa **XXVIII**, 141–152 (1999)
14. Keller, G., Liverani, C.: Rare events, escape rates and quasistationarity: some exact formulae. J. Stat. Phys. **135**, 519–534 (2009)
15. Keller, G.: Rare events, exponential hitting times and extremal indices via spectral perturbation. Dyn. Syst.-Int. J. **27**, 11–27 (2012)
16. Klebaner, F.C.: Introduction to Stochastic Calculus with Applications. Imperial College Press, London (2012). (**English**)
17. Kuntz, A.: On extremes of multidimensional stationary diffusion processes in Euclidean norm (2022). https://mediatum.ub.tum.de/doc/1097695/1097695.pdf
18. Lucarini, V., Faranda, D., Moreira de Freitas, A.C.G.M., Milhazes de Freitas, J.M., Holland, M., Kuna, T., Nicol, M., Todd, M., Vaienti, S.: Extremes and Recurrence in Dynamical Systems, p. xi + 296. Wiley, Hoboken (2016). (**English**)
19. Pène, F., Saussol, B., Zweimüller, R.: Recurrence rates and hitting-time distributions for random walks on the line. Ann. Prob. **41**(2), 619–635 (2013)
20. Rootzen, H., Leadbetter, M.R.: Extremal theory for stochastic processes. Ann. Prob. **16**, 431–478 (1988)
21. Saussol, B.: Absolutely continuous invariant measures for multidimensional expanding maps. Etude statistique de systèmes dynamiques dilatants. In: Universit é de Toulon et du Var (1998) http://www.math.univ-brest.fr/perso/benoit.saussol/art/these.pdf
22. Saussol, B.: Absolutely continuous invariant measures for multidimensional expanding maps. Israel J. Math. **116**, 223–248 (2000)
23. Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Classics in Mathematics, p. xii+338. Springer, Berlin (2006). (**Reprint of the 1997 edition**)
24. Todd, M., Freitas, A., Freitas, J.M.: Hitting time statistics and extreme value theory. Probab. Theory Relat. Fields **147**, 675–710 (2010)
25. Williams, D.: Probability with Martingales. Cambridge Mathematical Textbooks, University Press, Cambridge (1991). (**English**)
26. Zhang, X., Menozzi, S., Pesce, A.: Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift. J. Differ. Equ. **272**, 330–369 (2021)

Communicated by R. Seiringer