Nonlinearity

PAPER

Existence of noise induced order, a computer aided proof

IOP Publishing | London Mathematical Society Nonlinearity

Nonlinearity **33** (2020) 4237–4276 https://doi.org/10.1088/1361-6544/ab86cd

Existence of noise induced order, a computer aided proof

Stefano Galatolo[1,3], Maurizio Monge[2] and Isaia Nisoli[2]

[1] Dipartimento di Matematica, Universita di Pisa, Via Buonarroti 1, Pisa - Italy
[2] Instituto de Matemática, Univ. Fed. do Rio de Janeiro, Av. Athos da Silveira Ramos 149, Bloco C Cidade Universitária, Rio de Janeiro - Ilha do Fundão - Brazil

E-mail: stefano.galatolo@unipi.it

Received 16 November 2019, revised 20 March 2020
Accepted for publication 6 April 2020
Published 20 July 2020

Abstract

We prove the existence of noise induced order in the Matsumoto–Tsuda model, where it was originally discovered in 1983 by numerical simulations. This is a model of the famous Belousov–Zhabotinsky reaction, a chaotic chemical reaction, and consists of a one dimensional random dynamical system with additive noise. The simulations showed that an increase in amplitude of the noise causes the Lyapunov exponent to decrease from positive to negative; we give a mathematical proof of the existence of this transition. The method we use relies on some computer aided estimates providing a certified approximation of the system's stationary measure in the L^1 norm. This is realized by explicit functional analytic estimates working together with an efficient algorithm. The method is general enough to be adapted to any piecewise differentiable dynamical system on the unit interval with additive noise. We also prove that the stationary measure varies in a Lipschitz way if the system is perturbed and that the Lyapunov exponent of the system varies in a Hölder way when the noise amplitude increases.

Keywords: Lyapunov exponent, noise induced order, random dynamics, computer aided proof, quantitative statistical stability, interval arithmetics, Belousov–Zhabotinsky reaction

Mathematics Subject Classification numbers: 37H99, 37M25, 65G30, 92E20, 37C30.

(Some figures may appear in colour only in the online journal)

[3] Author to whom any correspondence should be addressed.

1. Introduction

The 'noise induced order' phenomenon was discovered in numerical simulations and experiments regarding systems modelled by a deterministic dynamics perturbed by noise. The somewhat surprising phenomenon emerging is that the system appears to be chaotic for very small noise intensity, but when the intensity increases the system begins to have a less and less chaotic behaviour, passing from a positive Lyapunov exponent to a negative one. A similar behaviour was also found for other indicators of chaos. This paper however will focus only on the Lyapunov exponent.

The phenomenon was first discovered by numerical simulations in [24] in a system related to the famous Belousov–Zhabotinsky reaction (see figure 1) modelled by a one dimensional map perturbed by additive noise. Real experiments confirmed the existence of the phenomenon appearing in the model of the reaction (see [34], and also [11, 21, 25, 33, 35, 36], for some examples of related works).

Despite the impact that the discovery of such noise induced phenomena had in the nonlinear science and physical literature (more than 390 citations to [24] on Google Scholar at the moment of writing this paper) to the best of our knowledge there is no mathematical literature about noise induced order or rigorous proofs of its actual existence in nontrivial systems.

The mathematical study of this phenomenon is difficult because in the deterministic part of the dynamics (see figure 2) there is a coexistence of strongly expanding and strongly contracting regions and the prevalence of expanding or contracting behaviour for typical orbits is a consequence of the global structure of the dynamics. We remark that this phenomenon is one dimensional and inherently nonlinear, thus mathematically not much related to the noise induced stabilization studied in [2] and following papers. With the help of some computer aided estimates we prove that the global structure of this random system, allows expansion to prevail when the noise amplitude is very small, but the appearance of quite a large noise allows the contraction to prevail.

Our approach is based on the fact that the presence of the noise simplifies the functional analytic properties of the transfer operator associated to the system, smoothing out fine resolution details, and making it well approximable by a finite resolution and finite dimensional one. This makes a computer aided proof possible, letting the computer manage the complexity of the deterministic part of the system at a *finite* resolution scale and understanding the global structure of the dynamics. However, the computational power required to perform these computations in a naive way is out of the range of current computers. This is true mostly for proving that the Lyapunov exponent is positive in some case of very small noise range. Indeed small noise corresponds to high resolution needed in the study of the system. Because of this we had to find some mathematical clever way to perform the needed estimates, using different functional spaces. This is the main part of the mathematical work contained in what follows and can be applied to many other dynamical systems perturbed by noise. The algorithm we develop in this work was indeed already used in [4, 10] for the study of other dynamical systems with additive noise considered as models of certain phenomena in climate science and neuroscience (other applications to linear response appear in [12]).

It is known that naive computer simulations of chaotic systems may not be reliable in some case (see [8, 16–18, 20] for examples of misleading naive simulations and a general discussion on the problem). Beside the pure mathematical interest of a rigorously proved result and a rigorously certified estimate, the study of inherently reliable methods for the numerical study of chaotic dynamical systems is strongly motivated.

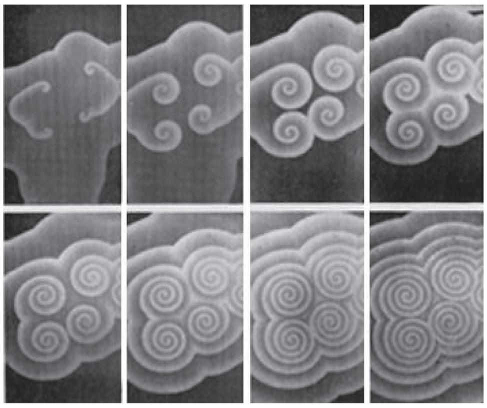

Figure 1. A sequence of pictures showing various stages of the evolution of the Belousov–Zhabotinsky reaction. Reprinted from [37], copyright (1973), with permission from Elsevier.

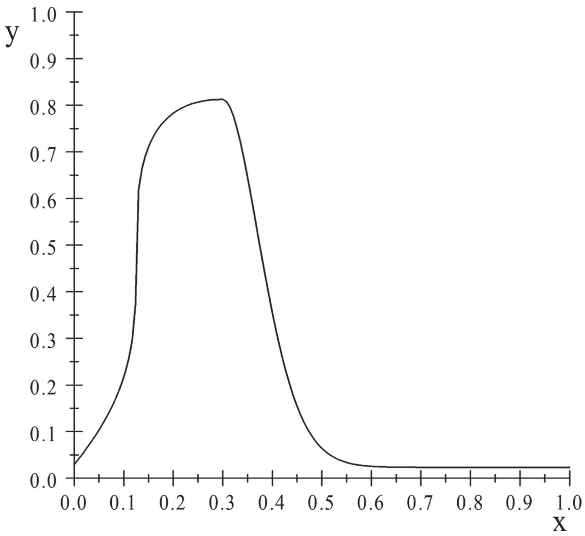

Figure 2. The map $T_{a,b,c}$.

Overview of the results. In this work we consider the model of the Belousov–Zhabotinsky reaction studied in [24] (see also [38] for explanations on how the model can be deduced from the chemical mechanism of the BZ reaction). This is a random dynamical system: a deterministic map with additive noise at each iteration. The deterministic part of the dynamics in the model is driven by a map $T_{a,b,c} : [0, 1] \to [0, 1]$ defined by

$$T_{a,b,c}(x) = \begin{cases} \left(a + \left(x - \frac{1}{8}\right)^{\frac{1}{3}}\right) \mathrm{e}^{-x} + b, & 0 \leqslant x \leqslant 0.3 \\ c\left(10xe^{\frac{-10x}{3}}\right)^{19} + b, & 0.3 < x \leqslant 1 \end{cases} \tag{1}$$

where

$$a \in 0.501\,319\,599\,371\,053\,047\,956\,980\,141\,736\,828\,203\,749\,380\,990\,114\,218\,225\,638\,827^{9}_{6}, \tag{2}$$

$$b \in 0.023\,288\,528\,303\,070\,320\,544\,781\,580\,440\,239\,187\,356\,699\,436\,480\,888\,526\,461\,231\,827\,398\,310\,225\,28^{213}_{158}, \tag{3}$$

$$c \in 0.121\,205\,692\,738\,975\,111\,744\,666\,848\,150\,620\,569\,782\,497\,212\,127\,938\,371\,936\,404\,761\,693\,002\,104\,361^{8}_{5}. \tag{4}$$

The graph of an example of $T_{a,b,c}$ is shown in figure 2. Following the *Inverval Arithmetics* formalism we represent intervals in the real line by subscript and superscript describing the decimal expansion of lower and upper bounds for x so that $x \in 0.klm^{xyz}_{tuv}$ means that x belongs to the interval $[0.klmtuv, 0.klmxyz]$. The choice of a, b, c follow the one made in [24], adding some more precision (see remark 5 for more details).

Remark 1. The interval arithmetics and its certified numerical methods (see [31] for an introduction) allow to obtain rigorous results as output of the computer aided estimates. Our computer aided estimates are implemented in this framework. We used SAGE [28] and the validated numerics packages shipped with it. (The interval package is a binding to MPFI [27].) Part of the numerical linear algebra was done using OpenCL [30] running on Nvidia graphic cards.

At each iteration of the map a *uniformly distributed* noise perturbation with span of size ξ is applied. Further details on the system are presented in section 2.

In the paper we prove that when the noise size ξ is contained in the interval $[\xi_1, 1/2]$ where $\xi_1 = \frac{8.73}{10^5}$ this random system has a unique ergodic absolutely continuous stationary measure μ_ξ (see proposition 46) and consider the associated *Lyapunov exponent*

$$\lambda_\xi := \int_0^1 \log|T'(x)|\mathrm{d}\mu_\xi.$$

We prove that the behaviour of λ_ξ in the system is similar to the one found by the numerical investigations of Matsumoto and Tsuda ([24]). *In particular there is a transition from positive to negative exponent as the noise amplitude increases.* We provide explicit examples of values for the noise amplitude having positive and negative Lyapunov exponent. In particular, the findings of the present work applied to the Belousov–Zhabotinsky model defined above allow to state the following

Theorem 2. *Let λ_ξ be the Lyapunov expontent of the system defined above*[4] *with noise of size ξ. For each $\alpha < 1, \lambda_\xi$ is $\alpha-$Hőlder continuous as a function of ξ when $\xi \in [\frac{8.73}{10^5}, 1/2]$; furthermore for $\xi_1 = \frac{8.73}{10^5}$ and $\xi_2 = \frac{8.60}{10^3}$ it holds*

*I*1 *the Lyapunov exponent $\lambda_{\xi_1} \in [\frac{8.365}{10^2}, \frac{8.917}{10^2}]$, hence it is rigorously certified to be positive;*

*I*2 *the Lyapunov exponent $\lambda_{\xi_2} \in [\frac{-6.036\,02}{10}, \frac{-6.035\,36}{10}]$, hence it is rigorously certified to be negative.*

[4] For every choice of the coefficients a, b, c of (1) as in (2)–(4).

Therefore, the system exhibits noise induced order.

We refer to section 6 for the results proving the change of sign for the Lyapunov exponent, while we refer to section 7 (see corollary 43) for the Hőlder regularity of the Lyapunov exponent. In the paper, we estimate similarly the Lyapunov exponent for many other values of the noise size. The results are reported in table 1 and described in figure 4 by a graph.

The method of proof and certification of these results relies on the approximation of the stationary measure μ_ξ of the system up to a certified error in the L^1 norm[5]. Using this approximation we can compute the Lyapunov exponent of the system up to a small certified error. The methods used however are general enough to be applied to any system formed by piecewise differentiable maps of the interval perturbed by additive noise having a bounded variation distribution (see sections 2 and 3.3 for a more detailed discussion).

Remark 3. Although theorem 2 certifies the existence of the noise induces order in the system, it does not give an intuitive explanation of 'why' this phenomenon occurs.

An argument which can give an approximate explanation of the existence of the noise induced order in a class of systems similar to the one studied in this paper might be the following: the map $T_{a,b,c}$ is a map which is *near* to a unimodal map T of Misiurewicz type. Those maps have positive Lyapunov exponent. If the map T is statistically stable under the adding of a small amount of noise (see e.g. [3]) we can expect that also in the case of small noise the resulting random system will have positive Lyapunov exponent.

On the other hand suppose that $\int \log|T'|\mathrm{d}m < 0$ where the integral is made with respect to the Lebesgue measure (this condition can be easily verified for a given map, and it is verified for example if the 'critical point' of the unimodal map is flat enough). When the noise is large enough we can expect the stationary measure of the system to converge to the uniform distribution, hence to the Lebesgue measure, letting the Lyapunov exponent of the resulting random system to become negative.

This heuristic argument can be made rigorous in interesting classes of systems, as it is done in the very recent work [26] and proving the existence of some kind of noise induced order in those systems.

However we remark that in the model of the Belousov Zhabotinsky reaction we consider, the adding of a large noise would have no physical sense, bringing the model's parameters outside its natural range. As a matter of fact, in the paper [24] the transition occurs for small values of the noise size, where the stationary measure (see figure 3) is far from the uniformly distributed one, hence this idea is far from explaining this transition. Furthermore the map considered in our paper and in the model is not necessarily exactly of Misurevicz type and indeed this is not used in our proof.

Plan of the paper. In section 2 we describe the systems under study and we introduce the technique we use for the approximation of the transfer operator. The techniques leading to the proof of theorem 2 are explained in a sequence of settings of decreasing generality where the needed assumptions on the system are listed (see settings 14 and 19). In section 3 we describe how to find an explicit bound on the approximation error for the computation of the stationary measure. In subsection 3.3 we describe an efficient procedure which exploits

[5] The computation of stationary measures for dynamical systems perturbed by noise was approached from different points of view in [19] where the convergence (without effective bounds on the approximation error) of an approximation scheme based on Fourier analysis was proved for certain classes of maps, and in [7], where the computability of the stationary measure up to a given error was considered in an abstract framework, giving bounds on the computational complexity of the problem. The rigorous computation of the stationary measure for expanding and contracting iterated function systems is considered in [15].

Table 1. The input and output of our computer aided estimates. In particular the first column shows the size of the noise, the fifth column shows the size of the approximation grid, the last column shows intervals enclosing the exact value for the Lyapunov exponent related to the noise size in the first column. See section 6 for explanations on all the other values.

ξ, noisesize	δ_{contr}	α_{contr}	n_{contr}	δ	α	$\sum_{i=0}^{n_{\text{contr}}} C_i$	*A priori* L^1 err. on measure	δ_{est}	Refined L^1 err. on measure	rig. estimate on Lyapunov exponent
0.860×10^{-2}	2^{-18}	0.023	56	2^{-27}	0.044	29.54	0.445×10^{-4}	2^{-14}	0.246×10^{-5}	$-6.03^{536}_{602} \times 10^{-1}$
0.785×10^{-2}	2^{-18}	0.017	58	2^{-27}	0.039	29.20	0.479×10^{-4}	2^{-14}	0.257×10^{-5}	$-5.499^{221}_{901} \times 10^{-1}$
0.721×10^{-2}	2^{-18}	0.013	60	2^{-27}	0.037	28.95	0.517×10^{-4}	2^{-14}	0.268×10^{-5}	$-4.985^{026}_{749} \times 10^{-1}$
0.661×10^{-2}	2^{-18}	0.015	60	2^{-27}	0.04	28.75	0.562×10^{-4}	2^{-14}	0.282×10^{-5}	$-4.467^{104}_{862} \times 10^{-1}$
0.606×10^{-2}	2^{-18}	0.022	54	2^{-27}	0.05	28.44	0.612×10^{-4}	2^{-14}	0.296×10^{-5}	$-3.95^{558}_{636} \times 10^{-1}$
0.556×10^{-2}	2^{-18}	0.015	58	2^{-27}	0.046	28.77	0.672×10^{-4}	2^{-14}	0.311×10^{-5}	$-3.462^{063}_{884} \times 10^{-1}$
0.509×10^{-2}	2^{-18}	0.014	60	2^{-27}	0.048	29.19	0.747×10^{-4}	2^{-14}	0.330×10^{-5}	$-2.97^{272}_{360} \times 10^{-1}$
0.466×10^{-2}	2^{-18}	0.019	59	2^{-27}	0.056	29.56	0.832×10^{-4}	2^{-14}	0.351×10^{-5}	$-2.52^{476}_{568} \times 10^{-1}$
0.426×10^{-2}	2^{-18}	0.032	53	2^{-27}	0.072	29.70	0.930×10^{-4}	2^{-14}	0.374×10^{-5}	$-2.10^{322}_{419} \times 10^{-1}$
0.391×10^{-2}	2^{-18}	0.025	57	2^{-27}	0.07	30.68	0.104×10^{-3}	2^{-14}	0.400×10^{-5}	$-1.74^{370}_{474} \times 10^{-1}$
0.359×10^{-2}	2^{-18}	0.023	60	2^{-27}	0.074	31.73	0.118×10^{-3}	2^{-14}	0.431×10^{-5}	$-1.42^{733}_{844} \times 10^{-1}$
0.329×10^{-2}	2^{-18}	0.028	60	2^{-27}	0.085	32.63	0.134×10^{-3}	2^{-14}	0.467×10^{-5}	$-1.1^{588}_{600} \times 10^{-1}$
0.302×10^{-2}	2^{-18}	0.034	60	2^{-27}	0.097	33.55	0.152×10^{-3}	2^{-14}	0.510×10^{-5}	$-9.^{399}_{411} \times 10^{-2}$
0.277×10^{-2}	2^{-18}	0.044	58	2^{-27}	0.11	34.26	0.173×10^{-3}	2^{-14}	0.560×10^{-5}	$-7.^{692}_{706} \times 10^{-2}$
0.252×10^{-2}	2^{-18}	0.053	56	2^{-27}	0.13	35.04	0.198×10^{-3}	2^{-14}	0.626×10^{-5}	$-6.3^{035}_{187} \times 10^{-2}$
0.232×10^{-2}	2^{-18}	0.053	56	2^{-27}	0.14	35.94	0.223×10^{-3}	2^{-14}	0.693×10^{-5}	$-5.4^{305}_{472} \times 10^{-2}$
0.212×10^{-2}	2^{-18}	0.055	56	2^{-27}	0.15	36.88	0.254×10^{-3}	2^{-14}	0.784×10^{-5}	$-4.7^{769}_{957} \times 10^{-2}$
0.192×10^{-2}	2^{-18}	0.051	58	2^{-27}	0.16	38.21	0.293×10^{-3}	2^{-14}	0.909×10^{-5}	$-4.3^{776}_{993} \times 10^{-2}$
0.177×10^{-2}	2^{-18}	0.053	58	2^{-27}	0.18	38.98	0.329×10^{-3}	2^{-14}	0.104×10^{-4}	$-4.2^{654}_{897} \times 10^{-2}$
0.162×10^{-2}	2^{-18}	0.054	58	2^{-27}	0.19	39.82	0.373×10^{-3}	2^{-14}	0.121×10^{-4}	$-4.3^{194}_{479} \times 10^{-2}$
0.150×10^{-2}	2^{-18}	0.055	58	2^{-27}	0.2	40.53	0.419×10^{-3}	2^{-14}	0.139×10^{-4}	$-4.^{488}_{521} \times 10^{-2}$
0.137×10^{-2}	2^{-18}	0.063	57	2^{-27}	0.23	41.00	0.476×10^{-3}	2^{-14}	0.164×10^{-4}	$-4.7^{613}_{992} \times 10^{-2}$
0.125×10^{-2}	2^{-18}	0.066	58	2^{-27}	0.25	42.14	0.554×10^{-3}	2^{-14}	0.200×10^{-4}	$-5.1^{293}_{750} \times 10^{-2}$
0.115×10^{-2}	2^{-18}	0.071	58	2^{-27}	0.27	43.12	0.636×10^{-3}	2^{-14}	0.239×10^{-4}	$-5.^{491}_{545} \times 10^{-2}$
0.105×10^{-2}	2^{-18}	0.079	58	2^{-27}	0.3	44.28	0.747×10^{-3}	2^{-14}	0.294×10^{-4}	$-5.9^{172}_{835} \times 10^{-2}$
0.960×10^{-3}	2^{-18}	0.086	58	2^{-27}	0.34	45.41	0.876×10^{-3}	2^{-14}	0.360×10^{-4}	a$-6.^{346}_{427} \times 10^{-2}$

Table 1. Continued

ξ, noisesize	δ_{contr}	α_{contr}	n_{contr}	δ	α	$\sum_{i=0}^{n_{\text{contr}}} C_i$	*A priori* L^1 err. on measure	δ_{est}	Refined L^1 err. on measure	rig. estimate on Lyapunov exponent
0.885×10^{-3}	2^{-18}	0.092	58	2^{-27}	0.37	46.43	0.102×10^{-2}	2^{-14}	0.436×10^{-4}	$-6.^{758}_{855} \times 10^{-2}$
0.810×10^{-3}	2^{-18}	0.095	58	2^{-27}	0.4	47.44	0.119×10^{-2}	2^{-14}	0.534×10^{-4}	$-7.^{200}_{320} \times 10^{-2}$
0.748×10^{-3}	2^{-18}	0.098	58	2^{-27}	0.43	48.25	0.138×10^{-2}	2^{-14}	0.643×10^{-4}	$-7.^{569}_{711} \times 10^{-2}$
0.686×10^{-3}	2^{-18}	0.1	57	2^{-27}	0.46	48.65	0.162×10^{-2}	2^{-14}	0.779×10^{-4}	$-^{7.888}_{8.061} \times 10^{-2}$
0.623×10^{-3}	2^{-18}	0.099	56	2^{-27}	0.5	49.04	0.192×10^{-2}	2^{-14}	0.956×10^{-4}	$-8.^{068}_{280} \times 10^{-2}$
0.573×10^{-3}	2^{-19}	0.07	60	2^{-27}	0.29	46.78	0.141×10^{-2}	2^{-15}	0.595×10^{-4}	$-8.^{076}_{209} \times 10^{-2}$
0.524×10^{-3}	2^{-19}	0.085	56	2^{-27}	0.33	46.32	0.162×10^{-2}	2^{-15}	0.696×10^{-4}	$-7.^{784}_{940} \times 10^{-2}$
0.480×10^{-3}	2^{-19}	0.07	58	2^{-27}	0.34	47.75	0.185×10^{-2}	2^{-15}	0.800×10^{-4}	$-7.^{312}_{490} \times 10^{-2}$
0.436×10^{-3}	2^{-19}	0.066	58	2^{-27}	0.36	48.66	0.215×10^{-2}	2^{-15}	0.935×10^{-4}	$-6.^{653}_{861} \times 10^{-2}$
0.399×10^{-3}	2^{-19}	0.06	60	2^{-27}	0.39	50.39	0.254×10^{-2}	2^{-15}	0.110×10^{-3}	$-^{5.964}_{6.208} \times 10^{-2}$
0.368×10^{-3}	2^{-19}	0.065	60	2^{-27}	0.43	51.36	0.299×10^{-2}	2^{-15}	0.128×10^{-3}	$-5.^{339}_{623} \times 10^{-2}$
0.343×10^{-3}	2^{-20}	0.041	65	2^{-27}	0.24	49.27	0.232×10^{-2}	2^{-16}	0.976×10^{-4}	$-^{4.871}_{5.086} \times 10^{-2}$
0.312×10^{-3}	2^{-20}	0.044	65	2^{-27}	0.26	50.33	0.270×10^{-2}	2^{-16}	0.112×10^{-3}	$-4.^{189}_{436} \times 10^{-2}$
0.287×10^{-3}	2^{-20}	0.048	65	2^{-27}	0.29	51.29	0.309×10^{-2}	2^{-16}	0.127×10^{-3}	$-3.^{565}_{844} \times 10^{-2}$
0.259×10^{-3}	2^{-20}	0.042	67	2^{-27}	0.31	53.14	0.368×10^{-2}	2^{-16}	0.149×10^{-3}	$-2.^{659}_{984} \times 10^{-2}$
0.237×10^{-3}	2^{-20}	0.046	67	2^{-27}	0.35	54.32	0.431×10^{-2}	2^{-17}	0.144×10^{-3}	$-1.^{686}_{998} \times 10^{-2}$
0.218×10^{-3}	2^{-20}	0.05	67	2^{-27}	0.38	55.49	0.504×10^{-2}	2^{-17}	0.165×10^{-3}	$-^{5.731}_{9.274} \times 10^{-3}$
0.199×10^{-3}	2^{-20}	0.051	68	2^{-27}	0.42	57.28	0.606×10^{-2}	2^{-17}	0.193×10^{-3}	$2.9887.138 \times 10^{-3}$
0.184×10^{-3}	2^{-20}	0.051	69	2^{-27}	0.46	59.07	0.725×10^{-2}	2^{-17}	0.225×10^{-3}	$1.^{844}_{362} \times 10^{-2}$
0.168×10^{-3}	2^{-20}	0.053	70	2^{-27}	0.5	61.10	0.896×10^{-2}	2^{-17}	0.272×10^{-3}	$2.^{977}_{394} \times 10^{-2}$
0.154×10^{-3}	2^{-21}	0.042	74	2^{-27}	0.29	57.78	0.650×10^{-2}	2^{-18}	0.708×10^{-4}	$3.^{689}_{533} \times 10^{-2}$
0.142×10^{-3}	2^{-21}	0.048	74	2^{-27}	0.32	59.11	0.757×10^{-2}	2^{-18}	0.822×10^{-4}	$4.^{462}_{282} \times 10^{-2}$
0.129×10^{-3}	2^{-21}	0.049	75	2^{-27}	0.36	60.98	0.901×10^{-2}	2^{-18}	0.988×10^{-4}	$5.^{239}_{023} \times 10^{-2}$
0.106×10^{-3}	2^{-21}	0.058	75	2^{-27}	0.45	64.34	0.135×10^{-1}	2^{-18}	0.154×10^{-3}	$6.^{965}_{626} \times 10^{-2}$
0.873×10^{-4}	2^{-21}	0.062	75	2^{-27}	0.55	67.55	0.209×10^{-1}	2^{-18}	0.252×10^{-3}	$8.^{917}_{365} \times 10^{-2}$

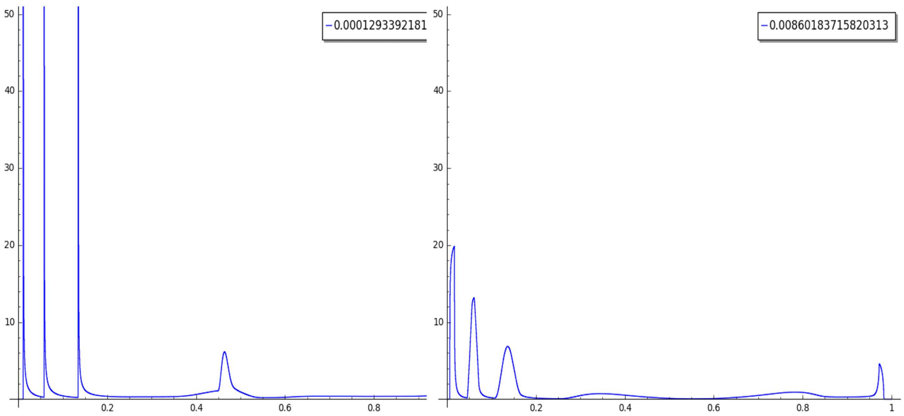

Figure 3. Plot of the computed invariant densities for the noise amplitudes in theorem 2. On the left, a plot of the approximated invariant density for $\xi_1 = 0.129 \times 10^{-3}$ (up to an error of 0.565×10^{-3} in the L^1 norm). On the right, a plot of of the approximated invariant density for $\xi_2 = 0.860 \times 10^{-2}$ (up to an error of 0.973×10^{-5} in the L^1 norm).

information coming from a coarse knowledge of the stationary measure in a bootstrap argument. This procedure uses several technical lemmas estimating the variation of certain densities and the norms of certain operators; such lemmas are listed in section 3 and proved in Appendix B to make reading easier. In section 4 we show an efficient procedure (which is used in the main algorithm) for the estimation of the rate of contraction of a finite rank transfer operator, when applied to zero average measures. This is a quantitative measure of the rate of convergence to equilibrium of the system which is involved in the quantitative estimate of its stability under perturbation. This is important in the estimation of the approximation error. In section 5 we show the estimates needed to compute the Lyapunov exponent of the system once we know the stationary measure. In section 6 we apply all these techniques to the system described in section 2, showing the results of our computer aided estimates, and proving the existence of noise induced order (and in particular, theorem 2). In section 7 we consider the stability of the stationary measure to changes in the system's parameters and of the Lyapunov exponent on changes of the noise amplitude proving the Hőlder continuity. Appendix A contains some definitions and generalities about random dynamical systems we include for completeness and to justify the correctness of the notion of Lyapunov exponent which is estimated in the paper.

2. The system and its transfer operator

In this section we describe more precisely the system which will be studied in the paper and the associated transfer operators. Basic notions on random dynamical systems, stationary measures and Lyapunov exponents we use in the paper are presented in Appendix A.

A random dynamical system with additive noise on [0, 1] and reflecting boundary conditions is a random perturbation of a deterministic map, defined by

$$x \to T(x)\hat{+}\omega_n \tag{5}$$

where $T : [0, 1] \to [0, 1]$ is a Borel measurable map and ω_n is an i.i.d. process distributed according to a probability density ρ_ξ and $\hat{+}$ is the 'reflecting boundaries sum' on [0, 1] defined as follows.

Definition 4. Let $\pi : \mathbb{R} \to [0, 1]$ be the piecewise linear map

$$\pi(x) = \min_{i\in\mathbb{Z}} |x - 2i|. \tag{6}$$

Let $a, b \in \mathbb{R}$ then

$$a\hat{+}b := \pi(a + b)$$

where $+$ is the usual sum operator on $\mathbb{R}$. By this $a\hat{+}b \in [0, 1]$.

As described in the introduction, the model studied in the present paper is the one studied in [24]. We consider a random dynamical system with additive noise, as in (5) where $T = T_{a,b,c}$ is defined in (1). At each iteration of the map a uniformly distributed noise perturbation with span of size ξ is added and the reflecting boundary condition applied.

Remark 5. The parameters a, b, c defined below (1) have been computed using interval arithmetic, in the implementation of our algorithm they are represented by intervals. We explain the motivation for the choice of the parameters in [24]. The parameter c is defined in a way to be nearby to a value for which $T(0.3^-) = T(0.3^+)$, making T continuous at 0.3. The exact value of c giving the continuity can be computed in a closed form as:

$$c = \frac{20}{3^{20} \cdot 7} \cdot \left(\frac{7}{5}\right)^{1/3} \cdot \mathrm{e}^{187/10}.$$

The parameter a is defined similarly in a way so that $T'(0.3^-) = 0$, making T' continuous at 0.3. The value of such a can be computed in a closed form as:

$$a = \frac{19}{42} \cdot \left(\frac{7}{5}\right)^{1/3}.$$

The choice of the parameter b in [24] is motivated by a parallel with the logistic map. Let us denote by T_b the map as only the parameter b varies; each T_b has a repelling fixed point p_b. In [24] this explicit value of b is chosen as an approximation of the parameter value for which $T^4(0.3) = p_b$ following the kneading sequence 'RLLL', i.e., a Misiurewicz condition. We computed a certified interval containing b using a Newton interval method [31]. The interval enclosing b is computed giving the result shown at (3).

We remark that all the results presented in this paper, including the numerical ones hold for each system whose coefficients are included the intervals considered at (2)–(4) hence including cases in which $T_{a,b,c}$ is discontinuous or not satisfying the Misiurewicz condition.

While in our computer assisted estimates we will consider the map given at (1) and uniform noise, the mathematical treatment about approximation of stationary measures in section 3 and following is more general.

We start considering a general random dynamical system where T is measurable and $\rho \in \mathrm{BV}$ giving general results and estimates we then improve using more assumptions on the system (i.e. T piecewise smooth) putting ourselves in different general settings which are stated precisely (see settings 14 and 19), to keep the exposition as clear and general as possible.

More precisely, will consider the case where the noise density ρ_ξ is the rescaling of some bounded variation kernel $\rho \in \mathrm{BV}\left[-\frac{1}{2}, \frac{1}{2}\right]$ with $\int \rho = 1$ in the interval $\left[-\frac{\xi}{2}, \frac{\xi}{2}\right]$, hence

$$\rho_\xi(x) = \frac{1}{\xi}\rho\left(\frac{1}{\xi}x\right).$$

(see definition 20 for a recall on the definition of bounded variation.) The case of uniformly distributed noise correspond to the case $\rho = 1$.

Remark 6. The reflecting boundary condition at (5) is not influent when the noise amplitude is smaller than the parameter b (as it will be for all the noise amplitudes considered in our computer aided proofs, see table 1).

The transfer operator. We study the statistical properties of the dynamical systems with additive noise, as defined at (5) through the study of the properties of their associated transfer operators. Let us recall that a measurable map $T : X \to X$, induces a map

$$L : \mathrm{SM}(X) \to \mathrm{SM}(X)$$

where SM(X) is the space of Borel signed measures on X. The associated map L is defined in the following way: if $\mu \in \mathrm{SM}(X)$ then:

$$L\mu(A) = \mu(T^{-1}(A)).$$

In the literature, L is also called the pushforward map associated to T, sometime denoted by T^*. It is a linear operator on the vector space SM(X) and it is also called the transfer operator associated to T. The space of measures with density in $L^1([0, 1])$ can be seen as a subspace of SM([0, 1]). If T is nonsingular, L can be considered as an operator $L^1([0, 1]) \to L^1([0, 1])$.

The (annealed) transfer operator L_ξ associated to the system with noise is given by the composition of the transfer operator L and a reflecting boundary convolution operator $N_\xi : \mathrm{SM}([0, 1]) \to L^1([0, 1])$ (a suitable modification of the usual convolution), defined by

$$N_\xi(f) := \rho_\xi \hat{*} f \tag{7}$$

where the 'reflecting boundaries convolution' $\hat{*}$ is defined similarly to the reflecting boundaries sum as

Definition 7. Let $\mu \in \mathrm{SM}(\mathbb{R})$. Let $\pi : \mathbb{R} \to [0, 1]$ be the piecewise linear map defined at (6) and π^* its associated pushforward map. We consider $\pi^*\mu \in \mathrm{SM}([0, 1])$ as the 'reflecting boundary' version of μ.

Definition 8. Let $f \in \mathrm{SM}([0, 1])$, $\rho_\xi \in \mathrm{BV}[-\xi, \xi]$. Let $\hat{f} \in \mathrm{SM}(\mathbb{R})$ defined by $\hat{f} = 1_{[0,1]}f$ and $\hat{\rho}_\xi \in L^1(\mathbb{R})$ by $\hat{\rho}_\xi = 1_{[-\xi,\xi]}\rho_\xi$. We define

$$\rho_\xi \hat{*} f = \pi^*(\hat{\rho}_\xi * \hat{f}) \tag{8}$$

where * stands for the usual convolution operator on $\mathbb{R}$.

This boundary reflecting convolution operator is regularizing, in particular $\rho_\xi \hat{*} f \in \mathrm{BV}([0, 1])$ if $f \in \mathrm{BV}$, and has properties similar to the usual convolution operator. For its basic properties see subsection 3.3.1.

Definition 9. The **annealed transfer operator** $L_\xi : \mathrm{SM}([0, 1]) \to L^1([0, 1])$ associated to a deterministic system with additive noise, as described at (5) is defined as

$$L_\xi := N_\xi L. \tag{9}$$

Remark 10. The annealed transfer operator is obtained by averaging the transfer operator L over all the possible noise perturbations. We refer to Appendix A for some basic facts on this operator. In the following we will mainly consider L_ξ as an operator $L^1([0, 1]) \to L^1([0, 1])$. In the notation we emphasize the dependence of the operator on the amplitude of the noise.

Definition 11. Let $\mu_\xi \in L^1([0,1])$ be a fixed probability measure for L_ξ, i.e.,

$$L_\xi \mu_\xi = \mu_\xi$$

we will call μ_ξ a **stationary measure for the system with additive noise** or an **invariant measure for L_ξ**.

By the regularizing properties of the convolution by a bounded variation kernel, and standard compactness arguments it is easy to see that the transfer operator L_ξ corresponding to a map with additive noise has at least one fixed point f_ξ in BV[0, 1] (see [12], lemma 23 for more details). Following [24] we will study the Lyapunov exponent of the system as ξ varies. The Lyapunov exponent is defined as follows:

Definition 12. The **average Lyapunov exponent** associated to noise size ξ is

$$\lambda_\xi := \int_0^1 \log|T'(x)| \mathrm{d}\mu_\xi. \tag{10}$$

When μ_ξ is ergodic the average Lyapunov exponent coincides almost everywhere with the pointwise Lyapunov exponent (see Appendix A). As a byproduct of our computer aided estimates, using the results given in sections 4 and 7 we will prove that the systems considered are ergodic (see proposition 46). This justifies the correctedness of the average Lyapunov exponent as an indicator of the behaviour of the system.

The Ulam Approximation. The main tool for the study of the behaviour of (10) in this work is the rigorous approximation of the stationary measure μ_ξ. This is done by approximating L_ξ by a finite dimensional operator $L_{\delta,\xi} : L^1([0,1]) \to L^1([0,1])$. The fixed points of L_ξ are then approximated by the ones of $L_{\delta,\xi}$ with a certified bound on the approximation error.

Let $\pi_\delta : L^1([0,1]) \to L^1([0,1])$ be a projection on a finite dimensional space defined in the following way: the space [0, 1] is discretized by a partition I_δ (with k elements); the projection considered is defined by the conditional expectation

$$\pi_\delta(f) = \mathbf{E}(f|F_\delta) \tag{11}$$

where F_δ is the $\sigma-$algebra associated to the partition I_δ.

The approximated operator is then defined by finite element approach, composing with π:

$$L_{\delta,\xi} := \pi_\delta N_\xi \pi_\delta L \pi_\delta.$$

The finite dimensional approximation of an operator based on the conditional expectation, as above, is commonly called *Ulam discretization* or *Ulam method*. This method was widely studied in the literature (see, e.g [5, 6, 9, 13, 22]).

Observe that

$$L^n_{\delta,\xi} = (\pi_\delta N_\xi \pi_\delta L)^n \pi_\delta,$$

taking into account that $\pi_\delta^2 = \pi_\delta$. We remark that since $\|\pi_\delta\|_{L^1\to L^1} \leqslant 1$ and $\|L\|_{L^1\to L^1} \leqslant 1$, then $\|L_{\delta,\xi}\|_{L^1\to L^1} \leqslant 1$.

Remark 13. Another discretization that could be used is

$$\tilde{L}_{\delta,\xi} = \pi_\delta N_\xi L \pi_\delta;$$

while this definition is reasonable, it is more difficult to implement and would force us to recompute the discretized operator for each size ξ of the noise. Our definition permits us to compute once and for all $\pi_\delta L\pi_\delta$ which is computationally expensive but leads to a sparse matrix, and then apply the operator $\pi_\delta N_\xi \pi_\delta$ which is independent of the dynamics.

3. Rigorous approximation of the stationary measure for dynamical systems with additive noise

The certified approximation of the Lyapunov exponent is based on the certified approximation of the stationary measure of the system in the L^1 norm. This is the main part of our general construction and is described in this section. The algorithm to approximate the stationary measure uses both *a priori* (analytical) and *a posteriori* (computer assisted) estimates on the measure and on the transfer operator. *For these estimates to be performed we do not need particular expansion or hyperbolicity properties of the one dimensional map driving the deterministic part of the dynamics.* We now introduce a general and simple algorithm which works for measurable maps perturbed by additive noise with a Bounded Variation kernel, in section 3.3 we refine this algorithm, assuming that the map is piecewise smooth and getting much better estimates. In this first part of the section we will hence work in the following setting.

Setting 14. *Let us suppose L_ξ is the transfer operator of a system with additive noise as considered in (5). We suppose the noise is distributed according to a bounded variation kernel ρ_ξ with support in $\left[-\frac{\xi}{2}, \frac{\xi}{2}\right]$ and the deterministic part of the system is driven by a measurable map T.*

Let $L_{\delta,\xi}$ be the Ulam approximation of L_ξ, defined by projecting on a partition of size δ. Let $f_{\delta,\xi}, f_\xi \in L^1$ respectively be invariant probability measures of $L_{\delta,\xi}$ and L_ξ. Since the measure $f_{\delta,\xi}$ is a fixed point of the finite dimensional operator $L_{\delta,\xi}$, it can be computed to any precision. We will treat now the issue of estimating $\|f_{\delta,\xi} - f_\xi\|_{L^1}$ effectively.

Lemma 15. *Suppose that for some $\overline{n} \in \mathbb{N}$*

$$\|L^{\overline{n}}_{\delta,\xi}|_V\|_{L^1\to L^1} \leqslant \alpha < 1 \tag{12}$$

where $V = \{\nu \in L^1, \nu([0,1]) = 0\}$ being the space of zero-average measures. Then

$$\|f_\xi - f_{\xi,\delta}\|_{L^1} \leqslant \frac{1}{1-\alpha}\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_\xi)f_\xi\|_{L^1}. \tag{13}$$

Remark 16. We remark that $L_{\delta,\xi}|_V$ is a finite dimensional operator and can be represented by a matrix. Thus is possible for a computer to verify that $\|L^{\overline{n}}_{\delta,\xi}|_V\|_{L^1\to L^1} \leqslant \alpha$ for some $\overline{n}$.

Proof (Of lemma 15). Since both $f_\xi, f_{\delta,\xi}$ are fixed points we can write

$$\begin{aligned}\|f_{\delta,\xi} - f_\xi\|_{L^1} &= \|L^n_{\delta,\xi} f_{\delta,\xi} - L^n_\xi f_\xi\|_{L^1} \\ &= \|L^n_{\delta,\xi} f_{\delta,\xi} - L^n_{\delta,\xi} f_\xi + L^n_{\delta,\xi} f_\xi - L^n_\xi f_\xi\|_{L^1} \\ &\leqslant \|L^n_{\delta,\xi}(f_{\delta,\xi} - f_\xi)\|_{L^1} + \|(L^n_{\delta,\xi} - L^n_\xi) f_\xi\|_{L^1}\end{aligned}$$

Then

$$\|f_\xi - f_{\xi,\delta}\|_{L^1} \leqslant \alpha\|f_\xi - f_{\xi,\delta}\|_{L^1} + \|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_\xi) f_\xi\|_{L^1}$$

implying the statement. □

3.1. An informal description of the main algorithm to compute the stationary measure up to a small given error

Based on lemma 15, a strategy to rigorously bound $\|f_{\delta,\xi} - f_\xi\|_{L^1}$ is the following. The computer will find an $\overline{n}$ such that (12) is satisfied, then (13) will give an estimate for the approximation error. We remark that if δ is small enough, $\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_{\xi})f_\xi\|_{L^1}$ has a chance of being small, since it is the difference of two nearby operators, both applied to the same regular (bounded variation) measure. This is where the size δ of the approximation grid has a role in the quality of the approximation. On the other hand, $\overline{n}$ may depend on δ. This is why *a priori* estimates on the approximation error are not trivial. Lemma 15 provides some *a posteriori* estimate on the error; the approximation error is known after the computer certifies the $\overline{n}$ and the α for which (12) is satisfied.

Hence the main algorithm for the approximation of the invariant measure will work as follows:

(a) Given the grid size δ, compute $L_{\delta,\xi}$ and $f_{\xi,\delta}$ up to some prescribed precision.

(b) Find a good $\overline{n}$ and α: we estimate $\|L^n_{\delta,\xi}|_V\|_{L^1\to L^1}$ in an efficient way, finding a good compromise between $\overline{n}$ and α, in a way that α is not too close to 1 and $\overline{n}$ not too big. We remark that the norm of the finite dimensional operator $L^n_{\delta,\xi}$ is directly computable in principle, but if δ is small then the size of the associated matrix is huge and this can be a hard computational task. For this we use a *coarse-fine* strategy which is explained in section 4, and which takes into account that the huge matrix representing $L_{\delta,\xi}$ is actually coming from a certain dynamical system with noise.

(c) Find a good estimate for $\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_{\xi})f_\xi\|_{L^1}$. We remark that in this formula f_ξ is not known, but still we can find enough information on it to estimate the difference of operators we are interested in. This will be done by a method using both *a priori* and *a posteriori* estimates, using an approximated knowledge of f_ξ and its variation. The procedure is explained in section 3.2 and in the following sections. A simple but not efficient bound (equation (16)) is proved in section 3.2.1; we refine the method in section 3.3 greatly improving the efficiency of the estimate with a bootstrap argument.

(d) Estimate the approximation error $\|f_{\xi,\delta} - f_\xi\|_{L^1}$ by lemma 15.

3.2. A bound for $\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_{\xi})f_\xi\|_{L^1}$

Having outlined the main algorithm we now show how to perform the main needed estimates. In this section we describe how to estimate the quantity appearing on the right-hand side of (13). This will be done by splitting this term in different parts which will be treated differently. We start estimating the term as a telescopic sum whose summands will be estimated in the following subsections.

Lemma 17. *Let L_ξ the transfer operator of the random system and $L_{\delta,\xi}$ its Ulam approximation, as defined in section 2. Let f_ξ be an invariant probability measure for L_ξ, it holds*

$$\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_{\xi})f_\xi\|_{L^1} \leqslant \|(\pi_\delta - 1)f_\xi\|_{L^1} + \left(\sum_{i=0}^{\overline{n}-1} \|L_{\delta,\xi}|_V^i\|_{L^1\to L^1}\right) \times \left(\|N_\xi(\pi_\delta - 1)Lf_\xi\|_{L^1} + \|N_\xi \pi_\delta L(\pi_\delta - 1)f_\xi\|_{L^1}\right) \qquad (14)$$

Proof. The proof is based on a telescopic decomposition. We have

$$\begin{aligned}L_{\delta,\xi} - L_{\xi} &= \pi_\delta N_\xi \pi_\delta L \pi_\delta - N_\xi L \\ &= N_\xi L - \pi_\delta N_\xi L + \pi_\delta N_\xi L - \pi_\delta N_\xi \pi_\delta L + \pi_\delta N_\xi \pi_\delta L - \pi_\delta N_\xi \pi_\delta L \pi_\delta.\end{aligned}$$

Performing a similar decomposition to $L^{\overline{n}}_{\delta,\xi} = (\pi_\delta N_\xi \pi_\delta L)^{\overline{n}} \pi_\delta$. Pairing in a suitable way the terms we inserted we obtain:

$$\begin{aligned}\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_{\xi}) f_\xi\|_{L^1} &= \left\| \left[(\pi_\delta N_\xi \pi_\delta L)^{\overline{n}} \pi_\delta - (N_\xi L)^{\overline{n}} \right] f_\xi \right\|_{L^1} \\ &\leqslant \sum_{i=0}^{\overline{n}} \|(\pi_\delta N_\xi \pi_\delta L)^i (\pi_\delta - 1)(N_\xi L)^{n-i} f_\xi\|_{L^1} \\ &\quad + \sum_{i=0}^{\overline{n}-1} \|(\pi_\delta N_\xi \pi_\delta L)^i \pi_\delta N_\xi (\pi_\delta - 1) L (N_\xi L)^{n-i-1} f_\xi\|_{L^1} \\ &= \sum_{i=0}^{\overline{n}} \|(\pi_\delta N_\xi \pi_\delta L)^i (\pi_\delta - 1) f_\xi\|_{L^1} \\ &\quad + \sum_{i=0}^{\overline{n}-1} \|(\pi_\delta N_\xi \pi_\delta L)^i \pi_\delta N_\xi (\pi_\delta - 1) L f_\xi\|_{L^1}\end{aligned}$$

considering that f_ξ is fixed by $N_\xi L = L_\xi$. Shifting indexes by 1 in the the first sum, the estimate can be written as

$$\begin{aligned}\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_{\xi}) f_\xi\|_{L^1} &\leqslant \|(\pi_\delta - 1) f_\xi\|_{L^1} + \sum_{i=0}^{\overline{n}-1} \|(\pi_\delta N_\xi \pi_\delta L)^i \pi_\delta|_V\|_{L^1\to L^1} \cdot \|N_\xi \pi_\delta L (\pi_\delta - 1) f_\xi\|_{L^1} \\ &\quad + \sum_{i=0}^{\overline{n}-1} \|(\pi_\delta N_\xi \pi_\delta L)^i \pi_\delta|_V\|_{L^1\to L^1} \cdot \|N_\xi (\pi_\delta - 1) L f_\xi\|_{L^1} \\ &= \|(\pi_\delta - 1) f_\xi\|_{L^1} + \left(\sum_{i=0}^{\overline{n}-1} \|L_{\delta,\xi}|^i_V\|_{L^1\to L^1} \right) \\ &\quad \times \left(\|N_\xi (\pi_\delta - 1) L f_\xi\|_{L^1} + \|N_\xi \pi_\delta L (\pi_\delta - 1) f_\xi\|_{L^1} \right)\end{aligned}$$

(notice that $(\pi_\delta - 1)g$ has always average zero for any g, and consequently belongs to V). □

3.2.1. An initial (a priori) bound for $\|f_\xi - f_{\xi,\delta}\|_{L^1}$. Now we show a strategy to get a simple effective bound for the approximation error $\|f_\xi - f_{\xi,\delta}\|_{L^1}$ based lemma 17, estimating the summands on the right-hand side of (14) by quantities which are known from the description of the system or can be computed by its approximated transfer operator. In the next section we will improve the method, using more information on T and f_ξ and getting much more efficient estimates.

Lemma 18. *Let f_ξ a stationary measure for a system defined as in (5) and $f_{\xi,\delta}$ a stationary measure for its Ulam approximation, as defined in section 2. If there is $\overline{n}$ such that*

$$\|L^{\overline{n}}_{\delta,\xi}|_V\|_{L^1\to L^1} \leqslant \alpha < 1 \tag{15}$$

then

$$\|f_\xi - f_{\xi,\delta}\|_{L^1} \leqslant \frac{1 + 2\sum_{i=0}^{\overline{n}-1} C_i}{2(1-\alpha)}\, \delta\xi^{-1}\ \mathrm{Var}(\rho). \tag{16}$$

where $0 \leqslant C_i \leqslant 1$ *are such that* $\|L^i_{\delta,\xi}|_V\|_{L^1\to L^1} \leqslant C_i$.

Proof. The proof of the lemma is based on the following estimates, proved in corollary 49 and proposition 50 (Appendix B) allowing a first estimate on the right-hand side of (13). We have

$$\|N_\xi(1-\pi_\delta)\|_{L^1\to L^1} \leqslant \frac{1}{2}\delta\xi^{-1}\,\mathrm{Var}(\rho). \tag{17}$$

$$\|(1-\pi_\delta)N_\xi\|_{L^1\to L^1} \leqslant \frac{1}{2}\delta\xi^{-1}\,\mathrm{Var}(\rho). \tag{18}$$

Now we apply (17), (18) to the summands of the right-hand side of (14). We see that all the items there have either a $N_\xi(1-\pi_\delta)$ or a $(1-\pi_\delta)N_\xi$ appearing. Indeed since $\|f_\xi\|_1 \leqslant 1$

$$\|(\pi_\delta - 1)f_\xi\|_{L^1} = \|(\pi_\delta - 1)N_\xi L f_\xi\|_{L^1} \leqslant \frac{1}{2}\delta\xi^{-1}\,\mathrm{Var}(\rho).$$

Similarly the other summands, can be estimated recalling that $\|\pi_\delta\|_{L^1\to L^1} \leqslant 1$ and $\|N_\xi\|_{L^1\to L^1} \leqslant 1$. Applying (13) we get the statement. □

The estimate given at (16) mainly depend on the ratio $\delta\xi^{-1}$ between the partition size and the noise amplitude. This estimate is obtained without any information on the deterministic part of the dynamics, only the information about the contraction rate of the approximated transfer operator $L_{\delta,\xi}$ (to obtain $\overline{n}$ and α). This would already allow to obtain a good approximation for the invariant density f_ξ, in principle, if we had enough computation power to carry on the computation with a very small δ. Unfortunately, in the Matsumoto–Tsuda system, positive Lyapunov exponent appears for very small sizes of the noise making the computation unfeasible even with the help of a supercomputer, due to the growth of the computational complexity as ξ becomes small.

Therefore, we have to apply a more subtle and complicated strategy where the bootstrap argument comes into play, i.e., using some information on f_ξ (and in particular about its variation in given intervals) which we can obtain with a preliminary computation.

3.3. A stronger (a posteriori) bound

In this section we analyse better (14) and see how using some more assumptions on T, more information on f_ξ and the use of the Wasserstein distance, we can drastically improve the estimates given in lemma 18. We remark that the explicit error bound provided by lemma 18 is proportional to $\delta\xi^{-1}$; the new error estimate will be a sum where most summands are proportional to $\delta^2\xi^{-1}$.

Setting 19. *From now on we will suppose we are in the framework of setting 14 and furthermore we suppose T being piecewise smooth. We suppose that there is a partition* $\{P_i\}_{1\leqslant i\leqslant k}$ *such that*

- *each* P_i *is an interval,*
- *on each* P_i *the branch* $T_i := T|_{P_i}$ *is monotonic and* C^2 *in the interior of* P_i
- *The limits of* $T_i'(x)$ *as x tend to the frontier of* P_i *exist in* $\mathbb{R}\cup\{-\infty,\infty\}$.

We will consider the transfer operator L related to the map and to every branch of the map. For each L^1 density g, we let $L_i g$ be the component of Lg coming from the ith monotone branch (the pushforward map related to T_i), that is

$$L_i g(x) = \begin{cases} \dfrac{g(T_i^{-1}(x))}{T'(T_i^{-1}(x))} & \text{if } x \in T_i(I_i), \\ 0 & \text{elsewhere.} \end{cases} \tag{19}$$

In this way we have $Lg = \sum_i L_i g$.

Once the discretized transfer operator $L_{\delta,\xi}$ is computed and has a unique fixed probability measure $f_{\delta,\xi}$, an approximation $\tilde{f}$ for $f_{\delta,\xi}$ can be computed up to any given small error in L^1 (this is the computation of the fixed point of the big matrix representing $L_{\delta,\xi}$). Let us assume that a numerical approximation $\tilde{f}$ of $f_{\delta,\xi}$ in the L^1 norm is computed. Using $\tilde{f}$, we will look for a strategy for estimating the total error $\|f_\xi - \tilde{f}\|_{L^1}$, that includes the appoximation error $\|f_\xi - f_{\delta,\xi}\|_{L^1}$ (because we approximated on a partition of size δ) and the numerical error $\|f_{\delta,\xi} - \tilde{f}\|_{L^1}$. We remark that

$$\|f_\xi - \tilde{f}\|_{L^1} \leqslant \|f_{\delta,\xi} - \tilde{f}\|_{L^1} + \|f_\xi - f_{\delta,\xi}\|_{L^1}. \tag{20}$$

For the estimation of $\|f_{\delta,\xi} - \tilde{f}\|_{L^1}$ in our algorithm we apply the same method described in [13]. To find a stronger bound for $\|f_\xi - f_{\xi,\delta}\|_{L^1}$ let us start again from the estimate given at lemma 17. We will estimate independently the terms

$$\|(\pi_\delta - 1)f_\xi\|, \quad \|N_\xi(\pi_\delta - 1)Lf_\xi\|_{L^1}, \quad \|N_\xi \pi_\delta L(\pi_\delta - 1)f_\xi\|_{L^1}, \tag{21}$$

appearing at (14), using the information we can extract from the approximation $\tilde{f}$. Except in the first case (that has the smallest weight in the estimate, according to (14)), the estimate will become roughly proportional to $\delta^2\xi^{-1}$, greatly improving the quality of the approximation certification. For each of the terms in (21) we prove in sections 3.3.2–3.3.4 bounds of the form

$$A\|\tilde{f} - f_\xi\|_{L^1} + B$$

where A and B depend on δ and become small when δ is small. We remark that these bounds **depend** on the error $\|\tilde{f} - f_\xi\|_{L^1}$ itself. This together with precise bounds, based on some approximation $\tilde{f}$, permits us to tighten the bounds on the error $\|\tilde{f} - f_\xi\|_{L^1}$, using a so called ‘bootstrapping’ process.

3.3.1. A summary of norms and estimates. Before entering in the details of the estimate of 21, we introduce some of the norms used in the paper, and we summarize some of the bounds that are used in this section and proved in appendix B.

Definition 20. Let $X \subset [0, 1]$ be a finite union of pairwise disjoint intervals, $X = \bigcup_j I_j$, $I_i \cap I_j = \emptyset$ for $i \neq j$. We define the **variation on** X of the function f (and denote it by $\mathrm{Var}_X(f)$) as follows:

- when X is an interval (the endpoints may be included or not), the variation is defined as

$$\mathrm{Var}_X(f) := \sup_{\{x_0<x_1<\ldots<x_k\}\subset X} \sum_{i=0}^{k-1} |f(x_{i+1}) - f(x_i)|$$

  (the supremum being over finite increasing sequences of any length k contained in X);

- when X is a finite union of pairwise disjoint intervals $X = \bigcup_j I_j$, $I_i \cap I_j = \emptyset$ for $i \neq j$, the variation is defined as $\mathrm{Var}_X(f) := \sum_j \mathrm{Var}_{I_j}(f)$.

Remark 21. As it is well known, if f has a weak derivative, then $\mathrm{Var}_X(f) = \|f'\|_{L^1(X)}$, see for instance [29, chapter 4, proposition 4.2].

We will also consider a norm which is weaker than the L^1 norm.

Definition 22. Let f be a function in $L^1([0,1])$ with zero average, we define the **Wasserstein-like norm** of f, as

$$\|f\|_W := \|F\|_{L^1}, \text{ where } F(x) = \int_0^x f(t)\mathrm{d}t. \tag{22}$$

Let $Y \subset [0,1]$ be a finite union of pairwise disjoint intervals $Y = \bigcup_j I_j$, $I_i \cap I_j = \emptyset$ for $i \neq j$, we denote by $W(Y)$ the space of L^1 functions with support contained in Y having zero average in each I_j. If $f \in W(Y)$ we define its norm by

$$\|f\|_{W(Y)} := \|f\|_W. \tag{23}$$

Next proposition contains a summary of the bounds used in the next proofs and the location of their proof in the paper; it can be used as an handy guideline throughout the paper, we refer to the cited lemmas and proposition for details.

Proposition 23 (*Summary of the bounds*). *Let π_δ be the Ulam projection on a homogeneous partition of size δ, as defined in (11), let I be a set which is a finite union of intervals of the partition then:*

(*a*) $\|1 - \pi_\delta\|_{\mathrm{Var}\to L^1} \leqslant \delta/2$, *lemma 48,*
(*b*) $\|1 - \pi_\delta\|_{L^1\to W} \leqslant \delta/2$, *lemma 52,*
(*c*) $\|1 - \pi_\delta\|_{\mathrm{Var}(I)\to W(I)} \leqslant \delta^2/8$, *lemma 53.*
Let N_ξ be the convolution operator then:

(*d*) $\|N_\xi\|_{L^1\to \mathrm{Var}} \leqslant \xi^{-1}\,\mathrm{Var}(\rho)$, *lemma 47,*
(*e*) $\|N_\xi\|_{W\to L^1} \leqslant \xi^{-1}\,\mathrm{Var}(\rho)$, *lemma 51,*
(*f*) $\|(1 - \pi_\delta)N_\xi\|_{L^1\to L^1} \leqslant \frac{1}{2}\delta\xi^{-1}\,\mathrm{Var}(\rho)$, *corollary 49,*
(*g*) $\|N_\xi(1 - \pi_\delta)\|_{L^1\to L^1} \leqslant \frac{1}{2}\delta\xi^{-1}\,\mathrm{Var}(\rho)$, *proposition 50.*
Let L be the transfer operator associated to T, and let $L_i g$ the component of Lg coming from the ith *branch as defined at (19), then:*

(*h*) $\|L\|_{W(I)\to W} \leqslant \|T'\|_{L^\infty(I)}$, *lemma 55*
(*i*) *in lemma 56 the* ***local variation inequality*** *is proved:*

$$\mathrm{Var}_I(L_i g) \leqslant \mathrm{Var}_{T_i^{-1}(I)}(g) \cdot \|\frac{1}{T'}\|_{L^\infty(T_i^{-1}(I))} + \|g\|_{L^1(T_i^{-1}(I))} \cdot \|\frac{T''}{T'^2}\|_{L^\infty(T_i^{-1}(I))} + \sum_{y\in\partial Dom(T_i):T(y)\in I} \left|\frac{g(y)}{T'(y)}\right|.$$

3.3.2. Estimate for $\|(\pi_\delta - 1)f_\xi\|_{L^1}$. We give now an estimate for the first item of (21).

Lemma 24. *Let π_δ be the Ulam projection on a homogeneous partition of size δ, then*

$$\|(\pi_\delta - 1)f_\xi\|_{L^1} \leqslant A_1 \cdot \|f_\xi - \tilde{f}\|_{L^1} + B_1 \tag{24}$$

for

$$A_1 = \frac{\delta}{2}\xi^{-1}\,\mathrm{Var}(\rho), \quad B_1 = \frac{\delta}{2}\cdot \mathrm{Var}(N_\xi L\tilde{f}). \tag{25}$$

Proof. We estimate:

$$\begin{aligned}\|(\pi_\delta - 1)f_\xi\|_{L^1} &= \|(\pi_\delta - 1)N_\xi L f_\xi\|_{L^1} \leqslant \|(\pi_\delta - 1)N_\xi L(f_\xi - \tilde{f})\|_{L^1} + \|(\pi_\delta - 1)N_\xi L\tilde{f}\|_{L^1}\\ &\leqslant \|(\pi_\delta - 1)N_\xi\|_{L^1} \cdot \|f_\xi - \tilde{f}\|_{L^1} + \|\pi_\delta - 1\|_{\mathrm{Var}\to L^1} \cdot \mathrm{Var}(N_\xi L\tilde{f})\\ &\leqslant \frac{\delta}{2}\xi^{-1}\,\mathrm{Var}(\rho)\cdot\|f_\xi - \tilde{f}\|_{L^1} + \frac{\delta}{2}\cdot\mathrm{Var}(N_\xi L\tilde{f})\end{aligned}$$

where in the last line we used proposition 23, items (6) and (1). □

An upper bound on $\mathrm{Var}(N_\xi L\tilde{f})$ will be estimated using the results in appendix B.2.2 and the explicit knowledge of the computed $\tilde{f}$.

3.3.3. Estimate for $\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1}$. We give now the estimate of the second item of (21). The main idea is to use a coarser partition Π made of intervals whose size is an integral multiple of δ. Then, instead of estimating $\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1}$ globally we bound this quantity on each interval of Π exploiting the approximate knowledge of the variation of f_ξ in the interval. This drastically improves the quality of the estimate in almost every interval of Π and allows the error-checking computation to be performed on a partition which is coarser than the initial partition of size δ.

Lemma 25. *Let Π be a uniform partition whose parts have size that is an integral multiple of δ, T piecewise monotonic with L_i defined as above. We have*

$$\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1} \leqslant A_2\cdot\|f_\xi - \tilde{f}\|_{L^1} + B_2, \tag{26}$$

with

$$\begin{aligned}A_2 &= \frac{\delta}{2}\mathrm{Var}(\rho_\xi)\\ B_2 &= \frac{\delta}{2}\mathrm{Var}(\rho_\xi)\sum_{I\in\Pi}\sum_i \min\left\{\frac{\delta}{4}\cdot\mathrm{Var}_I(L_i\tilde{f}),\ \|L_i\tilde{f}\|_{L^1(I)}\right\}.\end{aligned} \tag{27}$$

Proof. We can estimate as

$$\begin{aligned}\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1} &\leqslant \|N_\xi(1-\pi_\delta)L\tilde{f}\|_{L^1} + \|N_\xi(1-\pi_\delta)L(f_\xi - \tilde{f})\|_{L^1} \leqslant \|N_\xi(1-\pi_\delta)L\tilde{f}\|_{L^1}\\ &\quad + \|N_\xi(1-\pi_\delta)\|_{L^1}\cdot\|f_\xi - \tilde{f}\|_{L^1} \leqslant \|N_\xi(1-\pi_\delta)L\tilde{f}\|_{L^1}\\ &\quad + \frac{\delta}{2}\mathrm{Var}(\rho_\xi)\cdot\|f_\xi - \tilde{f}\|_{L^1}\end{aligned} \tag{28}$$

The first summand can be rewritten as

$$\begin{aligned}\|N_\xi(1-\pi_\delta)L\tilde{f}\|_{L^1} &\leqslant \|N_\xi\|_{W\to L^1}\cdot\|(1-\pi_\delta)L\tilde{f}\|_W\\ &\leqslant \|N_\xi\|_{W\to L^1}\cdot\sum_{I\in\Pi}\|(1-\pi_\delta)L\tilde{f}\cdot\chi_I\|_{W(I)}\end{aligned}$$

splitting on the intervals of the partition, and using lemma 56 (because each $I\in\Pi$ is a union of intervals of the partition of size δ)

$$\leqslant \|N_\xi\|_{W\to L^1} \cdot \sum_{I\in\Pi}\sum_i \min\Big\{\|1-\pi_\delta\|_{\mathrm{Var}_I\to W(I)}\cdot \mathrm{Var}_I(L_i\tilde{f}),$$
$$\|1-\pi_\delta\|_{L^1(I)\to W(I)}\cdot \|L_i\tilde{f}\|_{L^1(I)}\Big\}$$
$$\leqslant \mathrm{Var}(\rho_\xi)\cdot \sum_{I\in\Pi}\sum_i \min\left\{\frac{\delta^2}{8}\cdot \mathrm{Var}_I(L_i\tilde{f}), \frac{\delta}{2}\cdot \|L_i\tilde{f}\|_{L^1(I)}\right\}.$$

In the last step we used the property of the W norm stated in lemmas 52 and 53. Adding the second term of (28) we have proved the result. □

Therefore, once we have an approximation $\tilde{f}$ we can compute A_2, B_2 by a simple algorithm that evaluates the double summation in the last equation of the above lemma.

Remark 26. When computing B_2, the minimum could be always the one depending on $\|L_i\tilde{f}\|_{L^1(I)}$. If this happens, the sum adds up to 1, and the new bound is worse than the *a priori* estimate, since we are introducing a factor $\|f_\xi - \tilde{f}\|_{L^1} + 1 > 1$.

In practice, this does not happen. On each ith preimage of an interval I of the partition the new estimate will provide a better bound as soon as

$$\frac{\delta}{4}\cdot \mathrm{Var}_I(L_i\tilde{f}) < \|L_i\tilde{f}\|_{L^1(I)}.$$

First of all, the variation of $\tilde{f}$ has an *a priori* bound by

$$\mathrm{Var}(\tilde{f}) = \mathrm{Var}(\pi_\delta N_\xi \pi_\delta L \pi_\delta \tilde{f}) \leqslant \|\pi_\delta\|_{\mathrm{Var}}\cdot \|N_\xi\|_{L^1\to \mathrm{Var}}\cdot 1 \leqslant \xi^{-1}\,\mathrm{Var}(\rho).$$

Now, we can try to control $\mathrm{Var}_I(L_i\tilde{f})$ by using the local variation inequality in proposition 23. If the preimage $T_i^{-1}(I)$ does not contain a critical point or a singular point, we expect $(\delta^2/8)\cdot \mathrm{Var}_I(L_i\tilde{f})$ to be small.

For the intervals I where we cannot apply the local variation inequality or where it does not give us good enough bounds we fall back to the a-priori estimate depending on the L^1 mass rather than on the variation.

3.3.4. An estimate for $\|N_\xi\pi_\delta L(\pi_\delta - 1)f_\xi\|_1$. We give an estimate of the third item of (21). The general idea is similar to the one explained in the previous section, again, some required estimates are technical lemmas proved in appendix B.

Lemma 27. *Let Π be a uniform partition whose parts have size that is multiple of δ, we have:*

$$\|N_\xi\pi_\delta L(1-\pi_\delta)f_\xi\|_{L^1} \leqslant A_3\cdot \|f_\xi - \tilde{f}\|_{L^1} + B_3, \tag{29}$$

with

$$A_3 = \frac{\delta}{2}\xi^{-1}\,\mathrm{Var}(\rho),$$
$$B_3 = \sum_{I\in\Pi}\min\left\{\frac{\delta^2}{8}\xi^{-1}\,\mathrm{Var}(\rho)\cdot \|T'\|_{L^\infty(I)}, \frac{\delta}{2}\right\}\mathrm{Var}_I(N_\xi L\tilde{f}) \tag{30}$$
$$+\frac{\delta^2}{4}\xi^{-1}\,\mathrm{Var}(\rho)\cdot \mathrm{Var}(N_\xi L\tilde{f}).$$

Proof. We have

$$\begin{aligned}
\|N_\xi \pi_\delta L(1-\pi_\delta) f_\xi\|_{L^1} &= \|N_\xi \pi_\delta L(1-\pi_\delta) N_\xi L f_\xi\|_{L^1} \\
&\leqslant \|N_\xi \pi_\delta L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1} + \|N_\xi \pi_\delta L(1-\pi_\delta) N_\xi L (f_\xi - \tilde{f})\|_{L^1} \\
&\leqslant \|N_\xi L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1} \\
&\quad + \|N_\xi (1-\pi_\delta) L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1} \\
&\quad + \|N_\xi \pi_\delta L\|_{L^1} \cdot \|(1-\pi_\delta) N_\xi\|_{L^1} \cdot \|f_\xi - \tilde{f}\|_{L^1} \\
&\leqslant \|N_\xi L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1} \\
&\quad + \|N_\xi (1-\pi_\delta)\|_{L^1} \cdot \|1-\pi_\delta\|_{\mathrm{Var}\to L^1} \cdot \mathrm{Var}(N_\xi L \tilde{f}) \\
&\quad + \|(1-\pi_\delta) N_\xi\|_{L^1} \cdot \|f_\xi - \tilde{f}\|_{L^1} \\
&\leqslant \|N_\xi L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1} \qquad (31)\\
&\quad + \frac{\delta^2}{4} \mathrm{Var}(\rho_\xi) \cdot \mathrm{Var}(N_\xi L \tilde{f}) \qquad (32)\\
&\quad + \frac{\delta}{2} \mathrm{Var}(\rho_\xi) \cdot \|f_\xi - \tilde{f}\|_{L^1}. \qquad (33)
\end{aligned}$$

An algorithm for estimating for $\mathrm{Var}(N_\xi L\tilde{f})$ can be found in lemma 57. The term $\|N_\xi L(1-\pi_\delta)N_\xi L\tilde{f}\|_{L^1}$ can be split over the intervals $I \in \Pi$ and estimated as

$$\begin{aligned}
&\sum_{I\in\Pi} \|N_\xi L(1-\pi_\delta)[N_\xi L\tilde{f} \cdot \chi_I]\|_{L^1} \\
&\quad \leqslant \sum_{I\in\Pi} \min\{\|N_\xi\|_{W\to L^1} \cdot \|L\|_{W(I)\to W} \cdot \|1-\pi_\delta\|_{\mathrm{Var}_I \to W(I)}, \\
&\qquad \|1-\pi_\delta\|_{\mathrm{Var}_I\to L^1}\} \cdot \mathrm{Var}_I(N_\xi L\tilde{f}) \\
&\quad \leqslant \sum_{I\in\Pi} \min\left\{\frac{\delta^2}{8}\mathrm{Var}(\rho_\xi) \cdot \|T'\|_{L^\infty(I)}, \frac{\delta}{2}\right\} \mathrm{Var}_I(N_\xi L\tilde{f})
\end{aligned}$$

proving the statement thanks to lemma 55 (because each I is a union of intervals of the partition of size δ). □

Remark 28. To estimate B_3, we estimate computationally $\mathrm{Var}_I(N_\xi L\tilde{f})$ for each interval $I \in \Pi$ using the algorithms explained in appendices B.2.1 and B.2.2. As in lemma 25, we obtain a stronger estimate in the interval I as soon as

$$\|T'\|_{L^\infty(I)} < \frac{4}{\delta\xi^{-1}\,\mathrm{Var}(\rho)}.$$

Remark that in all our computations $\delta\xi^{-1}\mathrm{Var}(\rho)$ needs to be small, since it controls the approximation error (refer to proposition 23, items 6 and 7). This implies that the inequality above will be true for most of the intervals but those where T' becomes very big.

3.3.5. An estimate for the L^1 error $\|f_\xi - \tilde{f}\|_{L^1}$. In the previous sections we built the ingredients for estimating $\|f_\xi - f_{\xi,\delta}\|_{L^1}$, but we want to estimate $\|f_\xi - \tilde{f}\|_{L^1}$ where $\tilde{f}$ is the output of a computation approximating the fixed point of $L_{\delta,\xi}$. We will do so assuming that we have an estimate of the numerical error $\|f_{\xi,\delta} - \tilde{f}\|_{L^1}$ (such an estimate can be found in [13]).

Let A_i, B_i $(i = 1, 2, 3)$ be the constants defined as in (25), (27), (30). Plugging (24), (26), (29) (according to lemmas 25 and 27) into (14), we have that $\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_{\xi})f_\xi\|_{L^1}$ can be bounded as

$$\|(L^{\overline{n}}_{\delta,\xi} - L^{\overline{n}}_{\xi})f_\xi\|_{L^1} \leqslant A \cdot \|f_\xi - \tilde{f}\|_{L^1} + B,$$

where

$$A = A_1 + (A_2 + A_3) \cdot \sum_{i=0}^{\overline{n}-1} C_i \quad B = B_1 + (B_2 + B_3) \cdot \sum_{i=0}^{\overline{n}-1} C_i.$$

Thanks to (13) we have

$$\|f_\xi - f_{\xi,\delta}\|_{L^1} \leqslant C + D \cdot \|f_\xi - \tilde{f}\|_{L^1}.$$

for $C = A/(1-\alpha)$ and $D = B/(1-\alpha)$, α is appearing in (13). Therefore, by (20) we have

$$\|f_\xi - \tilde{f}\|_{L^1} \leqslant \|f_{\xi,\delta} - \tilde{f}\|_{L^1} + C + D \cdot \|f_\xi - \tilde{f}\|_{L^1},$$

which implies

$$\|f_\xi - \tilde{f}\|_{L^1} \leqslant \frac{1}{1-D} \cdot \left(\|f_{\xi,\delta} - \tilde{f}\|_{L^1} + C\right).$$

4. Contraction speed estimates via coarse-fine methods

In this section we show an efficient way to estimate the rate of contraction of the discretized transfer operator and find the suitable $\overline{n}$ and α described in section 3. Since $L_{\delta,\xi}$ is represented by a matrix, a first attempt to perform this task would be to iterate and estimate the norm of the iterate. This method is not very effective, since the matrix we should iterate is quite big. For this we implement a strategy in which we get information for the full matrix from the iterates of coarser versions of it. An earlier approach to this problem for deterministic systems can be found in [14].

In lemmas 51 and 52 and corollary 49 we have seen that

$$\|(1 - \pi_\delta)N_\xi\|_{L^1 \to L^1} \leqslant \delta/\xi, \quad \|N_\xi(1 - \pi_\delta)\|_{L^1 \to L^1} \leqslant \delta/\xi, \tag{34}$$

We now prove the following lemma which bounds the distance between the powers of L_ξ and $L_{\delta,\xi}$, provided that the noise has been applied at least once before the application of L_ξ and $L_{\delta,\xi}$.

Lemma 29. *Let $\|L^i_{\delta,\xi}|_V\|_{L^1} \leqslant C_i$; let σ be a linear operator such that $\sigma^2 = \sigma$, $\|\sigma\|_{L^1} \leqslant 1$, and $\sigma\pi_\delta = \pi_\delta\sigma = \pi_\delta$; let $\Lambda = \sigma N_\xi \sigma L$.*

Then $\forall n \geqslant 0$

$$\|(L^n_{\delta,\xi} - \Lambda^n)N_\xi\|_{L^1} \leqslant \frac{\delta}{\xi} \cdot \left(2\sum_{i=0}^{n-1} C_i + 1\right) \tag{35}$$

In particular the lemma applies if:

(a) $\sigma = Id$ and $\Lambda = L_\xi$;
(b) $\sigma = \pi_{\delta'}$ and $\Lambda = L_{\delta',\xi}$, for any δ' such that $n\delta' = \delta$ with $n \in \mathbb{R}$.

As a consequence, we obtain a way to bound the contraction rate of certain operator $L_{\delta',\xi}$ on the zero-average space V using the computed contraction rate for a coarser operator $L_{\delta,\xi}$. We remark that

$$\|L^{n+1}_{\delta',\xi}|_V\|_{L^1} \leqslant \|L^n_{\delta,\xi}|_V\|_{L^1} + \|(L^n_{\delta',\xi} - L^n_{\delta,\xi})\|_{L^1}. \tag{36}$$

We remark that on the left-hand side we have an $n+1$ in (36), which guarantees that the noise has been applied at least once; this permits us to use lemma 29 to estimate the second summand of the right-hand of (36).

Thus, if one is searching for an n such that $\|L^{n+1}_{\delta',\xi}|_V\|_{L^1} < 1$ this can be found and certified by using a suitable coarse version $L_{\delta,\xi}$, computing the norm of its iterates and using lemma 29 in a way that the second hand of (36) is smaller than 1.

Proof of Lemma 29. Notice that as a consequence of the hypotheses we have

$$\|\sigma g - \pi_\delta g\|_{L^1} \leqslant \mathrm{Var}(g)\delta/2, \quad \|\sigma g - \pi_\delta g\|_W \leqslant \|g\|_{L^1}\delta/2$$

because $\sigma g - \pi_\delta g = \sigma(1-\pi_\delta)g$ applying lemmas 52 and 53.

The proof is along the lines of what has been proved in section 3.2. Indeed, we have

$$\begin{aligned}
\|(L^n_{\delta,\xi} - \Lambda^n)N_\xi\|_{L^1} \leqslant{}& \|(\pi_\delta N_\xi \pi_\delta L)^n\|_{L^1}\cdot\|(\pi_\delta - \sigma)N_\xi\|_{L^1} \\
&+ \sum_{i=0}^{n-1}\|(\pi_\delta N_\xi \pi_\delta L)^i\|_{L^1}\cdot\|(\pi_\delta-\sigma)N_\xi\|_{L^1}\cdot\|(\sigma L\sigma N_\xi)^{n-i}\|_{L^1} \\
&+ \sum_{i=0}^{n-1}\|(\pi_\delta N_\xi \pi_\delta L)^i\pi_\delta\|_{L^1}\cdot\|N_\xi(\pi_\delta-\sigma)\|_{L^1}\cdot\|L\sigma N_\xi(\sigma L\sigma N_\xi)^{n-i-1}\|_{L^1} \\
\leqslant{}& \|(\pi_\delta-\sigma)N_\xi\|_{L^1} + \sum_{i=0}^{n-1}C_i\|(\pi_\delta-\sigma)N_\xi\|_{L^1}\cdot\|(\sigma L\sigma N_\xi)^{n-i}\|_{L^1} \\
&+ \sum_{i=0}^{n-1}C_i\|N_\xi(\pi_\delta-\sigma)\|_{L^1}\cdot\|L\sigma N_\xi(\sigma L\sigma N_\xi)^{n-i-1}\|_{L^1},
\end{aligned}$$

and the thesis follows from the fact that $\|L\|_{L^1} \leqslant 1$, $\|N_\xi\|_{L^1} \leqslant 1$, $\|\sigma\|_{L^1} \leqslant 1$. □

5. Estimating the average of an observable

As a result of the previous sections, we are able to obtain a precise approximation $\tilde{f}$ of f_ξ in the L^1 norm. This is not enough in order to estimate the Lyapunov exponent which we recall can be defined as $\lambda_\xi = \int h\,\mathrm{d}f_\xi$ where $h = \log|T'|$. This is because h is not in L^∞ in the whole interval. In the Belousov–Zhabotinsky case, there are two points where h goes to infinity: the critical point and the point where the $|T'|$ goes to $+\infty$. Outside of neighbourhoods of these two points h is bounded. Moreover, in these two neighbourhoods h still has bounded L^1 norm. This is enough to perform our estimates since we can have L^∞ bounds on the stationary measure f_ξ; this allows to compute the Lyapunov exponent using alternately L^1 and L^∞ estimates on f_ξ and h in different sets. Therefore, we can join all these observations together to obtain a rigorous approximation of λ_ξ applying the following strategy:

- we select a region E of $[0, 1]$ such that h is in L^∞ outside E;
- we estimate, on E, the quantities $\|f_\xi\|_{L^\infty(E)}$ and $\|h\|_{L^1(E)}$;
- we approximate (keeping rigorously track of the numerical errors) the integral $\int_{[0,1]} h\,\mathrm{d}f_\xi$ with $\int_{[0,1]\backslash E} h\,\mathrm{d}\tilde{f}$ (discarding the set E from the computation);
- we estimate the error in such an approximation in terms of $\|f_\xi - \tilde{f}\|_{L^1}$, $\|f_\xi\|_{L^\infty(E)}$, $\|h\|_{L^1(E)}$ and $\|h\|_{L^\infty([0,1]\backslash E)}$ (see corollary 30).

In subsection 5.1 we show how the error can be estimated in terms of the mentioned quantities. The remaining subsections are devoted to estimating $\|f_\xi\|_{L^\infty(E)}$ and $\|h\|_{L^1(E)}$, notice that in the case of uniform noise $\|f_\xi\|_{L^\infty(E)} \leqslant 1/\xi$, but we are able to obtain a better estimate via $\tilde{f}$.

5.1. Approximating the average using L^1 and L^∞ estimates

In this section we assume that $\tilde{f}$ is an approximation of f_ξ and both are probability measures, therefore $\tilde{f} - f_\xi$ has 0 average.

Corollary 30. *Let f and $\tilde{f}$ be probability densities on the measure space (X, m), both contained in L^1 and in L^∞. Let $E \subset X$ be a Borel subset, and H be an L^1 observable that is L^∞ in $X\backslash E$. Then*

$$\left|\int_X Hf\,\mathrm{d}m - \int_{X\backslash E} H\tilde{f}\,\mathrm{d}m\right| \leqslant \|H\|_{L^1(E)} \cdot \|f\|_{L^\infty(E)} + \frac{\sup_{X\backslash E} H + \inf H_{X\backslash E}}{2} \cdot \|f - \tilde{f}\|_{L^1}.$$

The proof of the corollary is straightforward, applying the following Lemma on the set $X\backslash E$

Lemma 31. *Let (X, m) be a measure space, let $H \in L^\infty(X)$, and let $v \in L^1(X)$ a function having 0 average. Then we have*

$$\left|\int_X H \cdot v\,\mathrm{d}m\right| \leqslant \frac{\sup H - \inf H}{2} \cdot \|v\|_{L^1}.$$

Proof. Indeed, for a constant c we have

$$\left|\int_X Hv\,\mathrm{d}m\right| \leqslant \left|\int_X (H - c)v\,\mathrm{d}m\right| + \left|\int_X cv\,\mathrm{d}m\right| \leqslant \|H - c\|_{L^\infty} \cdot \|v\|_{L^1},$$

because v has 0 average, and this is clearly optimized taking $c = (\sup H + \inf H)/2$. □

Remark 32. Corollary 30 yields immediately an algorithm for estimating an observable that is L^1, and is L^∞ outside a neighbourhood of a finite number of points s_i where it goes to ∞, as is the case with the observable $\log|T'|$ of the system we are studying. In fact, there is a trade-off on the size of E, and we attempt different sets E enclosing the s_i with intervals of different sizes on order to obtain the tightest possible estimate on the error. Every such choice of the set E yields an approximation of $\int_X H f \mathrm{d}m$ as $\int_{X\setminus E} H\tilde{f}\,\mathrm{d}m$ and a bound for the error.

5.2. L^∞ bounds for the stationary measure in an interval

To estimate the average of the unbonded observable h and apply corollary 30, in this subsection we obtain a bound for the L^∞ norm of the invariant measure f_ξ on intervals of a certain partition. We derive it as a byproduct of the rigorous estimate of the L^1 error, and the algorithms explained in appendices B.2.1 and B.2.2, that allow to bound $\mathrm{Var}_I(N_\xi L\tilde{f})$ for each $I \in \Pi$.

Lemma 33. *Let Π be a uniform partition, for each $I \in \Pi$ we have*

$$\|f_\xi\|_{L^\infty(I)} \leqslant \mathrm{Var}_I(N_\xi L\tilde{f}) + \frac{\|N_\xi L\tilde{f}\|_{L^1(I)}}{|I|} + \|\tilde{f} - f_\xi\|_{L^1} \cdot \|\rho_\xi\|_{L^\infty}.$$

Proof. Indeed,

$$\begin{aligned}\|f_\xi\|_{L^\infty(I)} &= \|N_\xi L f_\xi\|_{L^\infty(I)} \leqslant \|N_\xi L\tilde{f}\|_{L^\infty(I)} + \|N_\xi L(\tilde{f} - f_\xi)\|_{L^\infty}\\ &\leqslant \mathrm{Var}_I(N_\xi L\tilde{f}) + \|N_\xi L\tilde{f}\|_{L^1(I)}/|I| + \|N_\xi\|_{L^1\to L^\infty} \cdot \|L(\tilde{f} - f_\xi)\|_{L^1}\end{aligned}$$

and the estimate follows because $\|N_\xi\|_{L^1\to L^\infty} \leqslant \|\rho_\xi\|_{L^\infty}$. □

5.3. L^1 bounds on $\log|T'|$

In this section we compute explicit bounds for the L^1-norm of $\log|T'|$ for the map defined in section 2 in a neighbourhood of the points where it is not bounded, as required to apply remark 32. As can be deduced from its definition in section 2, we need to do so in intervals enclosing $x = 0.125$ and $x = 0.3$. We provide the proof only of the first one, as they are all very elementary.

Lemma 34. *For $0.125 - 2^{-6} < u < 0.125 < v < 0.125 + 2^{-6}$ we have*

$$\begin{aligned}\int_u^v \log|T'(x)|, \mathrm{d}x \in &-\frac{2}{3}\left[(0.125-u)(\log(0.125-u)-1) + (v-0.125)(\log(v-0.125)-1)\right]\\ &-(v-u)\log(3) - \frac{v^2-u^2}{2} + [0, (\log(5)-\log(4))(v-u)].\end{aligned}$$

Proof. By a direct computation we have

$$T'(x) = \left(-(x-0.125)^{1/3} - a + \frac{1}{3}|x-0.125|^{-2/3}\right)\mathrm{e}^{-x},$$

and

$$\log|T'(x)| = \log\left(\frac{1}{3}|x-0.125|^{-2/3} - (x-0.125)^{1/3} - a\right) - x.$$

Notice that for $|x - 0.125| < 2^{-6}$ we have

$$\frac{1}{3}|x-0.125|^{-2/3} \geqslant \frac{1}{3}2^4 \geqslant 5$$

while

$$a \leqslant \left|(x-0.125)^{1/3}+a\right| \leqslant 2^{-2}+a \leqslant 1.$$

Consequently in the interval $[0.125-2^{-6}, 0.125+2^{-6}]$ we have that $\log|T'(x)|>0$, and furthermore

$$\log|T'(x)| \leqslant \log\left(\frac{1}{3}|x-0.125|^{-2/3}\right)-x=-\frac{2}{3}\log|x-0.125|-\log(3)-x.$$

It follows that for $0.125-2^{-6}<u<0.125<v<0.125+2^{-6}$ we have

$$\int_u^v \log|T'(x)| \leqslant -\frac{2}{3}\left[(0.125-u)(\log(0.125-u)-1)+(v-0.125)(\log(v-0.125)-1)\right]$$
$$-(u-v)\log(3)-\frac{v^2-u^2}{2}.$$

In the same way we have that

$$\log|T'(x)| \geqslant \log\left(\frac{1}{3}|x-0.125|^{-2/3}-1\right)-x$$

and since

$$\frac{1}{5}\frac{1}{3}|x-0.125|^{-2/3} \geqslant 1$$

we have that

$$\log|T'(x)| \geqslant \log\left(\frac{4}{5}\frac{1}{3}|x-0.125|^{-2/3}\right)-x=\log\frac{4}{5}+\log\left(\frac{1}{3}|x-0.125|^{-2/3}\right)-x$$

consequently subtracting $(\log(5)-\log(4))(v-u)$ the same value above is also valid as a lower bound. □

Lemma 35. *For* $0.2<x<0.3$ *we have*

$$\int_x^{0.3} \log|T'(x)|, \mathrm{d}x \in (\log([d_1,d_2])+\log(0.3-x)-1)(0.3-x)-\frac{1}{2}(0.3^2-x^2)$$

for $d_1=\frac{1}{3}(\frac{2}{3}\cdot(0.175)^{-5/3}+(0.175)^{-2/3})$ *and* $d_2=\left[a+(x-0.125)^{1/3}-\frac{1}{3}|x-0.125|^{-2/3}\right]/(0.3-x)$*, where* $a=0.506\,073\,56\ldots$ *as in section 2.*

Lemma 36. *For* $0.3<x<0.303$ *we have*

$$\int_{0.3}^x \log|T'(x)|\mathrm{d}x=\left(\log a+\log 19+19\log 10-38+\log(10/3)\right)\cdot(x-0.3)$$
$$+18(x\log(x)+0.3\log(0.3))+(x-0.3)\log(x-0.3)-\frac{190}{6}(x-0.3)^2.$$

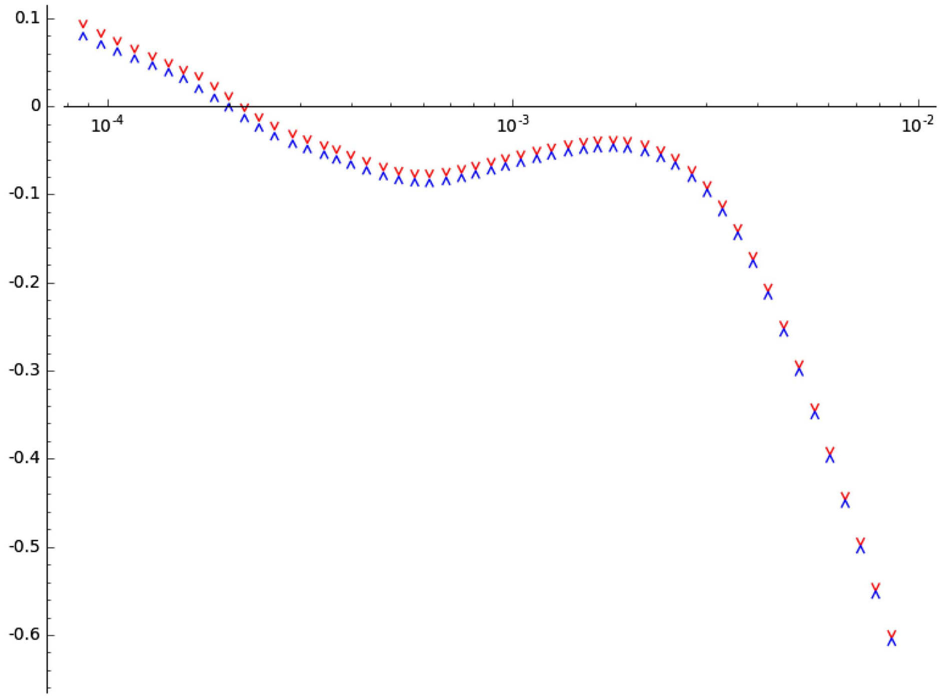

Figure 4. A plot of the intervals enclosing the values of the Lyapunov exponent for several sizes of the noise. The plotted values are listed in Table 1.

6. Computation details and results

We give here some details about the code performing our computer aided estimates and about the results. The main algorithm is written in Python using the Sage framework and interval arithmetics ([31]), some critical parts are written in C++ and uses (optionally) the GPU, this in particular is used for the iteration of large matrices[6] needed to apply the methods of section 4. Such parts have been optimized to use high performance computing; even so each contraction test has required a time of the of order of a week[7]. The code we used can be found at http://im.ufrj.br/∼maurizio.monge/wordpress/rigorous_computation_dyn/.

Table 1 contains the result of the computer aided estimates we performed and the values of the parameters used in these estimates. In figure 4 we summarize with a graph the most important information contained in the table. The graph shows intervals enclosing the Lyapunov exponent at the selected noise values. These are the final result of our computer aided estimates; it is worth to remark again that the estimated requiring more computational power are the ones performed to prove that the Lyapunov exponent is positive for small size of the noise.

[6] The matrices in this part of the computation have about 2^{20} lines and columns, see the second column of table 1 for the inverse of the size of the matrix at each computer aided estimate.

[7] The contraction time estimates were run on an Asus GeForce GTX 1050Ti, 4 GB of Ram GPU installed in a desktop computer with an AMD A4-6300 3.4 GHz processor and 8 GB of Ram. The matrices were assembled on a Dell R710 server with 2 sixcore Xeon 5660 2.8 GHz processors and 24 GB of Ram.

To ease the understanding of table 1, we now outline some more details of the implementation of our algorithm and how the parameters take place in the algorithm's execution; the columns are ordered as they are subsequently used in the algorithm, or deduced from previous quantities and via computations.

One of the main ingredients and goals of the computer aided estimates is to compute the number of iterates of the transfer operator needed to contract the zero average space (see item 1 of section 3.1). For this we apply the 'coarse fine' methods explained in section 4. For each value of the noise amplitude, denoted by ξ, we build a coarse discretization $L_{\xi,\delta_{\text{contr}}}$ of the operator L_ξ, on a partition of coarse size δ_{contr}. We then compute values of n_{contr} and α_{contr} represented in table 1, which satisfy

$$\|L^{n_{\text{contr}}}_{\xi,\delta_{\text{contr}}}|_V\|_{L^1} \leqslant \alpha_{\text{contr}} < 1;$$

and compute explicit bounds for $\|L^i_{\xi,\delta_{\text{contr}}}|_V\|_{L^1}$ (whose value does not appear in the table). The algorithm used for these finite dimensional estimates is the same as in [13] and there explained.

We consider then a finer partition size δ and use the coarse fine estimates of section 4 to compute α and $\sum_{i=0}^{n_{\text{contr}}} C_i$, where $C_i = \|L^i_{\xi,\delta}|_V\|_{L^1}$ and

$$\|L^{n_{\text{contr}}+1}_{\xi,\delta}|_V\|_{L^1} \leqslant \alpha.$$

The same bounds work for L_ξ, i.e.

$$\|L^{n_{\text{contr}}+1}_{\xi}|_V\|_{L^1} \leqslant \alpha, \quad \text{and} \quad \sum_{i=0}^{n_{\text{contr}}} \|L^i_\xi|_V\|_{L^1} \leqslant \sum_{i=0}^{n_{\text{contr}}} C_i.$$

Notice that the bigger the n_{contr}, the worse is going to be the estimate on the L^1 norm on $(L^n_\xi - L^n_{\xi,\delta\ \text{contr}})N_\xi$, i.e., the error coming from the coarse-fine inequality. While increasing n_{contr} permits us to find smaller α_{contr}, this may not imply that the corresponding α is smaller. Our algorithms attempts to find the best compromise. Needless to say, in practice this procedure may fail, if unable to detect any contraction in a reasonable time. This might happen for example if the original system is not mixing.

To estimate an upper bound to the L^1 error in the computation of the stationary measure we use the results shown in section 3. The column '*a priori*...' contains the *a priori* estimate on the L^1 error of the approximation of the measure on the partition of size δ as given using the results in subsection 3.2.1 while the column 'refined...' contains the L^1 error when we use the bootstrapping techniques of subsection 3.3.

Once we have a good approximation of the invariant measure, we compute an approximation of the Lyapunov exponent by computing an integral on a partition of size δ_{est}, using section 5; the computed intervals enclosing rigorously the Lyapunov exponent are contained in the last column. This allows to prove theorem 2.

Proof of Items I1 and I2 of Theorem 2. We refer to the values listed in table 1. For a noise size of $\xi_1 = 0.873 \times 10^{-4}$. By the results of sections 3–5 our algorithm certifies that the Lyapunov exponent $\lambda_{\xi_1} \in [8.365 \times 10^{-2}, 8.917 \times 10^{-2}]$. In particular, this proves that $\lambda_{\xi_1} > 0$.

For a noise size of $\xi_2 = 0.860 \times 10^{-2}$ our algorithm certifies that the Lyapunov exponent $\lambda_{\xi_2} \in [-6.036\,02 \times 10^{-1}, -6.035\,36 \times 10^{-1}]$. In particular, this proves that $\lambda_{\xi_2} < 0$. □

We remark that in the case of random diffeomorphisms there exists a dychotomy [23]: if the Lyapunov exponent is positive the system admits a *random strange attractor* (chaotic behaviour) while if the Lyapunov exponent is negative the system has a *random sink* (regular behaviour).

7. Quantitative stability of the system and of the Lyapunov exponent

In this section we study of the regularity of λ_ξ as a function of ξ and prove that this varies $\alpha-$Holder continuously for every $\alpha < 1$. We start showing a simple Lipschitz stability result for the fixed point of a Markov operator. This will show that the stationary measure is Lipschitz stable in L^1 when the noise amplitude change. This is not sufficient to deduce that the Lyapunov exponent is Lipschitz stable, because as we have seen the Lyapunov exponent is the average of an observable which is unbounded. In this case a control in L^1 of the stationary measure is not sufficient to control its average. For this reason we strengthen the estimates to an L^p quantitative stability statement with $p > 1$ which is sufficient to control the average of our observable.

7.1. Lipschitz stability of the stationary measure

A quantitative stability statement for the stationary measure follows from a general and elementary lemma about perturbations of Markov operators:

Lemma 37. *Let $L_1, L_2 : L^1 \to L^1$ be two Markov operators. Assume*

$$\|L_1^i|_V\|_{L^1\to L^1} \leqslant C_i$$

and suppose that $C_N < 1$ for a certain N. Let f_i be a fixed probability measure of L_i, then $\sum_{k=0}^{N-1} C_i < \infty$ and

$$\|f_2 - f_1\|_{L^1} \leqslant \frac{\sum_{k=0}^{N-1} C_i}{1 - C_N} \cdot \|L_1 - L_2\|_{L^1}. \tag{37}$$

Proof. The existence of N such that $C_N < 1$ easily implies that C_i decreases exponentially and $\sum_{k=0}^{N-1} C_i < \infty$. Since f_0, f_1 are fixed probability measures

$$\begin{aligned}\|f_2 - f_1\|_{L^1} &\leqslant \|L_2^N f_2 - L_1^N f_1\|_{L^1} \leqslant \|L_2^N f_2 - L_1^N f_2\|_{L^1} + \|L_1^N f_2 - L_1^N f_1\|_{L^1}\\ &\leqslant \|L_1^N(f_2 - f_1)\|_{L^1} + \|L_2^N f_2 - L_1^N f_2\|_{L^1}.\end{aligned}$$

Since $C_i \to 0$, and $f_2 - f_1 \in V$, we can chose N such that $C_N < 1$, we have $\|L_1^N(f_2 - f_1)\|_{L^1} \leqslant C_N\|f_2 - f_1\|_{L^1}$ and

$$\|f_2 - f_1\|_{L^1} \leqslant \frac{\|L_2^N f_2 - L_1^N f_2\|_{L^1}}{1 - C_N}.$$

Let us now consider the term $\|L_2^N f_2 - L_1^N f_2\|_{L^1}$. Since

$$(L_1^N - L_2^N) = \sum_{k=1}^{N} L_1^{N-k}(L_1 - L_2)L_2^{k-1}$$

then

$$\begin{aligned}-(L_2^N - L_1^N)f_2 &= \sum_{k=1}^{N} L_1^{N-k}(L_1 - L_2)L_2^{k-1} f_2\\ &= \sum_{k=1}^{N} L_1^{N-k}(L_1 - L_2)f_2\end{aligned}$$

and we have the statement. □

7.1.1. Stability of the measure under perturbation of the noise. Now let us see that the stationary measure also varies in a Lipschitz way with respect to perturbations of the noise: suppose T satisfies setting 19 and let ρ_1 and ρ_2 be two bounded variation noise kernels and associated transfer operators $L_i(f) = \rho_i \hat{*} L_T(f)$ for $i = 1, 2$.

Lemma 38. *Let L_i be defined as above, then*

$$\|(L_1 - L_2)f\|_{L^1} \leqslant \|\rho_1 - \rho_2\|_{L^1} \cdot \|f\|_{L^1}$$

Proof. Indeed,

$$\|(L_1 - L_2)f\|_{L^1} \leqslant \|[\rho_0 - \rho_1]\hat{*} L_T(f)\| \leqslant \|\rho_1 - \rho_2\|_{L^1} \cdot \|f\|_{L^1}.$$

□

From this and the classical L^p interpolation inequality, in the case of dynamical systems with additive noise we get the following L^p stability estimate

Corollary 39. *Let $L_1, L_2 : L^1 \to L^1$ be two transfer operators of deterministic systems with additive noise $L_i(f) = \rho_i \hat{*} L_T(f)$ for $i = 1, 2$. Assume*

$$\|L_1^i|_V\|_{L^1 \to L^1} \leqslant C_i$$

and suppose that $C_N < 1$ for a certain N. Let f_i be a fixed probability measure of L_i and Suppose $1 \leqslant r < \infty$. Then

$$\|f_2 - f_1\|_{L^r} \leqslant \left(\frac{\sum_{k=0}^{N-1} C_i}{1 - C_N} \cdot \|\rho_1 - \rho_2\|_{L^1} \right)^{\frac{1}{r}} (2\max(\|\rho_1\|_{BV}, \|\rho_2\|_{BV}))^{1-\frac{1}{r}}. \quad (38)$$

Proof. We get that for $i = 1$ and $i = 2$, $\|f_i\|_{L^\infty} \leqslant \|f_i\|_{BV} \leqslant \max(\|\rho_1\|_{BV}, \|\rho_2\|_{BV})$. Suppose $1 \leqslant r < \infty$ and $u \in L^1 \cap L^\infty$, the classical L^p interpolation inequality implies that $u \in L^r$ and

$$\|u\|_{L^r} \leqslant \|u\|_{L^1}^{\frac{1}{r}} \|u\|_{L^\infty}^{1-\frac{1}{r}}.$$

Applying this to (37)

$$\|f_2 - f_1\|_{L^r} \leqslant \left(\frac{\sum_{k=0}^{N-1} C_i}{1 - C_N} \cdot \|L_1 - L_2\|_{L^1} \right)^{\frac{1}{r}} (2\max(\|\rho_1\|_{BV}, \|\rho_2\|_{BV}))^{1-\frac{1}{r}}. \quad (39)$$

using the estimate for $\|L_1 - L_2\|_{L^1}$ given in lemma 38 we get (38). □

Lemma 40. *Suppose ρ_1, ρ_2 are the uniform kernel $\rho_1 = \xi^{-1} 1_{[-\xi/2, \xi/2]}$, $\rho_2 = \tilde{\xi}^{-1} 1_{[-\tilde{\xi}/2, \tilde{\xi}/2]}$ we have*

$$\|\rho_1 - \rho_2\|_{L^1} \leqslant \frac{2}{\max\{\xi, \tilde{\xi}\}} |\xi - \tilde{\xi}|.$$

Proof. Indeed,

$$\begin{aligned} \|\rho_1 - \rho_2\|_{L^1} &\leqslant \min\{\xi, \tilde{\xi}\} \left| \frac{1}{\xi} - \frac{1}{\tilde{\xi}} \right| + \frac{|\xi - \tilde{\xi}|}{\max\{\xi, \tilde{\xi}\}} \\ &= \frac{2|\xi - \tilde{\xi}|}{\max\{\xi, \tilde{\xi}\}}. \end{aligned}$$

□

If $\xi, \tilde{\xi} \geqslant \xi_1 = \frac{8.73}{10^5}$ (see theorem 2) we get $\|\rho_1 - \rho_2\|_{L^1} \leqslant \frac{2\times 10^5}{8.73}|\xi - \tilde{\xi}|$ and putting this in (38) considering that in this case $\|\rho_i\|_{BV} \leqslant \frac{410^5}{8.73} + 1$ we get

$$\|f_2 - f_1\|_{L^r} \leqslant \left(\frac{\sum_{k=0}^{N-1} C_i}{1 - C_N} \cdot \frac{2\times 10^5}{8.73}|\xi - \tilde{\xi}| \right)^{\frac{1}{r}} \left(\frac{4\times 10^5}{8.73} + 1 \right)^{1-\frac{1}{r}}. \tag{40}$$

We recall that the numbers C_i represents the contraction rate of one of the two operators. We show that even for these quantities we can have an uniform estimate when $\xi, \tilde{\xi} \geqslant \xi_1$. The following result allows to estimate the C_i constants (such that $\|L^i_{\xi}|_V\|_1 \leqslant C_i$) when the amplitude of the noise is increased.

Lemma 41. *Let* $\rho_\xi = \xi^{-1} 1_{[-\xi/2,\xi/2]}$ *and* N_ξ *the associated noise operator. If* $\hat{\xi} > \xi$ *and* $\|(N_\xi L)^i\|_{V\to L^1} \leqslant C_i < 1$, *then*

$$\|(N_{\hat{\xi}} L)^i\|_{V\to L^1} \leqslant C_i(\xi/\hat{\xi})^i + \left[1 - (\xi/\hat{\xi})^i\right] < 1. \tag{41}$$

Proof. Let $\hat{\xi} = \xi + \epsilon$ and $\rho_\xi = \xi^{-1} 1_{[-\xi/2,\xi/2]}$, we have

$$\begin{aligned} \rho_{\xi+\epsilon} &= (\xi + \epsilon)^{-1} 1_{[-(\xi+\epsilon)/2,(\xi+\epsilon)/2]} \\ &= \frac{\xi}{\xi+\epsilon}\rho_\xi + \frac{\epsilon}{\xi+\epsilon} \cdot \frac{1}{\epsilon} 1_{[-(\xi+\epsilon)/2,-\xi/2]\cup[\xi/2,(\xi+\epsilon)/2]} \end{aligned}$$

Therefore

$$N_{\xi+\epsilon} = \frac{\xi}{\xi+\epsilon} N_\xi + \frac{\epsilon}{\xi+\epsilon} M$$

where M is the Markov operator of convolution with $\epsilon^{-1} 1_{[-(\xi+\epsilon)/2,-\xi/2]\cup[\xi/2,(\xi+\epsilon)/2]}$.

In the same way we have that $N_{\hat{\xi}} L$ is a convex combination of the Markov operators $N_\xi L$ and ML with coefficients $\xi/\hat{\xi}$ and $1 - \xi/\hat{\xi}$, and

$$(N_{\hat{\xi}} L)^i = (\xi/\hat{\xi})^i (N_\xi L)^i + \left[1 - (\xi/\hat{\xi})^i\right] Q$$

for a suitable Markov operator Q formed by the remaining terms of the expansion. Considering the L^1 norm the inequality follows. □

We remark that when $C_i < 1$ it holds that the right hand of (41) is smaller than 1, by this we have immediately the following corollary

Corollary 42. *Let* ρ_ξ *and* N_ξ *as above. Suppose* $\|(N_\xi L)^i\|_{V\to L^1} < 1$, *then for each* $\xi \leqslant \hat{\xi} \leqslant 1$ *it holds*

$$\|(N_{\hat{\xi}} L)^i\|_{V\to L^1} < 1$$

and the system is mixing for every noise greater than ξ.

By continuity of the above estimates when $\hat{\xi}$ varies and compactness of $[\xi_1, 1]$ if follows that $\sum_{k=0}^{N-1} C_i$ and C_N have an uniform bound on $[\xi_1, 1]$ and then for each $r \geqslant 1$ there is $C \geqslant 0$ such that for each $\xi - \tilde{\xi} \in [\xi_1, 1]$

$$\|f_2 - f_1\|_{L^r} \leqslant C|\xi - \tilde{\xi}|^{\frac{1}{r}} \tag{42}$$

proving the Hőlder stability of the stationary measure in L^r.

7.2. Stability of the Lyapunov exponent

Now we see how from the Hőlder stability of the stationary measure proved in the previous section we can easily deduce the Hőlder stability of the Lyapunov exponent λ_ξ and the Hőlder continuity of λ_ξ as ξ varies. We recall that $\lambda_\xi := \int_0^1 \phi(x)\mathrm{d}\mu_\xi$ where $\phi(x) = \log|T'(x)|$.

Corollary 43. *For each $r > 1$ we have that $\phi \in L^{\frac{r}{r-1}}$ and there is $C \geqslant 0$ such that for each $\xi, \tilde{\xi} \in [\xi_1, 1]$,*

$$|\lambda_\xi - \lambda_{\tilde{\xi}}| \leqslant \|\phi\|_{L^{\frac{r}{r-1}}} C|\xi - \tilde{\xi}|^{\frac{1}{r}}.$$

Proof. We can get an explicit formula for ϕ, indeed

$$T'(x) = \begin{cases} -\dfrac{1}{6}\dfrac{\mathrm{e}^{-x}}{(8x-1)^{\frac{2}{3}}}\left(24x + 6a(8x-1)^{\frac{2}{3}} - 11\right) & 0 \leqslant x \leqslant 0.3 \\ -\dfrac{19 \times 10^{19}}{3} cx^{18}\mathrm{e}^{-\frac{190}{3}x}(10x-3) & 0.3 < x \leqslant 1 \end{cases}$$

by this $\phi \in L^p[0,1]$ for each $p \in [1, \infty)$ and then by the holder inequality

$$|\lambda_\xi - \lambda_{\tilde{\xi}}| \leqslant \int_0^1 \phi(x)[f_1 - f_2]\mathrm{d}m \leqslant \|\phi\|_{L^{\frac{r}{r-1}}} \|f_2 - f_1\|_{L^r}$$

from which the statement follows applying (42). □

Acknowledgments

The authors thank Yuzuru Sato for the nice seminar where we learned about the existence of Noise Induced Order and further explanations, EU Marie-Curie IRSES ‘Brazilian-European partnership in Dynamical Systems’ (FP7-PEOPLE-2012-IRSES 318999 BREUDS) for support during the research and The Abdus Salam International Centre for Theoretical Physics (ICTP) where the work started. IN was partially supported by CNPq, University of Uppsala and KAW Grant 2013.0315. IN thanks UFRJ, CAPES (through the programs PROEX and the CAPES-STINT project ‘Contemporary topics in non uniformly hyperbolic dynamics’ 88881.155748/2017-01). SG thanks Leverhulm Trust grant (IN-2014-021, ‘Statistical properties of non uniformly hyperbolic dynamical systems: computer assisted proofs and rigorous computation’) and GNAMPA/INDAM (‘Stabilita e instabilita in sistemi con rumore, un approccio computer assistito.’)

Appendix A. Generalities, ergodicity and Lyapunov exponents in random dynamics

In this section we recall some basic results and definitions in the ergodic theory of random transformations and the integral formula for the Lyapunov exponent. We refer to [32, chapter 5]; another classical reference is [1].

Let X be the interval $[-\xi, \xi]$ endowed with the Borel σ-algebra and p the uniform probability on X; let $\Xi = X^{\mathbb{N}}$ the space of sequences with values in this space endowed with the product σ-algebra Ω and the product measure $\mathbb{P} = p^{\mathbb{N}}$. Let ϕ be the shift acting on Ξ.

We endow the interval [0, 1] with the Borel σ-algebra β and we define the measurable skew product:

$$F : (\Xi, \Omega) \times ([0,1], \beta) \to (\Xi, \Omega) \times ([0,1], \beta) \quad F(\omega, x) = (\phi(\omega), T(x) + (\omega)_0).$$

This skew product models the evolution of the stochastic process

$$X_{n+1} = T(X_n) + \omega_n.$$

where ω_n is a sequence of i.i.d. random variables uniformly distributed in $[-\xi, \xi]$ endowed with the Borel σ-algebra.

In the following, let L_ξ be the annealed transfer operator as defined in section 2. This operator embodies how measures behave 'in average' under the action of the random dynamical system; let μ be a measure on [0, 1] and $T_\epsilon(x) = T(x) + \epsilon$, we have that

$$L_\xi \mu(B) = \int \mu(T_\epsilon^{-1}(B)) \mathrm{d}p(\epsilon).$$

Since ϕ is the one sided shift and ν is a stationary measure, i.e., $L_\xi \nu = \nu$, the product measure $\mathbb{P} \times \nu$ is invariant for F [32, proposition 5.4]. Since the transfer operator related to convolution with a bounded variation kernel is regularizing from L^1 to BV which is compactly immersed in L^1 it is easy to see that the transfer operator has at least one stationary measure f_ξ with density in BV (see [12] lemma 23 for more details).

Definition 44. A stationary measure ν is said to be ergodic if the measure $\mathbb{P} \times \nu$ is ergodic for F.

Remark 45. While we use this as our definition of ergodicity for the random dynamic, we refer to [32] for equivalent alternative definitions.

Proposition 46. *Let L_ξ be the transfer operator associated to the Belousov–Zhabotinsky map with additive noise of size ξ. Let $\xi_1 = 0.873 \times 10^{-4}$. For each $\xi \geqslant \xi_1$ there exists a unique, ergodic stationary measure μ_ξ for the operator L_ξ, and for $\mathbb{P} \times \mu_\xi$ almost every point (ω, x), we have that*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \log |T'^i(\omega, x))| = \int_0^1 \log(|T'(x)|) \mathrm{d}\mu_\xi(x) =: \lambda_\xi$$

Proof. The operator L_ξ is defined from SM(X) to $L^1([0, 1])$; therefore, all of its fixed points belong to $L^1([0, 1])$.

In the last row of table 1 the columns n and α shows that for the noise of amplitude ξ_1 it holds

$$\|L_{\xi_1}^{75}\|_{V \to L^1} \leqslant 0.55 \tag{43}$$

(the n in the table is valid for both the approximation L_{δ,ξ_1} and the original operator L_{ξ_1}, as explained after lemma 29).

Suppose f, g are two fixed probability measures for the operator, by (43):

$$\|f - g\|_{L^1} = \|L_{\xi_1}^{75}(f - g)\|_{L^1} \leqslant 0.55 \|f - g\|_{L^1}$$

which implies that $\|f - g\|_{L^1} = 0$. Thus L_{ξ_1} has a unique fixed probability measure μ_{ξ_1} in L^1, i.e., the only stationary measure is μ_{ξ_1}. Now the same holds for every $\xi \geqslant \xi_1$ thanks to corollary

42. The corollary indeed implies that $\|L_\xi^{75}\|_{V\to L^1} < 1$ and we can repeat the same reasoning as above, obtaining an unique stationary measure for each such L_ξ.

A kind of ergodic decomposition theorem is established for stationary measures [32, theorem 5.13]; the theorem states that every stationary measure can be written as a convex combination of ergodic stationary measures. Since μ_ξ is the unique stationary measure, it follows that μ_ξ is an ergodic stationary measure. Therefore $\mathbb{P} \times \mu_\xi$ is an ergodic measure for F and the result follows applying Birkhoff ergodic theorem to F. □

Appendix B. Operator norms and variation estimates

In this section we prove several technical lemmas and estimates about operator norms and variation of iterates of measures which are used in section 3.

B.1. Operator norms

The following lemma allows to estimate the variation of a boundary reflecting convolution.

Lemma 47. *Let $\rho_\xi(x)$ be a real function with bounded variation and support contained in $(-\xi, \xi)$. Let $b(x)$ a function with zero average on the unit interval. We have*

$$\mathrm{Var}(\rho_\xi \hat{*} f) \leqslant \mathrm{Var}(\rho_\xi) \cdot \|f\|_{L^1([0,1])}.$$

As a consequence we have that

$$\|N_\xi\|_{L^1\to\mathrm{Var}} \leqslant \mathrm{Var}(\rho_\xi) = \xi^{-1}\,\mathrm{Var}(\rho).$$

Proof. Since the support of $\hat{f}$ is contained in $[0, 1]$ we have that the support of $\rho_\xi * \hat{f}$ is contained in $[-\xi, 1 + \xi]$ and that:

$$\rho_\xi \hat{*} f(x) = \rho_\xi * \hat{f}(x) + \rho_\xi * \hat{f}(-x) + \rho_\xi * \hat{f}(2 - x).$$

We recall that:

$$\mathrm{Var}_{[-1,2]}(\rho_\xi * \hat{f}) \leqslant \mathrm{Var}(\rho_\xi) \cdot \|\hat{f}\|_{L^1([-1,2])} = \mathrm{Var}(\rho_\xi) \cdot \|f\|_{L^1([0,1])};$$

which implies

$$\mathrm{Var}_{[0,1]}(\rho_\xi \hat{*} f) \leqslant \mathrm{Var}_{[0,1]}(\rho_\xi * \hat{f}) + \mathrm{Var}_{[-1,0]}(\rho_\xi * \hat{f}) + \mathrm{Var}_{[1,2]}(\rho_\xi * \hat{f}) \leqslant \mathrm{Var}(\rho_\xi) \cdot \|f\|_{L^1([0,1])}.$$

□

The following lemma is a small improvement (by a factor 2) of a lemma which has already been used in [13, 22]. While it seems to be folklore, we prove it here for a matter of completeness.

Lemma 48. *For each $f \in \mathrm{BV}$*

$$\|f - \pi_\delta f\|_{L^1} \leqslant \frac{\delta}{2}\,\mathrm{Var}(f).$$

In other words

$$\|1 - \pi_\delta\|_{\mathrm{Var}\to L^1} \leqslant \delta/2.$$

Proof. Let I be an interval of the partition Π, and assume f to have variation v in I. In I, $\pi_\delta(f)$ is constant and equal to the average of f in I.

Subtracting a constant we can assume f to have average 0 in I, so that we will be estimating $\|f\|_{L^1(I)}$ assuming $\pi_\delta f$ to be 0 in I. We will disregard the variation at the boundary of I, and will only assume a bound on the variation on the interior.

We start by supposing f to be piecewise constant, with two pieces: f is $-x$ on an interval of size a and $v-x$ in an interval of size $\delta-a$. Since the average is 0 we have that

$$xa = (v-x)(\delta-a),$$

that implies $x = v - av/\delta$. The L^1 norm in I is

$$2xa = 2a(v - av/\delta),$$

and has derivative with respect to a equal to

$$2v - 4av/\delta,$$

which becomes zero for $a = \delta/2$. Consequently $x = v/2$, and a variation v contributed an $L^1(I)$ norm of $2ax = v\delta/2$.

We claim that the biggest ratio $\|f\|_{L^1(I)}/v$ is attained when f is piecewise constant attaining exactly the values $-x$ and $v-x$. Indeed, if this was not the case, we could build a new function $\tilde{f}$ selecting the region where f is nonnegative (or nonpositive), and setting as value the average of f in this region. In this way we obtain a $\tilde{f}$ that has the same $L^1(I)$-norm, but smaller variation. If these two regions are not a partition of I in two intervals, then the difference between the maximum and the minimum of f is smaller than v, and again the f is not optimal.

Applying this estimate to all the intervals of the partition we have that a total variation of v can give a total L^1 norm of $v\delta/2$, and consequently we have the lemma. □

By lemma 48, since $\|N_\xi\|_{L^1\to \mathrm{Var}} = \xi^{-1}\,\mathrm{Var}(\rho)$ by lemma 47 we get:

Corollary 49. *With the notations defined above, we have*

$$\|(1-\pi_\delta)N_\xi\|_{L^1\to L^1} \leqslant \frac{1}{2}\delta\xi^{-1}\,\mathrm{Var}(\rho).$$

This corollary is used in section 3.2.1.

B.1.1. Estimate for $\|N_\xi(1-\pi_\delta)\|_{L^1}$. To estimate this item (necessary in section 3.2.1) we will use the W norm, defined in definition 22.

Proposition 50. *We have*

$$\|N_\xi(1-\pi_\delta)\|_{L^1\to L^1} \leqslant \frac{1}{2}\delta\xi^{-1}\,\mathrm{Var}(\rho).$$

The proof will be postponed to the following lemmas. The first lemma relates the convolution with the $\|\|_W$ norm.

Lemma 51. *Let $a(x)$ be a real function with bounded variation with support contained in $(-1/2, 1/2)$, and $b(x)$ supported in $[0,1]$ and with zero average. We have*

$$\|a * b\|_{L^1([-1,2])} \leqslant \mathrm{Var}(a)\cdot\|b\|_W.$$

As a consequence we have that

$$\|a \hat{*} b\|_{L^1([0,1])} \leqslant \mathrm{Var}(a)\cdot\|b\|_W$$

and therefore

$$\|N_\xi\|_{W\to L^1} \leqslant \mathrm{Var}(\rho_\xi) = \xi^{-1}\,\mathrm{Var}(\rho).$$

Proof. Let's prove the lemma assuming first that $a(x)$ is absolutely continuous. Let $B(x) = \int_0^x b(t)\mathrm{d}t$; integrating by parts we have

$$(a*b)(x) = \int_{-1}^{1} a(-t)b(x+t)\mathrm{d}t = [a(-t)B(x+t)]_{-1}^{+1} - \int_{-1}^{1} -a'(-t)B(x+t)\mathrm{d}x,$$

the boundary term being 0 because $b(t)$ is zero-average in the interval.

The support of $a{*}b$ is contained in $[-1/2, 3/2]$; we compute now:

$$\begin{aligned}\|a*b\|_{L^1([-1,2])} &= \int_{-1}^{2}\left|\int_{-1}^{1} a'(t)B(x-t)\mathrm{d}t\right|\,\mathrm{d}x \leqslant \int_{-1}^{2}\int_{-1}^{1} |a'(t)B(x-t)|\,\mathrm{d}t\,\,\mathrm{d}x \\ &= \int_{-1}^{1}\int_{-1-t}^{2-t} |a'(t)B(u)|\,\mathrm{d}u\,\mathrm{d}t \leqslant \int_{-1}^{1} |a'(t)|\mathrm{d}t \int_{0}^{1} |B(u)|\,\mathrm{d}u\end{aligned}$$

(putting $u = t - x$ and using that $B(u)$ has support in $[0, 1]$)

$$\leqslant \int_{-1}^{1} |a'(t)|\mathrm{d}t \cdot \int_0^1 |B(u)|\mathrm{d}u \leqslant \mathrm{Var(a)}\cdot\|B\|_{L^1} \leqslant \mathrm{Var(a)}\cdot\|b\|_W.$$

When $a(x)$ is not absolutely continuous, let's just choose absolutely continuous functions a_n such that $a_n \to a$ in L^1 and $\mathrm{Var}(a_n) \to \mathrm{Var}(a)$, and apply Fatou's lemma.

Now, observing as before that:

$$\rho_\xi \hat{*} f(x) = \rho_\xi * \hat{f}(x) + \rho_\xi * \hat{f}(-x) + \rho_\xi * \hat{f}(2-x).$$

we have that

$$\|\rho_\xi \hat{*} f\|_{L^1([0,1])} \leqslant \|\rho_\xi * \hat{f}\|_{L^1([-1,2])} \leqslant \mathrm{Var}(\rho_\xi)\|f\|_W.$$

□

To prove the proposition 50 we also need a bound for $\|1 - \pi_\delta\|_{L^1\to W}$.

Lemma 52. *For the Ulam discretization of size δ we have*

$$\|1 - \pi_\delta\|_{L^1\to W} \leqslant \delta/2.$$

Proof. We will prove the lemma is true on the space of all measures in the interval. Then we can view any measure as a combination of point masses, using that $1 - \pi_\delta$ is a linear operator on signed measures, and W is a norm on zero-average measures.

Let Δ_t be the atomic measure centred in t and with weight 1 (Kronecker's δ_t, we use capital Δ to avoid confusion), then $(1 - \pi_\delta)\Delta_t = \Delta_t - \delta^{-1}\chi_I$, where $I = (p_i, p_{i+1})$ is the interval of the δ-sized partition containing t. To compute its W-norm we need to compute the L^1 norm of

$$\begin{aligned} u_t(x) &= \int_0^x \left(\Delta_t(y) - \delta^{-1}\chi_I(y)\right)\,\mathrm{d}y \\ &= \begin{cases} \delta^{-1}(p_i - x) & \text{for } x \in [p_i, t], \\ \delta^{-1}(p_{i+1} - x) & \text{for } x \in [t, p_{i+1}], \\ 0 & \text{elsewhere.} \end{cases}\end{aligned}$$

Its L^1 norm is computed as

$$\|u_t\|_{L^1} = \delta^{-1}\int_{p_i}^{t}(x-p_i)\mathrm{d}x + \delta^{-1}\int_{t}^{p_{i+1}}(p_{i+1}-x)\mathrm{d}x = \delta^{-1}\frac{1}{2}(t-p_i)^2 + \delta^{-1}\frac{1}{2}(p_{i+1}-t)^2 \leqslant \frac{\delta}{2}$$

because it is a quadratic function that reaches its maximum for $t \in \{p_i, p_{i+1}\}$, where its value is exactly $\delta/2$.

Now, the general case. We are applying a linear operator and a norm on measures, consequently we are applying a weak-$*$ lower semi-continuous function. We verified that such function is $\leqslant \delta/2$ on atomic measures, and this holds for all finite combinations of atomic measures. Since finite combinations of atomic measures are weak-$*$ dense in the space of all measure we have the lemma. □

Proof of Proposition 50. By lemmas 51 and 52 we have

$$\|N_\xi(1-\pi_\delta)\|_{L^1} \leqslant \|1-\pi_\delta\|_{L^1\to W}\cdot\|N_\xi\|_{W\to L^1} \leqslant \frac{1}{2}\delta\xi^{-1}\,\mathrm{Var}(\rho).$$

□

B.1.2. An estimate for $\|(1-\pi_\delta)|_X\|_{\mathrm{Var}\to W}$.

Lemma 53. *We have*

$$\|(1-\pi_\delta)\|_{\mathrm{Var}(X)\to W(X)} \leqslant \frac{\delta^2}{8}.$$

for each $X \subseteq [0,1]$ *that is a union of intervals of the partition.*

Proof. Let $g \in L^1$, let us estimate $\|(1-\pi_\delta)g\|_W$; assume g to have support contained in an interval of the partition I, and subtracting $\pi_\delta g$ assume its average in I to be 0, and as usual let $G = \int g$. Assume the variation to be v, and the maximum of G' be u. Then we have $u - v \leqslant G' \leqslant u$. Assume $a < b \in I$ are points such that $G(a) = G(b) = 0$ and G is nonnegative (the nonpositive case being symmetrical), then G is bounded by the functions

$$u(x-a), \quad (v-u)(b-x).$$

These linear functions form a triangle of height h that satisfies

$$h/u + h/(v-u) = b - a \leqslant \delta,$$

and therefore $h \leqslant \delta u(v-u)/v$. An estimate for the integral of G over $[a,b]$ is therefore obtained multiplying by $(b-a)/2$, and its integral over I is therefore bounded by

$$\frac{\delta^2 u(v-u)}{2v}.$$

Deriving with respect to the parameter u as usual, we have that the maximum is attained for $u = v/2$, and is equal to $v\delta^2/8$. Since v was the variation in the interval of the partition, we proved the lemma. □

Remark 54. In the proof we just used the total variation of g in the *interiors* of the intervals $I \in \Pi$, disregarding the possible jumps at the boundary between different intervals of the partition.

B.1.3. An estimate for $\|L\|_{W(I)\to W}$. We give here an estimate for $\|L\|_{W(I)\to W}$, that is the W norm of Lf for a function (having zero average and) whose support is contained in an interval I. This estimate is required in the proof of lemma 27.

Lemma 55. *We have*

$$\|L\|_{W(I)\to W} \leqslant \|T'\|_{L^\infty(I)}.$$

Proof. Observe that for each function $h \in L^1$, and putting $H = \int h$ so that $\|h\|_W = \|H\|_{L^1}$, we have

$$\int Lh = \sum_i \int_0^x \frac{h(T_i^{-1}(t))}{T'(T_i^{-1}(t))} \mathrm{d}t = \sum_i \int_0^{T_i^{-1}(x)} h(y)\mathrm{d}y \quad (\text{via } t = T_i(y)) = \sum_i H(T_i^{-1}(x))$$

(we assume the above integral to be extended as a constant value for x outside the image T_i).

Therefore

$$\|Lh\|_W = \left\|\int Lh\right\|_{L^1} \leqslant \sum_i \int_0^1 |H(T_i^{-1}(x))|\mathrm{d}x \leqslant \sum_i \int_{T_i^{-1}([0,1])} |H(y)| \cdot T'(y)\mathrm{d}y \quad (\text{via } x = T_i(y))$$

$$\leqslant \int_0^1 |H(y)| \cdot T'(y)\mathrm{d}y \leqslant \|H\|_{L^1} \cdot \|T'\|_{L^\infty(\mathrm{Supp}(H))} \leqslant \|h\|_W \cdot \|T'\|_{L^\infty(\mathrm{Supp}(I))}.$$

□

B.2. Variation estimates

In this section we collect several variation estimates which are used in section 3.

B.2.1. The variation of Lg in an interval I. We recall here how we can estimate the variation of Lg in an interval, the estimate will be used computationally for estimating the variation of $L\tilde{f}$ in intervals of some partition. This estimate is used to compute the bound provided by lemma 25.

Lemma 56 (*Local variation inequality*). *Let $I \subseteq [0, 1]$ be an interval. Let $L_i g$ be the component of Lg coming from the ith branch, defined in (19). We have* $\mathrm{Var}_I(Lg) \leqslant \sum_i \mathrm{Var}_I(L_i g)$, *and the variation of each component can be estimated as*

$$\mathrm{Var}_I(L_i g) \leqslant \mathrm{Var}_{T_i^{-1}(I)}(g) \cdot \left\|\frac{1}{T'}\right\|_{L^\infty(T_i^{-1}(I))} + \|g\|_{L^1(T_i^{-1}(I))} \cdot \left\|\frac{T''}{T'^2}\right\|_{L^\infty(T_i^{-1}(I))}$$
$$+ \sum_{y \in \partial Dom(T_i): T(y) \in I} \left|\frac{g(y)}{T'(y)}\right|.$$

Proof. Let $g \in C^1([0, 1])$, the bounded variation case following by density of C^1 in BV:

$$(L_i g)'(x) = \frac{g'(T_i^{-1}(x))}{T'(T_i^{-1}(x))^2} - \frac{g(T_i^{-1}(x)) \cdot T''(T_i^{-1}(x))}{T'(T_i^{-1}(x))^3}.$$

And consequently the variation of g in an interval I can be bounded as

$$\int I \left|\frac{g'(T_i^{-1}(x))}{T'(T_i^{-1}(x))^2}\right| + \left|\frac{g(T_i^{-1}(x))\cdot T''(T_i^{-1}(x))}{T'(T_i^{-1}(x))^3}\right| \mathrm{d}x = \int_{T_i^{-1}(I)} \left|\frac{g'(y)}{T'(y)}\right| + \left|\frac{g(y)\cdot T''(y)}{T'^2}\right| \mathrm{d}y$$

replacing $T^{-1}(x)$ by y, and $\mathrm{d}x$ by $T'(Y)\mathrm{d}y$, as usual.

Taking into account the value of $L_i g$ at the boundary of its support we obtain the estimate. Please remark that if T_i is not full branch the support of $L_i g$ is strictly contained in $[0, 1]$ □

B.2.2. The variation of $N_\xi g$. We deduce an algorithm for estimating $\mathrm{Var}_I(N_\xi g)$ provided that we have enough information about g.

Lemma 57. *Assume $\rho_\xi = \xi^{-1}\chi_{[-\xi/2,\xi/2]}(x)$. For any interval $I = [a, b]$ we have*

$$\mathrm{Var}_I(N_\xi g) \leqslant \xi^{-1}\cdot \min\left\{\|g\|_{L^1(I-\xi/2\cup I+\xi/2)}\,\mathrm{Var}(\rho), |I|\cdot \mathrm{Var}_{[a-\xi/2,b+\xi/2]}(g)\|\rho\|_\infty\right\}.$$

Hence in the case where $\rho_\xi = \xi^{-1}\chi_{[-\xi/2,\xi/2]}$

$$\mathrm{Var}_I(N_\xi g) \leqslant \xi^{-1}\cdot \min\left\{\|g\|_{L^1(I-\xi/2\cup I+\xi/2)}, |I|\cdot \mathrm{Var}_{[a-\xi/2,b+\xi/2]}(g)\right\}.$$

Proof. For an interval of the partition I (centred in t, say), let us set

$$\overline{I} = I - \xi/2 \cup I + \xi/2,$$

we have

$$\int_I |(\rho_\xi * g)'|\mathrm{d}x \leqslant \int_I \left(\int_{\overline{I}} |\rho_\xi'(x-y)g(y)|\mathrm{d}y\right)\mathrm{d}x \leqslant \int_{\overline{I}} |g(y)|\int_I |\rho_\xi'(x-y)|\mathrm{d}x\mathrm{d}y$$
$$\leqslant \int_{\overline{I}} |g(y)|\,\mathrm{Var}(\rho_\xi)\mathrm{d}y \leqslant \xi^{-1}\,\mathrm{Var}(\rho)\cdot\left(\|g\|_{L^1(I-\xi/2)} + \|g\|_{L^1(I+\xi/2)}\right).$$

Symmetrically putting $I = [a, b]$ we have

$$\int_I |(\rho_\xi * g)'|\mathrm{d}x \leqslant \int_I \left(\int |\rho_\xi(x-y)g'(y)|\mathrm{d}y\right)\mathrm{d}x \leqslant \int_I \|\rho_\xi\|_\infty \int_{x-\xi/2}^{x+\xi/2} |g'(y)|\mathrm{d}y\mathrm{d}x$$
$$\leqslant \int_I \xi^{-1}\|\rho_\xi\|_\infty \int_{a-\xi/2}^{b+\xi/2} |g'(y)|\mathrm{d}y\mathrm{d}x \leqslant \int_I \xi^{-1}\|\rho_\xi\|_\infty\,\mathrm{d}x \cdot \int_{a-\xi/2}^{b+\xi/2} |g'(y)|\mathrm{d}y$$
$$\leqslant |I|\xi^{-1}\|\rho_\xi\|_\infty \mathrm{Var}_{[a-\xi/2,b+\xi/2]}(g).$$

□

Remark 58. If the minimum was always obtained as the first part, we end up estimating the total variation of $N_\xi g$ as $2\xi^{-1} = \mathrm{Var}(\rho_\xi)$, getting the same bound as in the *a priori* estimate. If the second part is always bigger, the estimate is approximatively $\mathrm{Var}(g)$, with a small increase due to the fact that we will be integrating the variation over an interval of size $\xi + \delta$ rather than ξ.

ORCID iDs

Stefano Galatolo https://orcid.org/0000-0003-3934-5412
Isaia Nisoli https://orcid.org/0000-0003-2474-594X

References

[1] Arnold L 2003 *Random Dynamical Systems* (Springer Monographs in Mathematics) (Berlin: Springer)
[2] Arnold L, Crauel H and Wihstutz V 1983 Stabilization of linear systems by noise *SIAM J. Control Optim.* **21** 451–61
[3] Baladi V and Viana M 1996 Strong stochastic stability and rate of mixing for unimodal maps *Annales scientifiques de l'École Normale Supérieure* **29** 483–517
[4] Bioni Liberalquino R, Monge M, Galatolo S and Marangio L 2018 Chaotic Itinerancy in Random Dynamical System Related to Associative Memory Models *Mathematics* **6** 39
[5] Bose C and Murray R 2001 The exact rate of approximation in Ulam's method *Discrete Contin. Dynam. Systems* **7** 219–35
[6] Boyarsky A and Góra P 1997 *Laws of Chaos, Invariant Measures and Dynamical Systems in One Dimension* (Basel: Birkhäuser)
[7] Braverman M, Grigo A and Rojas C 2012 Noise vs computational intractability in dynamics *ITCS '12 Proc. of the 3rd Innovations in Theoretical Computer Science Conf.* pp 128–41
[8] Capinski M and Simo C 2012 Computer assisted proof for normally hyperbolic invariant manifolds *Nonlinearity* **25** 1997–2026
[9] Dellnitz M and Junge O 2002 Set oriented numerical methods for dynamical systems *Handbook of Dynamical Systems* vol 2 (Amsterdam: Elsevier)
[10] Marangio L, Sedro J, Galatolo S, Di Garbo A and Ghil M 2019 Arnold maps with noise: differentiability and non-monotonicity of the rotation number *J. Stat. Phys.* (https://doi.org/10.1007/s10955-019-02421-1)
[11] Doi S 1989 A chaotic map with a flat segment can produce a noise-induced order *J. Stat. Phys.* **55** 941–64
[12] Galatolo S and Giulietti P 2019 Linear response for dynamical systems with additive noise *Nonlinearity* **32** 6
[13] Galatolo S and Nisoli I 2014 An elementary approach to rigorous approximation of invariant measures *SIAM J. Appl. Dyn. Syst.* **13** 958–85
[14] Galatolo S, Nisoli I and Saussol S 2015 An elementary way to rigorously estimate convergence to equilibrium and escape rates *J. Comput. Dynam.* **2** 51–64
[15] Galatolo S, Monge M and Nisoli I 2016 Rigorous approximation of stationary measures and convergence to equilibrium for iterated function systems *J. Phys. A: Math. Theor.* **49** 274001
[16] Galatolo S, Nisoli I and Rojas C 2014 Probability, statistics and computation in dynamical systems *Math. Struct. Comput. Sci.* **24** e240304
[17] Garrido P L and Gallavotti G 1994 Billiards correlation functions *J. Stat. Phys.* **76** 549–85
[18] Galias Z and Tucker W 2015 Is the Hénon attractor chaotic? *Chaos* **25** 033102
[19] Gora P and Islam S 2011 Invariant measures of stochastic perturbations of dynamical systems using Fourier approximations *Int. J. Bifurcation Chaos* **21** 113–23
[20] Guiheneuf P A 2018 Physical measures of discretizations of generic diffeomorphisms *Ergod. Theor. Dyn. Syst.* **38** 1422–58
[21] Henzel H P and Ebeling W 1985 The decay of correlations in chaotic maps *Phys. Lett.* A **111** 1–4
[22] Liverani C 2001 Rigorous numerical investigations of the statistical properties of piecewise expanding maps-A feasibility study *Nonlinearity* **14** 463–90
[23] Young L-S 2013 Mathematical theory of Lyapunov exponents *J. Phys. A: Math. Theor.* **46** 254001
[24] Matsumoto K and Tsuda I 1983 Noise-induced order *J. Stat. Phys.* **31** 87–106
[25] Matsumoto K 1984 Noise-induced order II *J. Stat. Phys.* **34** 111–27
[26] Nisoli I Sufficient conditions for Noise induced order in 1-dimensional systems (arXiv:2003.08422)
[27] Revol N and Rouillier F 2005 Motivations for an arbitrary precision interval arithmetic and the MPFI library *Reliab. Comput.* **11** 275–90

[28] SageMath 2016 *The Sage Mathematics Software System* (*Version 7.1*) (The Sage Developers) http://www.sagemath.org
[29] Stein E M and Shakarchi R 2009 *Real Analysis: Measure Theory, Integration, and Hilbert Spaces* (Princeton, NJ: Princeton University Press)
[30] Stone E J, Gohara D and Shi G 2010 OpenCL: a parallel programming standard for heterogeneous computing systems *Comput. Sci. Eng.* **12** 66–73
[31] Tucker W 2011 *Auto-Validating Numerical Methods* (Frontiers in Mathematics) (Princeton, NJ: Princeton University Press)
[32] Viana M 2014 *Lectures on Lyapunov Exponents* (Cambridge Studies in Advanced Mathematics 145) (Cambridge: Cambridge University Press)
[33] Wackerbauer R 1999 When noise decreases deterministic diffusion *Phys. Rev.* E **59** 2872
[34] Yoshimoto M, Shirahama H and Kurosawa S 2008 Noise-induced order in the chaos of the Belousov–Zhabotinsky reaction *J. Chem. Phys.* **129** 014508
[35] Yamazaki H, Yamada T and Kai S 1998 Can stochastic resonance lead to order in chaos? *Phys. Rev. Lett.* **81** 4112
[36] Yoshimoto M, Yamaguchi T, Sato H, Kato K, Mahara H and Nagashima H 1998 Effect of Noise on the low flow rate chaos in the Belousov–Zhabotinsky reaction *J. Phys. Soc. Jpn.* **67** 103–11
[37] Zhabotisnky Z 1973 Autowave processes in a distributed chemical system *J. Theor. Biol.* **40** 45–61
[38] Zhang D, Györgyi L and Peltier W 1993 Deterministic chaos in the Belousov–Zhabotinsky reaction: experiments and simulations *Chaos* **3** 723