Probab. Theory Relat. Fields (2015) 162:233–274
DOI 10.1007/s00440-014-0571-y

Annealed and quenched limit theorems for random expanding dynamical systems

Romain Aimino · Matthew Nicol · Sandro Vaienti

Received: 16 October 2013 / Revised: 11 June 2014 / Published online: 1 July 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract In this paper, we investigate annealed and quenched limit theorems for random expanding dynamical systems. Making use of functional analytic techniques and more probabilistic arguments with martingales, we prove annealed versions of a central limit theorem, a large deviation principle, a local limit theorem, and an almost sure invariance principle. We also discuss the quenched central limit theorem, dynamical Borel–Cantelli lemmas, Erdös–Rényi laws and concentration inequalities.

Keywords Random dynamical systems · Limit theorems · Borel–Cantelli lemmas · Erdös–Rényi laws · Concentration inequalities

Mathematics Subject Classification (2010) 37H99 · 37A25 · 37D20

1 Introduction

1.1 Limit theorems for random dynamical systems: a brief survey

Statistical properties for deterministic uniformly expanding dynamical systems are by now pretty well understood, starting from the existence of an absolutely continuous

R. Aimino (✉) · S. Vaienti
Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France
e-mail: aimino@cpt.univ-mrs.fr

S. Vaienti
e-mail: vaienti@cpt.univ-mrs.fr

R. Aimino · S. Vaienti
Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France

M. Nicol
Department of Mathematics, University of Houston, Houston, TX, USA
e-mail: nicol@math.uh.edu

invariant probability [1,2], to exponential decay of correlations and limit theorems [3–5], and more refined properties, such as Erdös–Rényi laws [6,7], dynamical Borel–Cantelli lemmas [8,9] and concentration inequalities [10,11]. Most of these results are derived from the existence of a spectral gap of the transfer operator of the system, when acting on a appropriately chosen Banach space. The books [12,13] contain a nice overview and historical perspectives on the subject.

For random dynamical systems, the understanding of the situation is still unsatisfactory. A random dynamical system can be seen as a random composition of maps acting on the same space X, where maps are chosen according to a stationary process. When the process is a sequence of independent random maps, this gives rise to a Markov chain with state space X. The iid setting has been extensively studied in the book of Kifer [14], while the general case is treated in the book of Arnold [15]. The relevance of random dynamical systems is obvious from the fact that in most physical applications, it is very unlikely that the same map is iterated all along the time, and it is rather the case that different maps, very close to a fixed one, are iterated randomly. This topic of stochastic perturbations is very well covered in [16]. Random dynamical systems also arise naturally in the field of particle systems on lattices, where on a single site, the particle is subject to a local deterministic dynamic, but can jump from one site to another randomly, see [17,18].

The existence of stationary measures absolutely continuous with respect to Lebesgue measure was first studied by Pelikan [19] and Morita [20] in the case where X is the unit interval, and in the thesis of Hsieh [21] for the multidimensional case. This question has also been investigated for one-dimensional and multidimensional systems in the case of position dependent probabilities, see [22] and references therein. For limit theorems, the available literature is much more sparse. It should be first stressed that for random systems, limit theorems are of two kind: annealed results concern properties related to the skew-product dynamics, while quenched results describe properties for almost every realization. Annealed results follow from the spectral analysis of an annealed transfer operator, generalizing the successful approach for deterministic systems. In this line of spirit, we can cite the papers of Baladi [23], Baladi and Young [24] or Ishitani [25] and the thesis [18]. Quenched results are usually more difficult to prove. Exponential decay of correlation in a quenched regime has been proved using Birkhoff cones technique in [26–28], while a quenched central limit theorem and a law of iterated logarithm are studied by Kifer [29], using a martingale approximation. These results deal with more general stationary process, where absolutely continuous stationary measure can fail to exist and are replaced by a family of sample measures. Closer to our setting are the papers [30,31] which are concerned with random toral automorphisms, and a very recent work of M. Stenlund and H. Sulky (A coupling approach to random circle maps expanding on the average, preprint, 2013), where quenched exponential decay of correlations together with an annealed almost sure invariance principle are shown for iid expanding circle maps, using the coupling method.

1.2 Limit theorems: our new results

When the random dynamical system is contracting on average, the transition operator of the Markov chain admits a spectral gap on a space of Hölder functions, from which

one can deduce a large span of limit theorems following the Nagaëv's method, see for instance [32] and references therein. Nevertheless, for the applications we have in mind, the maps will instead be expanding on average. In this situation, the transition operator generally fails to admit a spectral gap and we will preferably rely on the quasi-compactness of an associated annealed transfer operator on an appropriate Banach space. In this paper, we provide an abstract functional framework, valid for several one dimensional and multidimensional systems, under which annealed limit theorems hold for smooth enough observables. More precisely, under a spectral gap assumption for the annealed transfer operator, we apply Nagaëv's perturbative method to obtain a central limit theorem with rate of convergence and a large deviation principle. A Borel–Cantelli argument allows us to derive immediately a quenched upper bound for the large deviation principle, but the question of whether a quenched lower bound holds remains open. We also show a local limit theorem under an abstract aperiodicity condition, and relate in most practical cases this condition to the usual one for individual maps. We apply Gouëzel's spectral method to prove an annealed almost sure invariance principle for vector valued observables: this is a strong reinforcement of the central limit theorem which has many consequences, such as the law of the iterated logarithm, the functional central limit theorem, and the almost sure central limit theorem [33].

Changing slightly our approach, we then adapt the martingale approximation method, which goes back to Gordin [34], and give an alternative proof of the annealed central limit theorem. This requires the introduction of a symbolic deterministic system on which the standard martingale procedure can be pursued. Decay of annealed correlations is the key ingredient here and allows us to show that the Birkhoff's sums can be written as the sum of a backwards martingale and a coboundary, from which the central limit theorem follows from the analogous result for martingales.

We next investigate dynamical Borel–Cantelli lemmas: if (f_n) is a bounded sequences of positive functions lying in the functional space, such as $\sum_n \int f_n d\mu = \infty$, where μ is the stationary measure, we prove that

$$\frac{\sum_{k=0}^{n-1} f_k(T_{\underline{\omega}}^k x)}{\sum_{k=0}^{n-1} \int f_k d\mu} \to 1,$$

for almost every realization $\underline{\omega}$ and almost every point $x \in X$, a property usually called strong Borel–Cantelli lemma in the literature. Of particular interest is the case where f_n are the characteristic functions of a sequence of decreasing sets, since this relates to recurrence properties of the system. The proof builds upon annealed decay of correlations, and is a consequence of the work of Kim [9]. This result can be seen as a generalization of the strong law of large numbers, and it is hence natural to study the nature of the fluctuations in this convergence. Provided we have precise enough estimates on the measure of the sets, we prove a central limit theorem. For this purpose, we employ the martingale technique already used before for Birkhoff sums, and make use of a central limit theorem for non stationary martingales from Hall and Heyde [35], mimicking the proof from [36] for the deterministic case.

We then turn to Erdös–Rényi laws: these limit laws give information on the maximal average gain in the case where the length of the time window ensures there is a non-

degenerate limit. This result was first formulated by Erdös and Rényi [37], and brought in dynamical context in [6,7,38] among others. Making use of the large deviation principle, we adapt the proof of [7] to show that an annealed Erdös–Rényi law holds true in the random situation, for one-dimensional transformations.

Importing a technique from the field of random walks in random environments, Ayyer et al. [31] proved a quenched central limit theorem for random toral hyperbolic automorphisms. Their approach consists in proving a spectral gap for the original system and for a "doubled" system acting on X^2, where maps are given by $\hat{T}_\omega(x, y) = (T_\omega x, T_\omega y)$, and driven by the same iid process. This allows to prove a quenched central limit theorem for subsequences of the Birkhoff sums by a Borel–Cantelli argument, and the large deviation principle helps to estimate the error occurring in the gaps. Unfortunately, this method needs a precise relation between the asymptotic variance of the observable on the original system, and the asymptotic variance of a deduced observable on the doubled system. This relation is easily shown when all maps preserve the same measure, as it is the case in [31], but is harder to prove, and possibly false, in full generality. Hence, in this paper, we restrict our attention to the case where all maps preserve the Lebesgue measure on the unit interval. Apart the trivial case where all maps are piecewise onto and linear, we show that we can include in the random compositions a class of maps introduced in [39], which have a neutral fixed point and a point where the derivative blows up. The general case remains open.

Concentration inequalities are a well known subject in probability theory, and have numerous and deep consequences in statistics. They concern deviations from the mean for non additive functionals and hence generalize large deviations estimates for ergodic sums. Furthermore, these inequalities are non-asymptotic. The price to pay is that they do not give precise asymptotics for the deviation function, in contrast to the large deviations principle. They were introduced in dynamical systems by Collet, Martinez and Schmitt [11] who prove an exponential inequality for uniformly piecewise expanding maps of the interval. The paper [10] covers a wide range of uniformly and non-uniformly expanding/hyperbolic dynamical systems which are modeled by a Young tower. For random dynamical systems, concentration inequalities were not previously studied. As far as the authors know, the only result available is [40], which covers the case of the observational noise. We attempt to fill this gap and prove an annealed exponential concentration inequality for randomly expanding systems on the interval, generalizing the approach of [11]. We then give an application to the rate of convergence of the empirical measure to the stationary measure.

1.3 Plan of the paper

The paper is outlined as follows. In Sect. 2, we described our abstract functional framework, and give several classes of one dimensional and multidimensional examples which fit the assumptions. In Sect. 3 we apply Nagaëv method to prove annealed limit theorems. In Sect. 4, we explain how the central limit theorem follows from a martingale approximation. In Sect. 5, we prove dynamical Borel–Cantelli lemmas and a central limit theorem for the shrinking target problem. In Sect. 6, we prove an Erdös–Rényi law for random one-dimensional systems. In Sect. 7, we consider the quenched

central limit theorem for specific one dimensional random systems. Finally, in Sect. 8, we prove an exponential concentration inequality and discuss its applications.

The letter C denotes a positive constant whose precise value has no particular importance and can change from one line to another.

2 Abstract framework and examples

Let $(\tilde{\Omega}, \tilde{\mathcal{T}}, \tilde{\mathbb{P}})$ be a probability space, and $\theta : \tilde{\Omega} \to \tilde{\Omega}$ be a measure preserving transformation. Let now $(X, \mathcal{A})$ be a measurable space. Suppose that to each $\underline{\omega} \in \tilde{\Omega}$ is associated a transformation $T_{\underline{\omega}} : X \to X$ such that the map $(\underline{\omega}, x) \mapsto T_{\underline{\omega}}(x)$ is measurable. We are then considering random orbits $T_{\theta^n \underline{\omega}} \circ \cdots \circ T_{\underline{\omega}} x$.

One can now define a skew-product transformation $F : \tilde{\Omega} \times X \to \tilde{\Omega} \times X$ by $F(\underline{\omega}, x) = (\theta \underline{\omega}, T_{\underline{\omega}} x)$. We will say that a probability measure μ on $(X, \mathcal{A})$ is a stationary measure if $\tilde{\mathbb{P}} \otimes \mu$ is invariant under F.

The simplest situation possible is the i.i.d. case: $(\tilde{\Omega}, \tilde{\mathcal{T}}, \tilde{\mathbb{P}})$ is a countable product space, namely $\tilde{\Omega} = \Omega^{\mathbb{N}}, \tilde{\mathcal{T}} = \mathcal{T}^{\otimes \mathbb{N}}, \tilde{\mathbb{P}} = \mathbb{P}^{\otimes \mathbb{N}}$ and θ is the full shift. If to each $\omega \in \Omega$ is associated a map T_ω on X such that $(\omega, x) \mapsto T_\omega(x)$ is measurable, then we define $T_{\underline{\omega}} = T_{\omega_1}$ for each $\underline{\omega} \in \tilde{\Omega}$, with $\underline{\omega} = (\omega_1, \omega_2, \ldots)$. This fits the framework described previously. It is easily seen that μ is a stationary measure iff $\mu(A) = \int_\Omega \mu(T_\omega^{-1}(A))\, d\mathbb{P}(\omega)$ for each $A \in \mathcal{A}$. Moreover, if we set $X_n(\underline{\omega}, x) = T_{\theta^n \underline{\omega}} \circ \cdots \circ T_{\underline{\omega}} x$, this defines a homogeneous Markov chain with state space $(X, \mathcal{A})$ and transition operator given by $U(x, A) = \mathbb{P}(\{\omega : T_\omega x \in A\})$ for any $x \in X$ and any set $A \in \mathcal{A}$.

From now, we will always consider this i.i.d. situation. Suppose now that $(X, \mathcal{A})$ is endowed with a probability measure m such that each transformation T_ω is non-singular w.r.t. m. We will investigate existence and statistical properties of stationary measures absolutely continuous w.r.t. m. To this end, we introduce averaged transfer and Koopman operators.

Since every transformation T_ω is non-singular, the transfer operator P_ω and the Koopman operator U_ω of T_ω are well defined, and act respectively on $L^1(m)$ and $L^\infty(m)$. We recall their definitions for the convenience of the reader, and refer to [12,13] for more properties. The Koopman operator of T_ω is acting on $L^\infty(m)$ by $U_\omega f = f \circ T_\omega$. Its action on conjugacy classes of functions is well defined since T_ω is non-singular w.r.t. m. The transfer operator, or Perron–Frobenius operator is acting on $L^1(m)$ in the following way : for $f \in L^1(m)$, define the complex measure m_f by $dm_f = f dm$. Then $P_\omega f$ is defined to be the Radon–Nykodim derivative of the push-forward measure $T_\star m_f$ w.r.t. m, which is well defined by non-singularity. The main relation between these two operators is given by the duality formula $\int_X P_\omega f(x) g(x) dm(x) = \int_X f(x) U_\omega g(x) dm(x)$, which holds for all $f \in L^1(m)$ and $g \in L^\infty(m)$.

We can now defined the averaged versions of these operators. For $f \in L^1(m)$, we define Pf by the formula $Pf(x) = \int_\Omega P_\omega f(x)\, d\mathbb{P}(\omega)$, and for $g \in L^\infty(m)$, we define Ug by $Ug(x) = \int_\Omega U_\omega g(x)\, d\mathbb{P}(\omega)$. The operator U just defined coincides with the transition operator U of the Markov chain (X_n), when acting on functions. Notice that for all $n \geq 0$ and $g \in L^\infty(m)$, one has

$U^n g(x) = \int_{\Omega^n} g(T_{\omega_n} \dots T_{\omega_1} x)\, d\mathbb{P}^{\otimes n}(\omega_1, \dots, \omega_n) = \int_{\tilde{\Omega}} g(T_{\omega_n} \dots T_{\omega_1} x)\, d\tilde{\mathbb{P}}(\underline{\omega})$, because $\tilde{\mathbb{P}}$ is a product measure. It is then straightforward to check that U is the dual operator of P, that is $\int_X Pf(x)g(x)\, dm(x) = \int_X f(x)Ug(x)\, dm(x)$ for all $f \in L^1(m)$ and $g \in L^\infty(m)$. An absolutely continuous probability measure is stationary iff its density is a fixed point of P.

We will assume that P has good spectral properties on some Banach space of functions. More precisely, we assume that there exists a Banach space $(\mathcal{B}, \|.\|)$ such that:

1. $\mathcal{B}$ is compactly embedded in $L^1(m)$;
2. Constant functions lie in $\mathcal{B}$;
3. $\mathcal{B}$ is a complex Banach lattice: for all $f \in \mathcal{B}$, $|f|$ and $\bar{f}$ belong to $\mathcal{B}$;
4. $\mathcal{B}$ is stable under P: $P(\mathcal{B}) \subset \mathcal{B}$, and P acts continuously on $\mathcal{B}$;
5. P satisfies a Lasota–Yorke inequality: there exist $N \geq 1$, $\rho < 1$ and $K \geq 0$ such that $\|P^N f\| \leq \rho \|f\| + K \|f\|_{L^1_m}$ for all $f \in \mathcal{B}$.

The LY inequality implies in particular that the spectral radius of P acting on $\mathcal{B}$ is less or equal than 1, and since m belongs to the topological dual of $\mathcal{B}$ by the first assertion, and is fixed by $P^\star$ the adjoint of P, we have that the spectral radius is 1. Hence, by Ionescu-Tulcea and Marinescu's theorem [41] (see also [42]), we have that the essential spectral radius of P is less or equal than $\rho < 1$, implying that P is quasi-compact on $\mathcal{B}$, since it has spectral radius 1. A standard argument using compactness proves that P^N and hence P has a positive fixed point: there is an element $h \in \mathcal{B}$ with $Ph = h, h \geq 0$ and $\int_X h\, dm = 1$. As a consequence 1 is an eigenvalue of P. Another consequence of quasi-compactness is the fact that the spectrum of P is constituted of a finite set of eigenvalues with modulus 1 of finite multiplicity and the remaining spectrum is contained in a disk of radius strictly less than 1. We will make the following assumption, that prevents the possibility of peripheral spectrum:

6. 1 is a simple (isolated) eigenvalue of P, and there is no other eigenvalue on the unit circle.

This assumption implies in particular that the absolutely continuous stationary measure is unique. We will denote it by μ, and its density by h, throughout the paper.

Usually, assertions 4 and 5 can be deduced from corresponding assertions for the operators P_ω if the constants appearing in the Lasota–Yorke inequality are uniform. Nevertheless, they can be established even if one of the maps T_ω is not uniformly expanding, as showed by the following class of examples:

Example 2.1 [Piecewise expanding one-dimensional maps] A Lasota–Yorke map is a piecewise C^2 map $T : [0, 1] \to [0, 1]$ for which $\lambda(T) := \inf |T'| > 0$.

We denote by P_T the transfer operator (with respect to Lebesgue measure) associated to T. One has

$$P_T f(x) = \sum_{Ty=x} \frac{f(y)}{|T'(y)|}$$

for all $f \in L^1(m)$. We will analyze the spectral properties of P_T acting on the space of functions of bounded variation. We recall the definition. A function $f : [0, 1] \to \mathbb{C}$

is of bounded variation if its total variation defined as

$$\mathrm{Var}(f) = \sup \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)|,$$

where the supremum is taken over all the finite partitions $0 = x_0 < \cdots < x_n = 1$, is finite.

For a class of equivalence $f \in L^1(m)$, we then define

$$\mathrm{Var}(f) = \inf\{\mathrm{Var}(g) \,/\, f = g \; m\text{−ae}\}.$$

The space $\mathrm{BV} = \{f \in L^1(m) \,/\, \mathrm{Var}(f) < \infty\}$ is endowed with the norm $\|f\| = \|f\|_{L^1_m} + \mathrm{Var}(f)$, which turns it into a Banach space satisfying assumptions 1, 2 and 3 above. Furthermore, this is a Banach algebra which embeds continuously into $L^\infty(m)$.

If T is a Lasota–Yorke map, the following inequality holds:

Proposition 2.2 (Lasota–Yorke inequality [1]) *For any* $f \in \mathrm{BV}$*, we have*

$$\mathrm{Var}(P_T f) \leq \frac{2}{\lambda(T)} \mathrm{Var}(f) + A(T)\|f\|_{L^1_m}$$

where $A(T)$ *is a finite constant depending only on* T.

Let Ω be a finite set, together with a probability vector $\mathbb{P} = \{p_\omega\}_{\omega\in\Omega}$ and a finite number of Lasota–Yorke maps $T = \{T_\omega\}_{\omega\in\Omega}$. We assume that $p_\omega > 0$ for all $\omega \in \Omega$. The system $(\Omega, \mathbb{P}, T)$ is called a random Lasota–Yorke system.

The random Lasota–Yorke system $(\Omega, \mathbb{P}, T)$ is expanding in mean if

$$\Lambda := \sum_{\omega\in\Omega} \frac{p_\omega}{\lambda(T_\omega)} < 1.$$

The annealed transfer operator associated to $(\Omega, \mathbb{P}, T)$ is $P = \sum_{\omega\in\Omega} p_\omega P_{T_\omega}$. It satisfies $P^n = \sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} P_{T^n_{\underline{\omega}}}$, where $\underline{\omega} = (\omega_1, \ldots, \omega_n)$, $p^n_{\underline{\omega}} = p_{\omega_1} \ldots p_{\omega_n}$ and $T^n_{\underline{\omega}} = T_{\omega_n} \circ \cdots \circ T_{\omega_1}$.

Proposition 2.3 *If* $(\Omega, \mathbb{P}, T)$ *is expanding in mean, then some iterate of the annealed transfer operator satisfies a Lasota–Yorke inequality on* BV.

Proof By the classical LY inequality and subadditivity of the total variation, one has

$$\mathrm{Var}(P^n f) \leq 2\theta_n \mathrm{Var}(f) + A_n \|f\|_{L^1_m}$$

for all $n \geq 1$ and all $f \in \mathrm{BV}$, where $\theta_n = \sum_{\underline{\omega}\in\Omega^n} \frac{p^n_{\underline{\omega}}}{\lambda(T^n_{\underline{\omega}})}$ and $A_n = \sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} A(T^n_{\underline{\omega}})$. Since $\lambda(T^n_{\underline{\omega}}) \geq \lambda(T_{\omega_1}) \ldots \lambda(T_{\omega_n})$, one obtains $\theta_n \leq \sum_{\underline{\omega}\in\Omega^n} \frac{p_{\omega_1} \cdots p_{\omega_n}}{\lambda(T_{\omega_1}) \ldots \lambda(T_{\omega_n})} = \Lambda^n$. Hence, $2\theta_n < 1$ for n large enough, while the corresponding A_n is finite. This concludes the proof. □

This implies assumptions 4 and 5.

Remark 2.4 1. Pelikan [19] showed that the previous Lasota–Yorke inequality still holds under the weaker assumption $\sup_x \sum_{\omega\in\Omega} \frac{p_\omega}{|T'_\omega(x)|} < 1$.
2. The result is still valid if the set Ω is infinite, assuming an integrability condition for the distortion. See Remark 5.1 in Morita [20].

From Ionescu-Tulcea and Marinescu theorem, it follows that the annealed transfer operator has the following spectral decomposition:

$$P = \sum_i \lambda_i \Pi_i + Q,$$

where all λ_i are eigenvalues of P of modulus 1, Π_i are finite-rank projectors onto the associated eigenspaces, Q is a bounded operator with a spectral radius strictly less than 1. They satisfy

$$\Pi_i \Pi_j = \delta_{ij} \Pi_i, \;\; Q\Pi_i = \Pi_i Q = 0.$$

This implies existence of an absolutely continuous stationary measure, with density belonging to BV. Standard techniques show that 1 is an eigenvalue and that the peripheral spectrum is completely cyclic. We'll give a concrete criterion ensuring that 1 is a simple eigenvalue of P, and that there is no other peripheral eigenvalue, hence implying assumption 6. In this case, we will say that $(\Omega, \mathbb{P}, T)$ is mixing.

Definition 2.5 The random LY system $(\Omega, \mathbb{P}, T)$ is said to have the Random Covering (RC) property if for any non-trivial subinterval $I \subset [0, 1]$, there exist $n \geq 1$ and $\underline{\omega} \in \Omega^n$ such that $T^n_{\underline{\omega}}(I) = [0, 1]$.

Proposition 2.6 *If $(\Omega, \mathbb{P}, T)$ is expanding in mean and has the (RC) property, then $(\Omega, \mathbb{P}, T)$ is mixing and the density of the unique a.c. stationary measure is bounded away from 0.*

Proof Since the peripheral spectrum of P consists of a finite union of finite cyclic groups, there exists $k \geq 1$ such that 1 is the unique peripheral eigenvalue of P^k. It suffices then to show that the corresponding eigenspace is one-dimensional. Standard arguments show there exists a basis of positive eigenvectors for this subspace, with disjoint supports. Let then $h \in \mathrm{BV}$ non-zero satisfying $h \geq 0$ and $P^k h = h$. There exist a non-trivial interval $I \subset [0, 1]$ and $\alpha > 0$ such that $h \geq \alpha \mathbb{1}_I$. Choose $n \geq 1$ and $\underline{\omega}^\star \in \Omega^{nk}$ such that $T^{nk}_{\underline{\omega}^\star}(I) = [0, 1]$. For all $x \in [0, 1]$, we have

$$\begin{aligned} h(x) = P^{nk} h(x) &\geq \alpha P^{nk} \mathbb{1}_I(x) = \alpha \sum_{\underline{\omega}\in\Omega^{nk}} p^{nk}_{\underline{\omega}} \sum_{T^{nk}_{\underline{\omega}} y = x} \frac{\mathbb{1}_I(y)}{|(T^{nk}_{\underline{\omega}})'(y)|} \\ &\geq \alpha p^{nk}_{\underline{\omega}^\star} \sum_{T^{nk}_{\underline{\omega}^\star} y = x} \frac{\mathbb{1}_I(y)}{|(T^{nk}_{\underline{\omega}^\star})'(y)|}. \end{aligned}$$

This shows that $h(x) \geq \alpha \frac{p^{nk}_{\underline{\omega}^\star}}{\|(T^{nk}_{\underline{\omega}^\star})'\|_{\sup}} > 0$, since there is always a $y \in I$ with $T^{nk}_{\underline{\omega}^\star} y = x$. This implies that h has full support, and concludes the proof. □

Some statistical properties of random one-dimensional systems, using the space BV, were studied in the thesis of Tümel [18].

Example 2.7 [Piecewise expanding multidimensional maps] We describe a class of piecewise expanding multi-dimensional maps introduced by Saussol [2]. Denote by m_d the d-dimensional Lebesgue measure, by $d(.,.)$ the euclidean distance and by γ_d the m_d-volume of the unit ball of $\mathbb{R}^d$. Let M a compact regular subset of $\mathbb{R}^d$ and let $T : M \to M$ be a map such that there exists a finite family of disjoint open sets $U_i \subset M$ and V_i with $\bar{U}_i \subset V_i$ and maps $T_i : V_i \to \mathbb{R}^d$ satisfying for some $0 < \alpha \leq 1$ and some small enough $\epsilon_0 > 0$:

1. $m_d(M \setminus \cup_i U_i) = 0$;
2. for all i, the restriction to U_i of T and T_i coincide, and $B_{\epsilon_0}(TU_i) \subset T_i(V_i)$;
3. for all i, T_i is a C^1-diffeomorphism from V_i onto $T_i V_i$, and for all $x, y \in V_i$ with $d(T_i x, T_i y) \leq \epsilon_0$, we have
$$|\det DT_i(x) - \det DT_i(y)| \leq c|\det DT_i(x)| d(T_i x, T_i y)^\alpha,$$
for some constant $c > 0$ independent of i, x and y;
4. there exists $s < 1$ such that for all $x, y \in V_i$ with $d(T_i x, T_i y) \leq \epsilon_0$, we have $d(x, y) \leq s d(T_i x, T_i y)$;
5. Assume that the boundaries of the U_i are included in piecewise C^1 codimension one embedded compact submanifolds. Define
$$Y = \sup_x \sum_i \sharp\{\text{smooth pieces intersecting } \partial U_i \text{ containing } x\},$$
and
$$\eta_0 = s^\alpha + \frac{4s}{1-s} Y \frac{\gamma_{d-1}}{\gamma_d}.$$
Then $\eta_0 < 1$.

The above conditions can be weakened in order to allow infinitely many domains of injectivity and also the possibility of fractal boundaries. We refer the interested reader to [2] for more details.

We will call such maps piecewise expanding maps in the sense of Saussol. The analysis of the transfer operators of these maps requires the introduction of the functional space called Quasi-Hölder. This space was first defined and studied by Keller [43] for one-dimensional transformations, and then extended to multidimensional systems by Saussol [2].

We give the definition of this space. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a measurable function. For a Borel subset $A \subset \mathbb{R}^d$, define $\operatorname{osc}(f, A) = \operatorname{ess\,sup}_{x,y \in A} |f(x) - f(y)|$. For any

$\epsilon > 0$, the map $x \mapsto \operatorname{osc}(f, B_\epsilon(x))$ is a positive lower semi-continuous function, so that the integral $\int_{\mathbb{R}^d} \operatorname{osc}(f, B_\epsilon(x))dx$ makes sense. For $f \in L^1(\mathbb{R}^d)$ and $0 < \alpha \leq 1$, define

$$|f|_\alpha = \sup_{0<\epsilon\leq\epsilon_0} \frac{1}{\epsilon^\alpha} \int_{\mathbb{R}^d} \operatorname{osc}(f, B_\epsilon(x))dx.$$

For a regular compact subset $M \subset \mathbb{R}^d$, define

$$V_\alpha(M) = \{f \in L^1(\mathbb{R}^d) \,/\, \operatorname{supp}(f) \subset M,\ |f|_\alpha < \infty\},$$

endowed with the norm $\|f\|_\alpha = \|f\|_{L^1_m} + |f|_\alpha$, where m is the Lebesgue measure normalized so that $m(M) = 1$. Note that while the norm depends on ϵ_0, the space V_α does not, and two choices of ϵ_0 give rise to two equivalent norms.

If T is a piecewise expanding map in the sense of Saussol and P_T is the transfer operator of T, a Lasota–Yorke type inequality holds:

Proposition 2.8 ([2, Lemma 4.1]) *Provided ϵ_0 is small enough, there exists $\eta < 1$ and $D < \infty$ such that for any $f \in V_\alpha$,*

$$|P_T f|_\alpha \leq \eta |f|_\alpha + D\|f\|_{L^1_m}.$$

Suppose now that Ω is a finite set, $\mathbb{P} = \{p_\omega\}_{\omega\in\Omega}$ a probability vector, and $\{T_\omega\}_{\omega\in\Omega}$ a finite collection of piecewise expanding maps on $M \subset \mathbb{R}^d$. This will be referred to as a random piecewise expanding multidimensional system. Take ϵ_0 small enough so that the inequalities $|P_{T_\omega}|_\alpha \leq \eta_\omega |f|_\alpha + D_\omega \|f\|_{L^1_m}$ for all $f \in V_\alpha$. Defining $\eta = \max \eta_\omega$ and $D = \max D_\omega$, so that $\eta < 1$ and $D < \infty$. Since $P = \sum_\omega P_{T_\omega}$, we immediately get $|Pf|_\alpha \leq \eta |f|_\alpha + \|f\|_{L^1_m}$, for all $f \in V_\alpha$. This shows that our abstract assumptions 1–5 are all satisfied. To prove that P is mixing, and hence check assumption 6, we can proceed as in the one-dimensional situation, introducing the same notion of random-covering.. Indeed, any positive non-zero element $h \in V_\alpha$ is bounded uniformly away from zero on some ball by Lemma 3.1 in [2], so we can mimic the proof of Proposition 2.6 and get:

Proposition 2.9 *If $(\Omega, \mathbb{P}, T)$ is a random piecewise expanding multidimensional system which has the random covering property in the sense that for all ball $B \subset M$, there exists $n \geq 1$ and $\underline{\omega} \in \Omega^n$ such that $T^n_{\underline{\omega}}(B) = M$, then $(\Omega, \mathbb{P}, T)$ is mixing and the density of the unique a.c. stationary measure is bounded away from* 0.

There are alternative functional spaces to study multidimensional expanding maps. One of them is the space of functions of bounded variations in higher dimension. Applications of this space to dynamical systems have been widely studied, see [44,45] among many others. We also mention the thesis of Hsieh [21] who investigates the application of this space to random multidimensional maps, using the same setting as us. Notice that maps studied there are the so-called Jabłoński maps, for which the

dynamical partition is made of rectangles. This kind of maps will appear later in this paper, when we will derive a quenched CLT. Nevertheless, the space BV in higher dimensions presents some drawbacks: it is not included in L^∞ and there exist some positive functions which are not bounded below on any ball, making the application of random covering difficult, in contrast to the the Quasi-Hölder space. Apart multi-dimensional BV, another possibility is to use fractional Sobolev spaces, as done in a deterministic setting by Thomine [46].

Example 2.10 [Random expanding piecewise linear maps] Building on a work by Tsujii [47], we considerer random compositions of piecewise linear maps. First recall a definition:

Definition 2.11 Let U be a bounded polyhedron in $\mathbb{R}^d$ with non-empty interior. An expanding piecewise linear map on U is a combination $(\mathcal{T}, \mathcal{U})$ of a map $\mathcal{T} : U \to U$ and a family $\mathcal{U} = \{U_k\}_{k=1}^l$ of polyhedra $U_k \subset U$, $k = 1, \ldots, l$, satisfying the conditions

1. the interiors of polyhedra U_k are mutually disjoint,
2. $\cup_{k=1}^l U_k = U$,
3. the restriction of the map $\mathcal{T}$ to the interior of each U_k is an affine map and
4. there exists a constant $\rho > 1$ such that $\|D\mathcal{T}_x(v)\| \geq \rho\|v\|$ for all $x \in \cup_{k=1}^l \mathrm{int}(U_k)$, and all $v \in \mathbb{R}^d$.

We will drop $\mathcal{U}$, writing merely $\mathcal{T}$, when the partition $\mathcal{U}$ is understood. A basic consequence of Tsujii [47], using the Quasi-Hölder space, is the following:

Proposition 2.12 *For any expanding piecewise linear map $\mathcal{T}$ on U, there exists constant $\epsilon_0 > 0, \theta < 1, C, K > 0$ such that, for any $n \geq 0$ and $f \in V_1$:*

$$|P_{\mathcal{T}}^n f|_1 \leq C\theta^n |f|_1 + K\|f\|_{L_m^1},$$

where $P_{\mathcal{T}}$ is the transfer operator of $\mathcal{T}$.

Let $T = \{(\mathcal{T}_\omega, \mathcal{U}_\omega)\}_{\omega\in\Omega}$ be a finite collection of expanding piecewise linear map on U, and $\mathbb{P} = \{p_\omega\}_{\omega\in\Omega}$ a probability vector.

Choosing $\epsilon_0 > 0, \theta < 1, C, D < \infty$ adequately, we get $|P_{\mathcal{T}}^n f|_1 \leq C\theta^n |f|_1 + K\|f\|_{L_m^1}$ for all $f \in V_1$, where P is the annealed transfer operator. If we assume furthermore that the system has the random covering property, then the Proposition 2.9 also holds true.

3 Spectral results

We assume here that there exists a Banach space $\mathcal{B}_0 \subset L^1(m)$, with norm $\|.\|_0$, and a constant $C > 0$ such that $\|fg\| \leq C\|f\|_0\|g\|$ for all $f \in \mathcal{B}_0$ and $g \in \mathcal{B}$. This assumption is clearly satisfied with $\mathcal{B} = \mathcal{B}_0$ if $\mathcal{B}$ is a Banach algebra, as it is the case for the space of functions of bounded variation in one dimension, or the Quasi-Hölder space. If $\mathcal{B}$ is the space of functions of bounded variation in $\mathbb{R}^d$, then we can take

$\mathcal{B}_0 = \mathrm{Lip}$, see lemma 6.4 in [46]. Elements of $\mathcal{B}_0$ will play the role of observables in the following.

The spectral decomposition of P yields $P = \Pi + Q$ where Π is the projection given by $\Pi f = (\int_X f\,dm)h$ and Q has spectral radius on $\mathcal{B}$ strictly less than 1 and satisfies $\Pi Q = Q\Pi = 0$. It follows that $P^n = \Pi + Q^n$, where $\|Q^n\| \leq C\lambda^n$, for some $C \geq 0$ and $\lambda < 1$. This implies exponential decay of correlations:

Proposition 3.1 *We have:*

1. *For all $f \in \mathcal{B}_0$ and $g \in L^\infty(m)$,*

$$\left| \int_X f\, U^n g\, d\mu - \int_X f\, d\mu \int_X g\, d\mu \right| \leq C\lambda^n \|f\|_0 \|g\|_{L^\infty_m},$$

2. *If $\mathcal{B}$ is continuously embedded in $L^\infty(m)$, then for all $f \in \mathcal{B}_0$ and $g \in L^1(m)$,*

$$\left| \int_X f\, U^n g\, d\mu - \int_X f\, d\mu \int_X g\, d\mu \right| \leq C\lambda^n \|f\|_0 \|g\|_{L^1_m},$$

3. *If $\mathcal{B}$ is continuously embedded in $L^\infty(m)$ and if the density of μ is bounded uniformly away from 0, then for all $f \in \mathcal{B}_0$ and $g \in L^1(\mu)$*

$$\left| \int_X f\, U^n g\, d\mu - \int_X f\, d\mu \int_X g\, d\mu \right| \leq C\lambda^n \|f\|_0 \|g\|_{L^1_\mu}.$$

The proof is classical, see appendix C.4 in [48] for the deterministic analogue.

We will now investigate limit theorems, namely a central limit theorem (CLT) and a large deviation principle (LDP), following Nagaev's perturbative approach. We refer to [49] for a full account of the theory. Let $\varphi \in \mathcal{B}_0$ be a bounded real observable with $\int_X \varphi\, d\mu = 0$. Define X_k on $\tilde{\Omega} \times X$ by $X_k(\underline{\omega}, x) = \varphi(T_{\omega_k} \dots T_{\omega_1} x)$ and $S_n = \sum_{k=0}^{n-1} X_k$. The first step is to prove the existence of the asymptotic variance.

Proposition 3.2 *The limit $\sigma^2 = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2)$ exists, and is equal to*

$$\sigma^2 = \int_X \varphi^2\, d\mu + 2\sum_{n=1}^{+\infty} \int_X \varphi\, U^n \varphi\, d\mu.$$

Proof We expand the term S_n^2 and get $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2) = \sum_{k,l=0}^{n-1} \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(X_k X_l)$.

Lemma 3.3 *For all integers k and l, one has $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(X_k X_l) = \int_X \varphi\, U^{|k-l|}\varphi\, d\mu$.*

Proof (Proof of the lemma) By symmetry, we can assume $k \geq l$. We have

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(X_k X_l) &= \int_X h(x) \int_{\tilde{\Omega}} X_k(\underline{\omega}, x) X_l(\underline{\omega}, x)\, d\tilde{\mathbb{P}}(\underline{\omega}) dm(x) \\
&= \int_X h(x) \int_{\tilde{\Omega}} (\varphi \circ T_{\omega_k} \circ \cdots \circ T_{\omega_{l+1}})(T_{\omega_l} \ldots T_{\omega_1} x)\varphi(T_{\omega_l} \ldots T_{\omega_1} x)\, d\tilde{\mathbb{P}}(\underline{\omega}) dm(x) \\
&= \int_X h(x) \int_{\tilde{\Omega}} U^l(\varphi(\varphi \circ T_{\omega_k} \circ \cdots \circ T_{\omega_{l+1}}))(x)\, d\tilde{\mathbb{P}}(\omega_{l+1}, \ldots) dm(x) \\
&= \int_{\tilde{\Omega}} \int_X P^l h(x)\varphi(x)\varphi(T_{\omega_k} \ldots T_{\omega_{l+1}} x)\, dm(x) d\tilde{\mathbb{P}}(\omega_{l+1}, \ldots) \\
&= \int_X \varphi(x) U^{k-l}\varphi(x)\, d\mu(x)
\end{aligned}$$

□

Applying this lemma, we get

$$\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2) = \sum_{k,l=0}^{n-1} \int_X \varphi\, U^{|k-l|}\varphi\, d\mu = n \int_X \varphi^2\, d\mu + 2\sum_{k=1}^{n}(n-k) \int_X \varphi\, U^k \varphi\, d\mu.$$

Since $\int_X \varphi\, U^k \varphi\, d\mu$ decays exponentially fast, we see immediately that $\frac{1}{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2)$ goes to the desired quantity. □

Let us mention that we have the following criteria to determine whether the asymptotic variance is 0. The proof follows along the same lines as Lemma 4.1 in [31].

Proposition 3.4 *The asymptotic variance satisfies $\sigma^2 = 0$ if and only if there exists $\psi \in L^2(\mu)$ such that, for $\mathbb{P}$-a.e. ω, $\varphi = \psi - \psi \circ T_\omega \mu$-a.e.*

Denote by $\mathcal{M}_{\mathcal{B}}$ the set of all probability measures on $(X, \mathcal{A})$ which are absolutely continuous w.r.t. m, and whose density lies in $\mathcal{B}$. For a measure $\nu \in \mathcal{M}_{\mathcal{B}}$, we will denote by $\|\nu\|$ the $\mathcal{B}$-norm of the density $\frac{d\nu}{dm}$. Now, we are able to state the main theorems of this section:

Theorem 3.5 (Central limit theorem) *For every probability measure $\nu \in \mathcal{M}_{\mathcal{B}}$, the process $(\frac{S_n}{\sqrt{n}})_n$ converges in law to $\mathcal{N}(0, \sigma^2)$ under the probability $\tilde{\mathbb{P}} \otimes \nu$.*

Theorem 3.6 (Large deviation principle) *Suppose that $\sigma^2 > 0$. Then there exists a non-negative rate function c, continuous, strictly convex, vanishing only at 0, such that for every $\nu \in \mathcal{M}_{\mathcal{B}}$ and every sufficiently small $\epsilon > 0$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \tilde{\mathbb{P}} \otimes \nu(S_n > n\epsilon) = -c(\epsilon)$$

In particular, these theorems are valid for both the reference measure m and the stationary one μ, with the same asymptotic variance and the same rate function.

We introduce Laplace operators, which will encode the moment-generating function of the process. For every $z \in \mathbb{C}$, we define P_z by $P_z(f) = P(e^{z\varphi} f)$. Thanks to our assumption on $\mathcal{B}_0$, this a well defined and continuous operator on $\mathcal{B}$, and the map $z \mapsto P_z$ is complex-analytic on $\mathbb{C}$: indeed, if we define $C_n(f) = P(\varphi^n f)$, then $P_z = \sum_{n\geq 0} \frac{z^n}{n!} C_n$, and this series is convergent on the whole complex plane since $\|C_n\| \leq (C\|\varphi\|_0)^n \|P\|$.

We have the following fundamental relation:

Lemma 3.7 *For every $n \geq 0$ and every $f \in \mathcal{B}$, we have*

$$\int_{\tilde{\Omega}} \int_X e^{zS_n(\underline{\omega},x)} f(x)\, dm(x)\, d\tilde{\mathbb{P}}(\underline{\omega}) = \int_X P_z^n(f)\, dm.$$

Proof We proceed by induction on n. The case $n = 0$ is trivial. Assume that the relation is valid for some $n \geq 0$, and all $f \in \mathcal{B}$. Let f be a member of $\mathcal{B}$. Since $P_z(f)$ belongs to $\mathcal{B}$, the induction hypothesis gives

$$\begin{aligned}
\int_X P_z^{n+1}(f)\, dm &= \int_X P_z^n(P_z(f))\, dm = \int_{\tilde{\Omega}} \int_X e^{zS_n(\underline{\omega},x)} P_z(f)(x)\, dm(x) d\tilde{\mathbb{P}}(\underline{\omega}) \\
&= \int_{\tilde{\Omega}} \int_X e^{zS_n(\underline{\omega},x)} P(e^{z\varphi} f)(x)\, dm(x) d\tilde{\mathbb{P}}(\underline{\omega}) \\
&= \int_{\tilde{\Omega}} \int_X U(e^{zS_n(\underline{\omega},\,\cdot)})(x) e^{z\varphi(x)} f(x)\, dm(x) d\tilde{\mathbb{P}}(\underline{\omega}) \\
&= \int_X \int_{\tilde{\Omega}} \int_{\Omega} e^{z(\varphi(x)+S_n(\underline{\omega},T_\omega x))}\, d\mathbb{P}(\omega) d\tilde{\mathbb{P}}(\underline{\omega}) f(x) dm(x)
\end{aligned}$$

But $\varphi(x) + S_n(\underline{\omega}, T_\omega x) = S_{n+1}(\omega\underline{\omega}, x)$, where $\omega\underline{\omega}$ stands for the concatenation $(\omega, \omega_1, \omega_2, \ldots)$ if $\underline{\omega} = (\omega_1, \omega_2, \ldots)$. As $\int_{\tilde{\Omega}} \int_{\Omega} e^{zS_{n+1}(\omega\underline{\omega},x)}\, d\mathbb{P}(\omega) d\tilde{\mathbb{P}}(\underline{\omega}) = \int_{\tilde{\Omega}} e^{zS_{n+1}(\underline{\omega},x)} d\tilde{\mathbb{P}}(\underline{\omega})$ because of the product structure of $\tilde{\mathbb{P}}$, we obtain the formula for $n+1$ and f. □

If f is the density w.r.t. m of a probability measure $\nu \in \mathcal{M}_{\mathcal{B}}$, by the previous lemma, we know that the moment-generating function of S_n under the probability measure $\tilde{\mathbb{P}} \otimes \nu$ is given by $\int_X P_z^n(f)\, dm$ This leads us to understand the asymptotic behavior of the iterates of the Laplace operators P_z. Since they are smooth perturbations of the quasi-compact operator P, one can apply here the standard theory of perturbations for linear operators (see for instance theorem III.8 in [49]), and get the following:

Lemma 3.8 *There exists $\epsilon_1 > 0, \eta_1 > 0, \eta_2 > 0$, and complex-analytic functions $\lambda(.), h(.), m(.), Q(.)$, all defined on $\mathbb{D}_{\epsilon_1} = \{z \in \mathbb{C} / |z| < \epsilon_1\}$, which take values respectively in $\mathbb{C}, \mathcal{B}, \mathcal{B}^\star, \mathcal{L}(\mathcal{B})$ and satisfying for all $z \in \mathbb{D}_{\epsilon_1}$:*

1. $\lambda(0) = 1, h(0) = h, m(0) = m, Q(0) = Q$;
2. $P_z(f) = \lambda(z)\langle m(z), f\rangle h(z) + Q(z)f$ *for all* $f \in \mathcal{B}$;
3. $\langle m(z), h(z)\rangle = 1$;
4. $Q(z)h(z) = 0$ *and* $m(z)Q(z) = 0$;
5. $|\lambda(z)| > 1 - \eta_1$;
6. $\|Q(z)^n\| \leq C(1 - \eta_1 - \eta_2)^n$.

Furthermore, $|\langle m, Q(z)^n f\rangle| \leq C|z|(1 - \eta_1 - \eta_2)^n \|f\|$ *for all* $f \in \mathcal{B}$ *and* $z \in \mathbb{D}_{\epsilon_1}$.

For all $n \geq 0$, we hence have $P_z^n(f) = \lambda(z)^n \langle m(z), f\rangle h(z) + Q(z)^n f$. The asymptotic behavior of P_z^n is clearly intimately related to the behavior of the leading eigenvalue $\lambda(z)$ in a neighborhood of 0. We have the following:

Lemma 3.9 *The leading eigenvalue* $\lambda(.)$ *satisfies* $\lambda'(0) = \int \varphi d\mu = 0$ *and* $\lambda''(0) = \sigma^2 \geq 0$.

Proof By corollary III.11 in [49], $\lambda'(0) = \langle m(0), P'(0)h(0)\rangle$. As $m(0) = m$, $h(0) = h = \frac{d\mu}{dm}$ and $P'(0)f = C_1(f) = P(\varphi f)$ for any $f \in \mathcal{B}$, since $P(z) = \sum_{n\geq 0} \frac{z^n}{n!} C_n$ with $C_n(f) = P(\varphi^n f)$, the formula for $\lambda'(0)$ reads as

$$\lambda'(0) = \langle m, P(\varphi h)\rangle = \langle m, \varphi h\rangle = \int \varphi d\mu = 0.$$

Using again corollary III.11 in [49], we have

$$\lambda''(0) = \langle m(0), P''(0)h(0)\rangle + 2\langle m(0), P'(0)\tilde{h}\rangle,$$

where $\tilde{h}$ is the unique element of $\mathcal{B}$ satisfying $\langle m(0), \tilde{h}\rangle = 0$ and $(\lambda(0) - P(0))\tilde{h} = (P'(0) - \lambda'(0))h(0)$.

This implies that $\tilde{h}$ is the unique element of $\mathcal{B}$ satisfying $\langle m, \tilde{h}\rangle = 0$ and $(I - P)\tilde{h} = P(\varphi h)$. By corollary III.6 in [49], $\tilde{h}$ is given by $\tilde{h} = \sum_{n\geq 0} Q(0)^n(\varphi h)$. But $Q(0)^n(\varphi h) = P^n(\varphi h) - \langle m, \varphi h\rangle h = P^n(\varphi h)$, since $\langle m, \varphi h\rangle = \int \varphi d\mu = 0$. Hence $\tilde{h} = \sum_{n\geq 0} P^n(P(\varphi h)) = \sum_{n\geq 1} P^n(\varphi h)$.

On one hand, we have $\langle m(0), P''(0)h(0)\rangle = \langle m, C_2(h)\rangle = \langle m, P(\varphi^2 h)\rangle = \int \varphi^2 d\mu$. On the other hand,

$$\begin{aligned}\langle m(0), P'(0)\tilde{h}\rangle = \langle m, P(\varphi\tilde{h})\rangle = \langle m, \varphi\tilde{h}\rangle &= \sum_{n\geq 1}\langle m, \varphi P^n(\varphi h)\rangle \\ &= \sum_{n\geq 1}\int \varphi P^n(\varphi h) dm = \sum_{n\geq 1}\int U^n\varphi\, \varphi d\mu.\end{aligned}$$

Summing these two parts, we recognize the formula for σ^2 given by Proposition 3.2. □

Then, $\lambda(\frac{it}{\sqrt{n}})^n = (1 - \frac{\sigma^2 t^2}{2n} + o(\frac{1}{n}))^n$ goes to $e^{-\frac{\sigma^2 t^2}{2}}$, from which it follows that $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{i\frac{t}{\sqrt{n}}S_n}) = \lambda(\frac{it}{\sqrt{n}})^n \langle m(\frac{it}{\sqrt{n}}), f\rangle\langle m, h(\frac{it}{\sqrt{n}})\rangle + \langle m, Q(\frac{it}{\sqrt{n}})f\rangle$ goes also to $e^{-\frac{\sigma^2 t^2}{2}}$,

for each $t \in \mathbb{R}$, when $n \to \infty$, which it implies the CLT by Lévy's continuity theorem. Remark that the previous identity holds for any measure $\nu \in \mathcal{M}_{\mathcal{B}}$ and their associated density f.

We can furthermore prove a rate of convergence in the CLT, when $\sigma^2 > 0$:

Lemma 3.10 *There exists $C > 0$ and $\rho < 1$ such that for all $t \in \mathbb{R}$ and $n \geq 0$ with $\frac{|t|}{\sqrt{n}}$ sufficiently small, and all $\nu \in \mathcal{M}_{\mathcal{B}}$, we have*

$$|\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{i\frac{t}{\sqrt{n}}S_n}) - e^{-\frac{1}{2}\sigma^2 t^2}| \leq C\|\nu\| \left(e^{-\frac{\sigma^2 t^2}{2}}\left(\frac{|t|+|t|^3}{\sqrt{n}}\right) + \frac{|t|}{\sqrt{n}}\rho^n\right).$$

Proof This follows from the third order differentiability of $\lambda(.)$ at 0: for $\frac{t}{\sqrt{n}}$ small enough, $\lambda(\frac{it}{\sqrt{n}})^n = (1 - \frac{\sigma^2 t^2}{2n} + \mathcal{O}(\frac{|t|^3}{n\sqrt{n}}))^n = e^{-\frac{\sigma^2 t^2}{2}} + \mathcal{O}(e^{-\frac{\sigma^2 t^2}{2}}\frac{|t|^3}{\sqrt{n}})$. Recalling that $f = \frac{d\nu}{dm}$ and $\|\nu\| = \|f\| \geq C$, where the constant C comes from the continuous embedding $\mathcal{B} \subset L^1(m)$ and is independent of ν, we have

$$\begin{aligned}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{i\frac{t}{\sqrt{n}}S_n}) &= \lambda\left(\frac{it}{\sqrt{n}}\right)^n \left\langle m\left(\frac{it}{\sqrt{n}}\right), f\right\rangle \left\langle m, h\left(\frac{it}{\sqrt{n}}\right)\right\rangle + \left\langle m, Q\left(\frac{it}{\sqrt{n}}\right)^n f\right\rangle \\ &= \left(e^{-\frac{\sigma^2 t^2}{2}} + \mathcal{O}\left(e^{-\frac{\sigma^2 t^2}{2}}\frac{|t|^3}{\sqrt{n}}\right)\right)\left(1 + \mathcal{O}\left(\frac{|t|}{\sqrt{n}}\|f\|\right)\right) \\ &\quad + \mathcal{O}\left(\frac{|t|}{\sqrt{n}}\rho^n\|f\|\right),\end{aligned}$$

where the first line follows from Lemma 3.7 applied to f and $z = \frac{it}{\sqrt{n}}$ and item 2 of Lemma 3.8. This implies the result. □

From this lemma, we deduce that $|\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{i\frac{t}{\sqrt{n}}S_n}) - e^{-\frac{1}{2}\sigma^2 t^2}| = \mathcal{O}(\frac{1+|t|^3}{\sqrt{n}})$, which will be useful latter, when proving a quenched CLT. The precise estimate of the lemma also implies a rate of convergence of order $\frac{1}{\sqrt{n}}$ in the CLT, using the Berry–Esséen inequality. We refer to [49] or [50] for a scheme of proof:

Theorem 3.11 *If $\sigma^2 > 0$, there exists $C > 0$ such that for all $\nu \in \mathcal{M}_{\mathcal{B}}$:*

$$\sup_{t\in\mathbb{R}}\left|\tilde{\mathbb{P}}\otimes\nu\left(\frac{S_n}{\sqrt{n}} \leq t\right) - \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{u^2}{2\sigma^2}}du\right| \leq \frac{C\|\nu\|}{\sqrt{n}}.$$

We turn now to the proof of the LDP. For this, we will show the convergence of $\frac{1}{n}\log\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\theta S_n})$ for small enough $\theta \in \mathbb{R}$ and then apply Gartner-Ellis theorem [51,52]. Proofs are a verbatim copy of those from [53].

Lemma 3.12 *There exists $0 < \epsilon_2 < \epsilon_1$ such that for every $\theta \in \mathbb{R}$ with $|\theta| < \epsilon_2$, we have $\lambda(\theta) > 0$. Furthermore, the functions $h(.)$ and $m(.)$ can be redefined in such a way that they still satisfy conclusions of lemma 3.8, while they also verify $h(\theta) \geq 0$, $m(\theta) \geq 0$ for $\theta \in \mathbb{R}$.*

Proof As P_θ is a real operator, we have $P_\theta \overline{f} = \overline{P_\theta f}$ for all $f \in \mathcal{B}$. So, we have $P_\theta \overline{h(\theta)} = \overline{P_\theta h(\theta)} = \overline{\lambda(\theta) h(\theta)}$. Since $\lambda(\theta)$ is the unique eigenvalue of P_θ with maximal modulus, we get $\overline{\lambda(\theta)} = \lambda(\theta)$, and hence $\lambda(\theta) \in \mathbb{R}$. Since $\lambda(0) = 1$, by a continuity argument, we obtain $\lambda(\theta) > 0$ for small θ. For $z \in \mathbb{C}$ small enough, $\langle m(z), \mathbb{1}\rangle \neq 0$. We define $\tilde{h}(z) = \langle m(z), \mathbb{1}\rangle h(z)$ and $\tilde{m}(z) = \langle m(z), \mathbb{1}\rangle^{-1} m(z)$. Those new eigenfunctions satisfy obviously the conclusions of the previous proposition. We have just to prove that $\tilde{h}(\theta)$ and $\tilde{m}(\theta)$ are positive for $\theta \in \mathbb{R}$ small enough. By the spectral decomposition of P_θ, we see that $\lambda(\theta)^{-n} P_\theta^n \mathbb{1}$ goes to $\tilde{h}(\theta)$ in $\mathcal{B}$, and hence in $L^1(m)$. We then get $\tilde{h}(\theta) \geq 0$ because P_θ is a positive operator and $\lambda(\theta)$ is positive too. Now, let $\psi(\theta) \in \mathcal{B}^\star$ positive such that $\langle \psi(\theta), \tilde{h}(\theta)\rangle = 1$[1]. Then, $\lambda(\theta)^{-n}(P_\theta^\star)^n \psi(\theta)$ goes to $\langle \psi(\theta), h(\theta)\rangle m(\theta) = \tilde{m}(\theta)$, which proves that $\tilde{m}(\theta)$ is a positive linear form. □

We denote $\Lambda(\theta) = \log \lambda(\theta)$. We then have

Proposition 3.13 *For every $\nu \in \mathcal{M}_{\mathcal{B}}$, there exists $0 < \epsilon_3 < \epsilon_2$ such that for every $\theta \in \mathbb{R}$ with $|\theta| < \epsilon_3$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\tilde{\mathbb{P}}_{\otimes\nu}}(e^{\theta S_n}) = \Lambda(\theta)$$

Proof Let $f \in \mathcal{B}$ be the density $\frac{d\nu}{dm}$. We have the identity

$$\begin{aligned}\mathbb{E}_{\tilde{\mathbb{P}}_{\otimes\nu}}(e^{\theta S_n}) = \langle m, P_\theta^n(f)\rangle &= \lambda(\theta)^n \langle m(\theta), f\rangle \langle m, h(\theta)\rangle + \langle m, Q(\theta)^n f\rangle \\ &= \lambda(\theta)^n \left(\langle m(\theta), f\rangle \langle m, h(\theta)\rangle + \lambda(\theta)^{-n}\langle m, Q(\theta)^n f\rangle\right)\end{aligned}$$

All involved quantities are positive, hence we can write

$$\frac{1}{n} \log \mathbb{E}_{\tilde{\mathbb{P}}_{\otimes\nu}}(e^{\theta S_n}) = \log \lambda(\theta) + \frac{1}{n} \log \left(\langle m(\theta), f\rangle \langle m, h(\theta)\rangle + \lambda(\theta)^{-n}\langle m, Q(\theta)^n f\rangle\right)$$

Since $\lim_{\theta\to 0}\langle m(\theta), f\rangle \langle m, h(\theta)\rangle = 1$ and since the spectral radius of $Q(\theta)$ is strictly less than $\lambda(\theta)$, it's easy to see that for θ small enough, we have

$$\lim_{n\to\infty} \frac{1}{n} \log \left(\langle m(\theta), f\rangle \langle m, h(\theta)\rangle + \lambda(\theta)^{-n}\langle m, Q(\theta)^n f\rangle\right) = 0.$$

□

To complete the proof, it suffices to prove that Λ is a differentiable function, strictly convex in a neighborhood of 0, which is indeed the case since $\lambda(.)$ is complex-analytic and we have supposed $\Lambda''(0) = \lambda''(0) = \sigma^2 > 0$. A local version of the Gartner-Ellis theorem (a precise statement can be found e.g. in lemma XIII.2 in [49]) finishes the proof.

It is interesting to notice that the annealed LDP implies almost immediately a quenched upper bound, with the same rate function for almost every realization:

[1] Choose $\psi(\theta) = \alpha(\theta)^{-1} m$, where $\alpha(\theta) = \langle m, \tilde{h}(\theta)\rangle$ is positive, since $\tilde{h}(\theta)$ is a positive element of $L^1(m)$.

Proposition 3.14 *For every $\nu \in \mathcal{M}_{\mathcal{B}}$, for every small enough $\epsilon > 0$ and for $\tilde{\mathbb{P}}$-almost every $\underline{\omega}$, we have*

$$\limsup_{n\to\infty} \frac{1}{n} \log \nu(\{x \in X \,/\, S_n(\underline{\omega}, x) > n\epsilon\}) \leq -c(\epsilon)$$

Proof Let $\epsilon > 0$ small enough such that the annealed LDP holds. Let $0 < \gamma < 1$ and define

$$A_n = \{\underline{\omega} \in \tilde{\Omega} \,/\, \nu(\{x \in X \,/\, S_n(\underline{\omega}, x) > n\epsilon\}) \geq e^{-n(1-\gamma)c(\epsilon)}\}.$$

By the annealed LDP, we have $\tilde{\mathbb{P}} \otimes \nu(S_n > n\epsilon) \leq Ce^{-n(1-\frac{\gamma}{2})c(\epsilon)}$ for some $C = C(\gamma, \epsilon)$, and hence Markov inequality yields

$$\tilde{\mathbb{P}}(A_n) \leq e^{n(1-\gamma)c(\epsilon)} \tilde{\mathbb{P}} \otimes \nu(S_n > n\epsilon) \leq Ce^{-n\frac{\gamma}{2}c(\epsilon)}.$$

By the Borel–Cantelli lemma, we have that $\tilde{\mathbb{P}}$-almost every $\underline{\omega}$ lies in finitely many A_n whence $\limsup_{n\to\infty} \frac{1}{n} \log \nu(\{x \in X \,/\, S_n(\underline{\omega}, x) > n\epsilon\}) \leq -(1-\gamma)c(\epsilon)$ for $\tilde{\mathbb{P}}$-almost every $\underline{\omega}$. As γ can be a rational number arbitrarily close to 0, we get $\limsup_{n\to\infty} \frac{1}{n} \log \nu(\{x \in X \,/\, S_n(\underline{\omega}, x) > n\epsilon\}) \leq -c(\epsilon)$ for $\tilde{\mathbb{P}}$-almost every $\underline{\omega}$. □

We can also prove a local limit theorem.

Definition 3.15 We will say that φ is aperiodic if for all $t \neq 0$, the spectral radius of P_{it} is strictly less than 1.

Theorem 3.16 (Local Limit Theorem) *If $\sigma^2 > 0$ and φ is aperiodic, then, for all $\nu \in \mathcal{M}_{\mathcal{B}}$ and all bounded interval $I \subset \mathbb{R}$,*

$$\lim_{n\to\infty} \sup_{s\in\mathbb{R}} \left| \sigma\sqrt{n}\, \tilde{\mathbb{P}} \otimes \nu(s + S_n \in I) - \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2n\sigma^2}} |I| \right| = 0.$$

Proof We follow the proof given by Breiman [54] in the iid case. See also Rousseau-Egele [5] for a proof in a dynamical context. By a density argument, it is sufficient to prove that, uniformly in $s \in \mathbb{R}$, $|\sigma\sqrt{n}\, \mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(g(S_n + s)) - \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2n\sigma^2}} \int_{\mathbb{R}} g(u)du|$ goes to 0 as $n \to \infty$, for all $g \in L^1(\mathbb{R})$ for which the Fourier transform $\hat{g}$ is continuous with compact support. Using Fourier's inversion formula, we first write

$$\sigma\sqrt{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(g(S_n + s)) = \frac{\sigma\sqrt{n}}{2\pi} \int_{\mathbb{R}} e^{its} \hat{g}(t) \left(\int_X P_{it}^n(f)\, dm \right) dt.$$

Let $\delta > 0$ be such that the support of $\hat{g}$ is included in $[-\delta, +\delta]$, and, remembering that $\lambda(it) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2)$ and $\langle m(it), f\rangle\langle m, h(it)\rangle = 1 + \mathcal{O}(|t|)$, choose

$0 < \tilde{\delta} < \delta$ small enough in such a way that $|\lambda(it)| \leq 1 - \frac{\sigma^2 t^2}{4} \leq e^{-\frac{t^2\sigma^2}{4}}$ and $|\langle m(it), f\rangle\langle m, h(it)\rangle - 1| \leq C|t|$ for $|t| < \tilde{\delta}$. Using

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2n\sigma^2}} \int_{\mathbb{R}} g(u)du = \frac{\hat{g}(0)\sigma}{2\pi} \int_{\mathbb{R}} e^{\frac{its}{\sqrt{n}}} e^{-\frac{\sigma^2 t^2}{2}} dt$$

and

$$\int_X P_{it}^n(f)\, dm = \lambda(it)^n \langle m, h(it)\rangle\langle m(it), f\rangle + \langle m, Q(it)^n f\rangle$$

for $|t| < \tilde{\delta}$, we can write

$$\begin{aligned}
&\sigma\sqrt{n}\,\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(g(S_n + s)) - \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2n\sigma^2}} \int_{\mathbb{R}} g(u)du \\
&= \frac{\sigma}{2\pi}\Bigg(\int_{|t|<\tilde{\delta}\sqrt{n}} e^{\frac{its}{\sqrt{n}}} \left(\hat{g}\left(\frac{t}{\sqrt{n}}\right)\lambda\left(\frac{it}{\sqrt{n}}\right)^n - \hat{g}(0)e^{-\frac{\sigma^2 t^2}{2}}\right) dt \\
&\quad + \int_{|t|<\tilde{\delta}\sqrt{n}} e^{\frac{its}{\sqrt{n}}} \hat{g}\left(\frac{t}{\sqrt{n}}\right)\lambda\left(\frac{it}{\sqrt{n}}\right)^n \left(\left\langle m\left(\frac{it}{\sqrt{n}}\right), f\right\rangle\left\langle m, h\left(\frac{it}{\sqrt{n}}\right)\right\rangle - 1\right) dt \\
&\quad + \sqrt{n} \int_{|t|<\tilde{\delta}} e^{its}\hat{g}(t)\langle m, Q(it)^n f\rangle dt \\
&\quad + \sqrt{n} \int_{\tilde{\delta}\leq|t|\leq\delta} e^{its}\hat{g}(t)\langle m, P_{it}^n f\rangle dt - \hat{g}(0) \int_{|t|\geq\tilde{\delta}\sqrt{n}} e^{\frac{its}{\sqrt{n}}} e^{-\frac{\sigma^2 t^2}{2}} dt \Bigg) \\
&= \frac{\sigma}{2\pi}(A_n^{(1)}(s) + A_n^{(2)}(s) + A_n^{(3)}(s) + A_n^{(4)}(s) + A_n^{(5)}(s)).
\end{aligned}$$

One has $|A_n^{(1)}(s)| \leq \int_{|t|<\tilde{\delta}\sqrt{n}} |\hat{g}(\frac{t}{\sqrt{n}})\lambda(\frac{it}{\sqrt{n}})^n - \hat{g}(0)e^{-\frac{\sigma^2 t^2}{2}}|dt$, so $\sup_{s\in\mathbb{R}} |A_n^{(1)}(s)| \to 0$ by dominated convergence, since $|\hat{g}(\frac{t}{\sqrt{n}})\lambda(\frac{it}{\sqrt{n}})^n - \hat{g}(0)e^{-\frac{\sigma^2 t^2}{2}}| \leq \|\hat{g}\|_{\sup}(e^{-\frac{\sigma^2 t^2}{4}} + e^{-\frac{\sigma^2 t^2}{2}})$ is integrable on $\mathbb{R}$ and $\hat{g}(\frac{t}{\sqrt{n}})\lambda(\frac{it}{\sqrt{n}})^n \to \hat{g}(0)e^{-\frac{\sigma^2 t^2}{2}}$.

For the second term, we can bound it by $\frac{C\|\hat{g}\|_{\sup}}{\sqrt{n}} \int_{|t|<\tilde{\delta}\sqrt{n}} |t|e^{-\frac{\sigma^2 t^2}{4}} dt = \mathcal{O}(\frac{1}{\sqrt{n}})$ and so $\sup_{s\in\mathbb{R}} |A_n^{(2)}(s)| \to 0$.

The third term is bounded by $C\sqrt{n}\|\hat{g}\|_{\sup}\rho^n$, and so $\sup_{s\in\mathbb{R}} |A_n^{(3)}(s)| \to 0$.

By dominated convergence, we have clearly $\sup_{s\in\mathbb{R}} |A_n^{(5)}(s)| \to 0$, so it remains to deal with the fourth term. This is where the aperiodicity assumption plays a role.

Denoting by $r(P_{it})$ the spectral radius of the operator P_{it}, we know that the u.s.c. function $t \mapsto r(P_{it})$ reaches its maximum on the compact set $\{\tilde{\delta} \leq |t| \leq \delta\}$, which is then < 1 by assumption. Since the set $\{P_{it}\}_{\tilde{\delta} \leq |t| \leq \delta}$ is bounded, there exists C and $\theta < 1$ such that $\|P_{it}^n\| \leq C\theta^n$ for all $\tilde{\delta} \leq |t| \leq \delta$ and all $n \geq 0$. Then one has $\sup_{s \in \mathbb{R}} |A_n^{(4)}(s)| \leq C\sqrt{n}\theta^n \|\hat{g}\|_{\sup} \|m\| \|f\| \to 0$, which concludes the proof. □

We give a concrete criterion to check the aperiodicity assumption:

Proposition 3.17 *Assume that the stationary measure μ is equivalent to m, and that the spectral radius (resp. the essential spectral radius) of P_{it} is less (resp. strictly less) than 1 for all $t \in \mathbb{R}$. If φ is not aperiodic, then there exists $t \neq 0$, $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ and a measurable function $g : X \to \mathbb{C}$ such that $gh \in \mathcal{B}$ and $\lambda g(T_\omega x) = e^{it\varphi(x)} g(x)$ for m-ae x and $\mathbb{P}$-ae ω.*

Proof Suppose that the spectral radius of P_{it} is greater or equal than 1 for some $t \neq 0$. By the assumptions on the spectral radius, this implies that there is an eigenvalue λ of P_{it} satisfying $|\lambda| = 1$. Let $f \in \mathcal{B}$ a corresponding eigenvector, and define $g = \frac{f}{h}$. This definition makes sense m-ae, by the assumption on μ. We then have $P(\phi gh) = gh$, where $\phi = \bar{\lambda} e^{it\varphi}$. We then lift this relation to the skew-product: by Lemma 4.3, we have $P_S(\phi_\pi g_\pi h_\pi) = g_\pi h_\pi$, where P_S is the transfer operator for the skew-product system, defined w.r.t. the measure $\tilde{\mathbb{P}} \otimes m$. See Sect. 4 for the notations. By Proposition 1.1 in Morita [55], we deduce that $g_\pi \circ S = \phi_\pi g_\pi$, $\tilde{\mathbb{P}} \otimes m$-ae. We conclude the proof by writing explicitly this relation. □

Remark 3.18 1. The assumptions on the spectral radius of P_{it} are usually proved by mean of a Lasota–Yorke inequality for each P_{it}, which usually follow in the same way we prove a Lasota–Yorke inequality for the transfer operator P. See the works of Rousseau-Egele [5], Morita [55], Broise [56] or Aaronson et al. [57] for one-dimensional deterministic examples.

2. The previous Proposition shows that if φ is not aperiodic for the random system, then it is not aperiodic for almost each deterministic system T_ω in the usual sense, and that almost all aperiodicity equations share a common regular solution g. For instance, if the set Ω is finite and if we know that φ is aperiodic for one map T_ω, then it is aperiodic for the random system. This can be checked using known techniques, see [56–58] among many others for more details.

We conclude this section with an annealed vector-valued almost sure invariance principle. First recall the definition.

Definition 3.19 For $\lambda \in (0, \frac{1}{2}]$, and Σ^2 a (possibly degenerate) symmetric semi-positive-definite $d \times d$ matrix, we say that an $\mathbb{R}^d$-valued process $(X_n)_n$ satisfies an almost sure invariance principle (ASIP) with error exponent λ and limiting covariance Σ^2 if there exist, on another probability space, two processes $(Y_n)_n$ and $(Z_n)_n$ such that:

1. the processes $(X_n)_n$ and $(Y_n)_n$ have the same distribution;
2. the random variables Z_n are independent and distributed as $\mathcal{N}(0, \Sigma^2)$;
3. almost surely, $|\sum_{k=0}^{n-1} Y_k - \sum_{k=0}^{n-1} Z_k| = o(n^\lambda)$.

The ASIP has a lot of consequences, such as a functional central limit theorem, a law of the iterated logarithm, etc... See Melbourne and Nicol [59] and references therein for more details.

Let $\varphi : X \to \mathbb{R}^d$ be a bounded vector-valued observable such that each component $\varphi_j : X \to \mathbb{R}$, $j = 1, \dots, d$, belongs to $\mathcal{B}_0$, with $\int_X \varphi_j \, d\mu = 0$. Define as before $X_k(\underline{\omega}, x) = \varphi(T^k_{\underline{\omega}} x)$.

Theorem 3.20 *The covariance matrix* $\frac{1}{n}\mathrm{cov}(\sum_{k=0}^{n-1} X_k)$ *converges to a matrix* Σ^2 *and the process* $(X_n)_n$*, defined on the probability space* $(\tilde{\Omega} \times X, \tilde{\mathbb{P}} \otimes \mu)$*, satisfies an ASIP with limiting covariance* Σ^2*, for any error exponent* $\lambda > \frac{1}{4}$.

Proof We will apply results from Gouëzel [60]. Namely, we construct a family of operators $(\mathcal{L}_t)_{t \in \mathbb{R}^d}$ acting on $\mathcal{B}$ which codes the characteristic function of the process $(X_n)_n$ and we check assumptions (I1) and (I2) from [60]. For $k \geq 0$ and $j_1, \dots, j_k \in \{1, \dots, d\}$, define $C_{j_1,\dots,j_k}$ by $C_{j_1,\dots,j_k}(f) = P(\varphi_{j_1} \dots \varphi_{j_k} f)$. The assumptions on $\mathcal{B}_0$ and $\mathcal{B}$ show that $C_{j_1,\dots,j_k}$ acts continuously on $\mathcal{B}$, with a norm bounded by $C^k \|\varphi_{j_1}\|_0 \dots \|\varphi_{j_k}\|_0$. Now, for $t = (t_1, \dots, t_d) \in \mathbb{R}^d$, define

$$\mathcal{L}_t = \sum_{k=0}^{\infty} \frac{i^k}{k!} \sum_{j_1,\dots,j_k=1}^{d} t_{j_1} \dots t_{j_k} C_{j_1,\dots,j_k}.$$

This defines on $\mathbb{R}^d$ a real-analytic family of bounded operators on $\mathcal{B}$, since

$$\sum_{k=0}^{\infty} \| \frac{i^k}{k!} \sum_{j_1,\dots,j_k=1}^{d} t_{j_1} \dots t_{j_k} C_{j_1,\dots,j_k} \| \leq \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j_1,\dots,j_k=1}^{d} |t_{j_1}| \dots |t_{j_k}| \|C_{j_1,\dots,j_k}\|$$
$$\leq e^{C \sum_{j=0}^{d} |t_j| \|\varphi_j\|_0} < \infty.$$

For $t \in \mathbb{R}^d$ and $f \in \mathcal{B}$, we have $\mathcal{L}_t(f) = P(e^{i\langle t, \varphi \rangle} f)$, and so the family $\{\mathcal{L}_t\}_{t \in \mathbb{R}^d}$ codes the characteristic function of the process $(X_n)_n$ in the sense of [60], as easily seen using Lemma 3.7. Since $\mathcal{L}_0 = P$ has a spectral gap on $\mathcal{B}$, this implies (I1). To check (I2), we only need, by proposition 2.3 in [60], the continuity of the map $t \mapsto \mathcal{L}_t$ at $t = 0$, but this follows immediately from the real-analyticity of this map. □

4 Annealed central limit theorem via a martingale approximation

The main goal of this section is to show that the classical martingale approach to the CLT, see Gordin [34] and Liverani [61], can be easily adapted to the random setting, leading to a new proof of Theorem 3.5 for the stationary measure μ, together with a generalization, in the next section, where a sequence of observables is considered, rather than a single one.

In this section, the annealed transfer operator P and Koopman operator U are defined by duality with respect to the stationary measure μ, instead of the measure m.

We assume moreover that we have decay of correlations for observables in $\mathcal{B}$ against $L^1(\mu)$, in the sense that

$$\left| \int_X f\, U^n g\, d\mu - \int_X f\, d\mu \int_X g\, d\mu \right| \leq C\lambda^n \|f\| \|g\|_{L^1_\mu}$$

for all $f \in \mathcal{B}$ and $g \in L^1(\mu)$. This is the case for instance if we assume that the density h of the stationary measure is bounded away from 0 and that $\mathcal{B}$ is continuously embedded in $L^1(m)$, see Proposition 3.1.

Recall that we have the Markov operator U which acts on functions defined on X by $Uf(x) = \int_\Omega f(T_\omega x)\, d\mathbb{P}(\omega)$. To U is associated a transition probability on X defined by $U(x, A) = U(\mathbb{1}_A)(x) = \mathbb{P}(\{\omega \,/\, T_\omega x \in A\})$. Recall also that the stationary measure μ satisfies $\mu U = \mu$, by definition. We can then define the canonical Markov chain associated to μ and U:

Let $\Omega^\star = X^{\mathbb{N}_0} = \{\underline{x} = (x_0, x_1, x_2, \ldots, x_n, \ldots)\}$, endowed with the σ-algebra $\mathcal{F}$ generated by cylinder sets. As X is Polish, this is also the Borel σ-algebra associated with the product topology. The Ionescu-Tulcea's theorem asserts there exists an unique probability measure μ_c on $\Omega^\star$ such that

$$\int_{\Omega^\star} f(\underline{x})\, d\mu_c(\underline{x}) = \int_X \mu(dx_0) \int_X U(x_0, dx_1) \ldots \int_X U(x_{n-1}, dx_n) f(x_0, \ldots, x_n)$$

for every n and every bounded measurable function $f : \Omega^\star \to \mathbb{R}$ which depends only on $x_0, \ldots, x_n$. If we still denote by x_n the map which associates to each $\underline{x}$ its nth coordinate x_n, then $\{x_n\}_{n\geq 0}$ is a Markov chain defined on the probability space $(\Omega^\star, \mathcal{F}, \mu_c)$, with initial distribution μ, and transition probability U. By stationarity of the measure μ, each x_n is distributed accordingly to μ.

We can define an unilateral shift τ on $\Omega^\star$. By stationarity, it preserves μ_c. Recall also that we have a skew-product system S on $\Omega^{\mathbb{N}} \times X$, defined by $S(\underline{\omega}, x) = (\sigma\underline{\omega}, T_{\omega_1} x)$, where σ is the unilateral shift on $\Omega^{\mathbb{N}}$. S preserves the probability measure $\tilde{\mathbb{P}} \otimes \mu = \mathbb{P}^{\otimes\mathbb{N}} \otimes \mu$. This system is related to the shift on $\Omega^\star$ in the following way:

Define $\Phi : \Omega^{\mathbb{N}} \times X \to \Omega^\star$ by $\Phi(\underline{\omega}, x) = (x, T_{\omega_1} x, T_{\omega_2} T_{\omega_1} x, \ldots, T_{\omega_n} \ldots T_{\omega_1} x, \ldots) = \{p(S^n(\underline{\omega}, x))\}_{n\geq 0}$, where $p(\underline{\omega}, x) = x$. We have the following:

Lemma 4.1 *Φ is measurable, sends $\tilde{\mathbb{P}} \otimes \mu$ on μ_c, and satisfies $\Phi \circ S = \tau \circ \Phi$.*

Proof The only non-trivial thing to prove is that $\tilde{\mathbb{P}} \otimes \mu$ is sent on μ_c. For this, it is sufficient to prove that

$$\begin{aligned} &\int_X \mu(dx) \int_{\Omega^{\mathbb{N}}} f_0(x) f_1(T^1_{\underline{\omega}} x) \ldots f_n(T^n_{\underline{\omega}} x)\, d\tilde{\mathbb{P}}(\underline{\omega})\, d\mu(x) \\ &\quad = \int_X \mu(dx_0) f_0(x_0) \int_X U(x_0, dx_1) f_1(x_1) \ldots \int_X U(x_{n-1}, dx_n) f_n(x_n) \end{aligned}$$

for all $n \geq 0$ and all bounded measurable functions $f_0, \ldots, f_n : X \to \mathbb{R}$. We proceed by induction on n, the case $n = 0$ being obvious. We have

$$\begin{aligned}
&\int_X \mu(dx_0) f_0(x_0) \int_X U(x_0, dx_1) f_1(x_1) \ldots \int_X U(x_{n-1}, dx_n) f_n(x_n) \\
&\quad \times \int_X U(x_n, dx_{n+1}) f_{n+1}(x_{n+1}) \\
&= \int_X \mu(dx_0) f_0(x_0) \int_X U(x_0, dx_1) f_1(x_1) \ldots \int_X U(x_{n-1}, dx_n) f_n(x_n) U f_{n+1}(x_n) \\
&= \int_X \mu(dx) \int_{\Omega^{\mathbb{N}}} f_0(x) f_1(T^1_{\underline{\omega}} x) \ldots f_n(T^n_{\underline{\omega}} x) U f_{n+1}(T^n_{\underline{\omega}} x) \, d\tilde{\mathbb{P}}(\underline{\omega}) \, d\mu(x) \\
&= \int_X \mu(dx) \int_{\Omega^{\mathbb{N}}} f_0(x) f_1(T^1_{\underline{\omega}} x) \ldots f_n(T^n_{\underline{\omega}} x) f_{n+1}(T^{n+1}_{\underline{\omega}} x) \, d\tilde{\mathbb{P}}(\underline{\omega}) \, d\mu(x)
\end{aligned}$$

which concludes the proof. □

Let $\pi_n : \Omega^\star \to X$ be the projection operator $\pi_n(x_0, \ldots, x_n, \ldots) = x_n$ and $\pi = \pi_0$ We lift each $\phi : X \to \mathbb{R}$ on $\Omega^\star$ by $\phi_\pi = \phi \circ \pi$. We then have $\mathbb{E}_\mu(\phi) = \mathbb{E}_{\mu_c}(\phi_\pi)$. Notice that $\pi_n = \pi \circ \tau^n$ and $p \circ S^n = \pi_n \circ \Phi$.

For a fixed observable $\phi : X \to \mathbb{R}$ with zero μ-mean, define $X_k = \phi \circ p \circ S^k$ and $S_n = \sum_{k=0}^{n-1} X_k$. We have $X_k = \phi \circ \pi_k \circ \Phi = \phi_\pi \circ \tau^k \circ \Phi$. Hence $S_n = (\sum_{k=0}^{n-1} \phi_\pi \circ \tau^k) \circ \Phi$, and so the law of S_n under $\tilde{\mathbb{P}} \otimes \mu$ is the law of the nth Birkhoff sum of ϕ_π under μ_c. So, to prove the CLT for S_n under the probability measure $\tilde{\mathbb{P}} \otimes \mu$, it suffices to prove it for the Birkhoff sum of the observable ϕ_π for the symbolic system $(\Omega^\star, \tau, \mu_c)$.

To this end, we introduce the Koopman operator $\tilde{U}$ and the transfer operator $\tilde{P}$ associated to $(\Omega^\star, \tau, \mu_c)$. Since this system is measure-preserving, those operators satisfy $\tilde{P}^k \tilde{U}^k f = f$ and $\tilde{U}^k \tilde{P}^k f = \mathbb{E}_{\mu_c}(f | \mathcal{F}_k)$ for every μ_c integrable f, where $\mathcal{F}_k = \tau^{-k} \mathcal{F} = \sigma(x_k, x_{k+1}, \ldots)$

We have the following

Lemma 4.2 *For every $\phi : X \to \mathbb{R}$, we have $\tilde{P}(\phi_\pi) = (P\phi)_\pi$.*

We will deduce this result from the corresponding statement for the transfer operator P_S of the skew product:

Lemma 4.3 *For every $\phi : X \to \mathbb{R}$, we have $P_S(\phi \circ p) = (P\phi) \circ p$.*

Proof Let $\psi : \Omega^{\mathbb{N}} \times X \to \mathbb{R}$ be an arbitrary element of $L^\infty(\tilde{\mathbb{P}} \otimes \mu)$. We have to show that $\int_{\Omega^{\mathbb{N}} \times X} (\phi \circ p)(\psi \circ S) \, d(\tilde{\mathbb{P}} \otimes \mu) = \int_{\Omega^{\mathbb{N}} \times X} ((P\phi) \circ p) \psi \, d(\tilde{\mathbb{P}} \otimes \mu)$. But

$$\int_{\Omega^{\mathbb{N}}\times X} (\phi\circ p)(\psi\circ S)\, d(\tilde{\mathbb{P}}\otimes\mu) = \int_{\Omega^{\mathbb{N}}}\int_X \phi(x)\psi(\sigma\underline{w}, T_{\omega_1}x)\, d\mu(x) d\tilde{\mathbb{P}}(\underline{\omega})$$
$$= \int_{\Omega^{\mathbb{N}}}\int_X P_{\omega_1}\phi(x)\psi(\sigma\underline{\omega}, x)\, d\mu(x) d\tilde{\mathbb{P}}(\underline{\omega})$$
$$= \int_X\int_{\Omega^{\mathbb{N}}}\int_{\Omega} P_{\omega_1}\phi(x)\psi((\omega_2,\omega_3,\ldots), x)\, d\mathbb{P}(\omega_1) d\tilde{\mathbb{P}}(\omega_2,\omega_3,\ldots) d\mu(x)$$
$$= \int_X P\phi(x)\int_{\Omega^{\mathbb{N}}}\psi(\underline{\omega}, x)\, d\tilde{\mathbb{P}}(\underline{\omega}) d\mu(x) = \int_{\Omega^{\mathbb{N}}\times X}((P\phi)\circ p)\psi\, d(\tilde{\mathbb{P}}\otimes\mu)$$

□

Proof (Proof of Lemma 4.2) Let $\psi : \Omega^\star \to \mathbb{R}$ be an arbitrary element of $L^\infty(\mu_c)$. We have to show that $\int_{\Omega^\star}\phi_\pi(\psi\circ\tau)\, d\mu_c = \int_{\Omega^\star}(P\phi)_\pi\psi\, d\mu_c$. We have

$$\int_{\Omega^\star}\phi_\pi(\psi\circ\tau)\, d\mu_c = \int_{\Omega^{\mathbb{N}}\times X}(\phi\circ\pi\circ\Phi)(\psi\circ\tau\circ\Phi)\, d(\tilde{\mathbb{P}}\otimes\mu)$$
$$= \int_{\Omega^{\mathbb{N}}\times X}(\phi\circ p)(\psi\circ\Phi\circ S)\, d(\tilde{\mathbb{P}}\otimes\mu) = \int_{\Omega^{\mathbb{N}}\times X} P_S(\phi\circ p)(\psi\circ\Phi)\, d(\tilde{\mathbb{P}}\otimes\mu)$$
$$= \int_{\Omega^{\mathbb{N}}\times X}(P\phi\circ p)(\psi\circ\Phi)\, d(\tilde{\mathbb{P}}\otimes\mu) = \int_{\Omega^{\mathbb{N}}\times X}(P\phi\circ\pi\circ\Phi)(\psi\circ\Phi)\, d(\tilde{\mathbb{P}}\otimes\mu)$$
$$= \int_{\Omega^\star}(P\phi)_\pi\psi\, d\mu_c$$

□

This helps us to construct a martingale approximation for the Birkhoff sums of ϕ_π. We first remark upon the stationary case. By our assumption on decay of correlations, the series $w = \sum_{n=1}^\infty P^n\phi$ is convergent in $L^\infty(\mu)$ if $\phi\in\mathcal{B}$. We define $\chi = \phi_\pi + w_\pi - w_\pi\circ\tau$ on $\Omega^\star$. χ satisfies $\tilde{P}\chi = \tilde{P}(\phi_\pi) + \tilde{P}(w_\pi) - \tilde{P}\tilde{U}w_\pi = (P\phi)_\pi + (Pw)_\pi - w_\pi = 0$, since $Pw = w - P\phi$.

We claim that $\{\chi\circ\tau^k\}_{k\geq 0}$ is a reverse martingale difference with respect to the decreasing filtration $\{\mathcal{F}_k\}_{k\geq 0}$. Indeed, we have $\mathbb{E}_{\mu_c}(\chi\circ\tau^k|\mathcal{F}_{k+1}) = \tilde{U}^{k+1}\tilde{P}^{k+1}\tilde{U}^k\chi = \tilde{U}^{k+1}\tilde{P}\chi = 0$. Uniqueness of the stationary measure also yields that the associated martingale is ergodic, and hence satisfies a CLT (see Billingsley [62]).

Since $\sum_{k=0}^{n-1}\phi_\pi\circ\tau^k = \sum_{k=0}^{n-1}\chi\circ\tau^k + w_\pi\circ\tau^n - w_\pi$, and $\frac{w_\pi\circ\tau^n - w_\pi}{\sqrt{n}}$ goes to zero in probability (because it goes to 0 in the L^1 norm), it follows that $\frac{1}{\sqrt{n}}S_n$ converges to the gaussian law $\mathcal{N}(0,\sigma^2)$ in distribution , where $\sigma^2 = \mathbb{E}_{\mu_c}(\chi^2)$, since $\sum_{j=0}^{n-1}\frac{1}{\sqrt{n}}\chi\circ\tau^j$ converges to $\mathcal{N}(0,\sigma^2)$ in distribution.

5 Dynamical Borel–Cantelli lemmas

In this section, we make the same assumptions as in the previous one. Recall the following result from [63]:

Theorem 5.1 *Let f_k be a sequence of non-negative measurable functions on a measure space (Y, ν), and let $\bar{f}_k$, φ_k be sequences of real numbers such that $0 \leq \bar{f}_k \leq \varphi_k \leq M$ for all $k \geq 1$ and some $M > 0$. Suppose that*

$$\int_Y \left(\sum_{m<k\leq n} \left(f_k(y) - \bar{f}_k\right) \right)^2 d\nu(y) \leq C \sum_{m<k\leq n} \varphi_k$$

for arbitrary integers $m < n$ and some $C > 0$. Then

$$\sum_{1\leq k\leq n} f_k(y) = \sum_{1\leq k\leq n} \bar{f}_k + O(\Phi^{1/2}(n) \log^{3/2+\epsilon} \Phi(n))$$

for ν-a.e. $y \in Y$, for all $\epsilon > 0$ and $\Phi(n) = \sum_{1\leq k\leq n} \varphi_k$.

Applying this result to the probability space $(\tilde{\Omega} \times X, \tilde{\mathbb{P}} \otimes \mu)$, and using decay of correlations, we get:

Proposition 5.2 *If ϕ_n is a sequence of non-negative functions in $\mathcal{B}$, with $\sup_n \|\phi_n\| < \infty$ and $E_n \to \infty$, where $E_n = \sum_{j=0}^{n-1} \int \phi_n \, d\mu$, then*

$$\lim_{n\to\infty} \frac{1}{E_n} \sum_{j=0}^{n-1} \phi_j(S^j(x, \omega)) \to 1$$

for $\tilde{\mathbb{P}} \otimes \mu$ a.e. $(\omega, x) \in \tilde{\Omega} \times X$.

See theorem 2.1 in Kim [9] for a completely analogue proof in a deterministic setting. The annealed version of the Strong Borel–Cantelli property clearly implies a quenched version, namely for $\tilde{\mathbb{P}}$-a.e. ω for μ-a.e. $x \in X$,

$$\lim_{n\to\infty} \frac{1}{E_n} \sum_{j=0}^{n-1} \phi_j(S^j(x, \omega)) \to 1$$

We now show how to prove a CLT, following our martingale approach described in the previous section.

CLT for Borel–Cantelli sequences

Let $p \in X$ and let $B_n(p)$ be a sequence of nested balls about p such that for $0 \leq \gamma_2 \leq \gamma_1 \leq 1$ and constants $C_1, C_2 > 0$ we have $\frac{C_1}{n^{\gamma_1}} \leq \mu(B_n(p)) \leq \frac{C_2}{n^{\gamma_2}}$.

Let $\phi_n = \mathbb{1}_{B_n(p)}$ be the characteristic function of $B_n(p)$. We assume that ϕ_n is a bounded sequences in $\mathcal{B}$, which is clearly the case when $\mathcal{B}$ is BV or Quasi-Hölder. We will sometimes write $\mathbb{E}[\phi]$ or $\int \phi$ for the integral $\int \phi \, d\mu$ when the context is understood.

First we lift ϕ_i to Ω^* and define $(\phi_i)_\pi = \phi_i \circ \pi$ and then we normalize and write $\tilde{\phi}_j = (\phi_j)_\pi - \int (\phi_j)_\pi d\mu_c$.

We are almost in the setting of [36, Proposition 5.1] which states,

Proposition 5.3 *Suppose (T, X, μ) is an ergodic transformation whose transfer operator P satisfies, for some constants $C > 0, 0 < \theta < 1$,*

$$\|P^n \phi\|_{\mathcal{B}} \leq C\theta^n \|\phi\|_{\mathcal{B}}$$

for all ϕ such that $\phi \, d\mu = 0$. Let $B_i := B_i(p)$ be nested balls about a point p such that for constants $0 \leq \gamma_2 \leq \gamma_1 \leq 1$, $C_1 > 0$, $C_2 > 0$ we have $\frac{C_1}{n^{\gamma_1}} \leq \mu(B_n(p)) \leq \frac{C_2}{n^{\gamma_2}}$. Let

$$a_n^2 := E\left(\sum_{j=1}^{n} (1_{B_i} \circ T^i - \mu(B_i))\right)^2$$

Then

$$\liminf \frac{a_n^2}{E_n} \geq 1$$

and

$$\frac{1}{a_n} \sum_{j=1}^{n} \left(\phi_j - \int \phi_j \, d\mu\right) \circ T^j \to \mathcal{N}(0, 1)$$

As a fairly direct corollary we may show in our setting:

Corollary 5.4

$$\frac{1}{a_n} \sum_{j=1}^{n} \tilde{\phi}_j \circ \tau^j \to \mathcal{N}(0, 1)$$

and hence

$$\frac{1}{a_n} \sum_{j=1}^{n} \left(\phi_j - \int \phi_j \, d\mu\right) \circ S^j \to \mathcal{N}(0, 1)$$

Proof Define $\phi_0 = 1$ and for $n \geq 1$

$$w_n = P\phi_{n-1} + P^2\phi_{n-2} + \cdots P^n\phi_0$$

so that $w_1 = P\phi_0$, $w_2 = P\phi_1 + P^2\phi_0$, $w_3 = P\phi_2 + P^2\phi_1 + P^3\phi_0$ etc... For $n \geq 1$ define

$$\psi_n = (\phi_n)_\pi - (w_{n+1})_\pi \circ \tau + (w_n)_\pi$$

An easy calculation shows that $\tilde{P}\psi_n = 0$ and hence $X_{ni} := \psi_i \circ \tau^i/(a_n)$ is a reverse martingale difference array with respect to the filtration $\mathcal{F}_i$.

We have exponential decay of correlations in the sense that if $j > i$ then

$$\left|\int \tilde{\phi}_i \circ \tau^i \tilde{\phi}_j \circ \tau_j d\mu_c\right| = \left|\int \tilde{\phi}_i \tilde{\phi}_j \circ \tau^{j-i} d\mu_c\right|$$
$$\leq C\theta^{j-i} \left\|\phi_j - \int \phi_j \, d\mu\right\|_{\mathcal{B}} \|\tilde{\phi}_j\|_1$$

where $\|\phi_j - \int \phi_j \, d\mu\|_{\mathcal{B}}$ is bounded uniformly over j.

The proof of [36, Proposition 5.1] may now be followed exactly to establish conditions (a), (b), (c) and (d) of Theorem 3.2 from Hall and Heyde [35] as well as show that the variance a_n is unbounded. □

6 Erdös–Rényi laws

Erdös–Rényi limit laws give information on the maximal average gain precisely in the case where the length of the time window ensures there is a non-degenerate limit. Recall the following proposition from [7]:

Proposition 6.1 *Let (X, T, μ) be a probability preserving transformation, and $\varphi : X \to \mathbb{R}$ be a mean-zero μ-integrable function. Let $S_n(\varphi) = \varphi + \cdots + \varphi \circ T^{n-1}$.*

1. *Suppose that φ satisfies a large deviation principle with rate function I defined on the open set U. Let $\alpha > 0$ and set*

$$l_n = l_n(\alpha) = \left[\frac{\log n}{I(\alpha)}\right] \quad n \in \mathbb{N}.$$

Then the upper Erdös–Rényi law holds, that is, for μ a.e. $x \in X$

$$\limsup_{n\to\infty} \max\{S_{l_n}(\varphi) \circ T^j(x)/l_n : 0 \leq j \leq n - l_n\} \leq \alpha.$$

2. *Suppose that for every $\epsilon > 0$ the series $\sum_{n>0} \mu(B_n(\epsilon))$, where $B_n(\epsilon) = \{\max_{0\leq m\leq n-l_n} S_{l_n} \circ T^m \leq l_n(\alpha - \epsilon)\}$ is summable.*
Then the lower Erdös–Rényi law holds, that is, for μ a.e. $x \in X$

$$\liminf_{n\to\infty} \max\{S_{l_n}(\varphi) \circ T^j(x)/l_n : 0 \leq j \leq n - l_n\} \geq \alpha.$$

Remark 6.2 Assumptions (a) and (b) of Proposition 6.1 together imply that

$$\lim_{n\to\infty} \max_{0\le m\le n-l_n} \frac{S_{l_n}\circ T^m}{l_n} = \alpha.$$

In this section, we will suppose that $X = [0, 1]$, and that the Banach space $\mathcal{B}$ is BV, the space of functions of bounded variation on $[0, 1]$. All maps T_ω are piecewise C^2, and we assume an uniform upper bound $L > 0$ for their derivatives. We will apply the previous proposition to the symbolic system $(\Omega^\star, \tau, \mu_c)$ introduced before.

Theorem 6.3 *Suppose $\phi : X \to \mathbb{R}$ is of bounded variation with $\int_X \phi\, d\mu = 0$ and define $S_n = \sum_{j=0}^{n-1} \phi_\pi \circ \tau^j$. Let $\alpha > 0$ and set*

$$l_n = l_n(\alpha) = \left[\frac{\log n}{I(\alpha)}\right] \quad n \in \mathbb{N}$$

where $I(.)$ is the rate function associated to ϕ, which exists by Theorem 3.6. Then

$$\lim_{n\to\infty} \max_{0\le m\le n-l_n} \frac{S_{l_n}\circ T^m}{l_n} = \alpha.$$

Proof Since ϕ satisfies an annealed LDP, which can be immediately lifted to a LDP for ϕ_π, we need only prove (2). As in the section on the logistic map in [7] we use a blocking argument to establish (2).

For all $s > 0$, define $A_n^s = \{S_{l_n} \le l_n(\alpha - s)\}$. We fix $\epsilon > 0$, and consider A_n^ϵ and $A_n^{\epsilon/2}$. Let $0 < \eta < 1$. We define φ_ϵ to be a Lipschitz function with Lipschitz norm at most $L^{(1+\eta)l_n}$ satisfying $\mathbb{1}_{A_n^\epsilon} \le \varphi_\epsilon \le 1$ and $\mu(A_n^\epsilon) < \int_X \varphi_\epsilon\, d\mu < \mu(A_n^{\epsilon/2})$, in the same way as in the proof of Theorem 3.1 in [7].

Define $C_m(\epsilon) = \{S_{l_n} \circ \tau^m \le l_n(\alpha - \epsilon)\}$ and $B_n(\epsilon) = \bigcap_{m=0}^{n-l_n} C_m(\epsilon)$. We use a blocking argument to take advantage of decay of correlations and intercalate by blocks of length $(\log n)^\kappa$, $\kappa > 1$. We define

$$E_n^0(\epsilon) := \bigcap_{m=0}^{[(n-(\log n)^\kappa)/(\log n)^\kappa)]} C_{m[(\log n)^\kappa]}(\epsilon)$$

and in general for $0 \le j < [\frac{n}{(\log n)^\kappa}]$

$$E_n^j(\epsilon) := \bigcap_{m=0}^{[(n-(j+1)(\log n)^\kappa)/(\log n)^\kappa)]} C_{m[(\log n)^\kappa]}(\epsilon).$$

Note that $\mu(B_n(\epsilon)) \le \mu(E_n^0(\epsilon))$. For each j, let $\psi_j = \mathbb{1}_{E_n^j(\epsilon)}$ denote the characteristic function of $E_n^j(\epsilon)$.

By decay of correlations we have

$$\begin{aligned}\mu(E_n^0(\epsilon)) &\le \int \varphi_\epsilon \cdot \psi_1 \circ \tau^{[(\log n)^\kappa]} d\mu_c \\ &\le C\theta^{(\log n)^\kappa} \|\varphi_\epsilon\|_{BV} \|\psi_1\|_1 + \int \varphi_\epsilon \, d\mu_c \int \psi_1 \, d\mu_c \\ &\le \int \varphi_\epsilon \, d\mu_c \int \psi_1 \, d\mu_c + C\theta^{(\log n)^\kappa} (L^{(1+\eta)l_n}).\end{aligned}$$

Applying decay of correlations again to $\int \psi_1 \, d\mu_c$ we iterate and conclude

$$\mu(E_n^0(\epsilon)) \le nC\theta^{(\log n)^\kappa} L^{(1+\eta)l_n} + \mu(A_n^{\epsilon/2})^{n/(\log n)^\kappa}.$$

The term $nC\theta^{(\log n)^\kappa} L^{(1+\eta)l_n}$ is clearly summable since $\kappa > 1$. □

Remark 6.4 We would obtain a quenched Erdös–Rényi law as well, if we could establish exponential decay of correlations for $\tilde{\mathbb{P}}$ almost every $\underline{\omega}$, together with a quenched LDP for functions of bounded variation.

7 Quenched CLT for random one dimensional systems

In this section, we consider the quenched CLT, that is a CLT holding for almost every fixed sequence $\underline{\omega}$. We first state a general result, which is basically a consequence of [31]. Let $\{T_\omega\}_{\omega\in\Omega}$ be a iid random dynamical system acting on X, with a stationary measure μ. Let $\varphi : X \to \mathbb{R}$ be an observable with $\int_X \varphi d\mu = 0$, and define as before $X_k(\underline{\omega}, x) = \varphi(T_{\underline{\omega}}^k x)$ and $S_n = \sum_{k=0}^{n-1} X_k$. We will need to introduce a auxiliary random system defined as follows : the underlying probability space is still $(\Omega, \mathbb{P})$, while the auxiliary system acts on X^2, with associated maps $\hat{T}_\omega$ given by $\hat{T}_\omega(x, y) = (T_\omega x, T_\omega y)$. Define then a new observable $\hat{\varphi} : X^2 \to \mathbb{R}$ by $\hat{\varphi}(x, y) = \varphi(x) - \varphi(y)$, and denote its associated Birkhoff sums by $\hat{S}_n$.

Theorem 7.1 *Assume there exists $\sigma^2 > 0$ and a constant $C > 0$ such that for all $t \in \mathbb{R}$ and $n \ge 1$ with $\frac{t}{\sqrt{n}}$ small enough:*

(1) $|\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(e^{i\frac{t}{\sqrt{n}}S_n}) - e^{-\frac{t^2\sigma^2}{2}}| \le C\frac{1+|t|^3}{\sqrt{n}}$,

(2) $|\mathbb{E}_{\tilde{\mathbb{P}}\otimes(\mu\otimes\mu)}(e^{i\frac{t}{\sqrt{n}}\hat{S}_n}) - e^{-t^2\sigma^2}| \le C\frac{1+|t|^3}{\sqrt{n}}$.

Suppose also that for $n \ge 1$ and $\epsilon > 0$:

(3) $\tilde{\mathbb{P}} \otimes \mu(|\frac{S_n}{n}| \ge \epsilon) \le Ce^{-C\epsilon^2 n}$.

Then, the quenched CLT holds: for $\tilde{\mathbb{P}}$-a.e. sequence $\underline{\omega} \in \Omega^{\mathbb{N}}$ one has

$$\frac{\sum_{k=0}^{n-1} \varphi \circ T_{\underline{\omega}}^k}{\sqrt{n}} \Longrightarrow_\mu \mathcal{N}(0, \sigma^2).$$

The first and third assumptions can be proved using the spectral approach described in this paper. Indeed the first one corresponds to Lemma 3.10, while the third follows from the LDP. To prove the second assumption, one must employ again the spectral technique, but with the auxiliary system introduced above and the observable $\hat{\varphi}$. There are mainly two difficulties: the obvious one is that the auxiliary system acts on a space whose dimension is twice the dimension of X, so that we have to use more complicated functional spaces. The other difficulty, less apparent, is that the asymptotic variance of $\hat{\varphi}$ has to be twice the asymptotic variance of φ. We do not see any reason for this to be true in full generality. In the particular case where all maps T_ω preserve the measure μ, this can be proved using Green–Kubo formula from Proposition 3.2: assuming that the auxiliary system is mixing and has a spectral gap on an appropriated Banach space, the stationary measure is then given by $\mu \otimes \mu$ (since it is preserved by all maps $\hat{T}_\omega$), and an algebraic manipulation using Proposition 3.2 shows that the asymptotic variance of $\hat{\varphi}$ is given by $2\sigma^2$. See [31] for a similar computation.

In the general situation, the stationary measure of the auxiliary measure can be different from $\mu \otimes \mu$, and it seems hard to compute the asymptotic variance of $\hat{\varphi}$ from Green–Kubo formula. Even though this condition can seem unnatural, it is also necessary in order for the quenched central limit theorem to be true in the form we have stated it, as can be seen from the proof of [31]. We state this as a lemma:

Lemma 7.2 *Using the same notations introduced above, assume that there exists $\sigma^2 > 0$ and $\hat{\sigma}^2 > 0$ such that*

1. $\frac{S_n}{\sqrt{n}}$ *converges in law to $\mathcal{N}(0, \sigma^2)$ under the probability $\tilde{\mathbb{P}} \otimes \mu$,*
2. $\frac{\hat{S}_n}{\sqrt{n}}$ *converges in law to $\mathcal{N}(0, \hat{\sigma}^2)$ under the probability $\tilde{\mathbb{P}} \otimes (\mu \otimes \mu)$,*
3. *for a.e. $\underline{\omega}$, $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \varphi \circ T_{\underline{\omega}}^k$ converges in law to $\mathcal{N}(0, \sigma^2)$ under the probability μ.*

Then $\hat{\sigma}^2 = 2\sigma^2$.

Proof Define $S_{n,\underline{\omega}} = S_n(\underline{\omega}, .) = \sum_{k=0}^{n-1} \varphi \circ T_{\underline{\omega}}^k$. Following [31], we write for any $t \in \mathbb{R}$ and $n \geq 1$:

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbb{P}}}(|\mu(e^{i\frac{t}{\sqrt{n}}S_{n,\underline{\omega}}}) - e^{-\frac{\sigma^2 t^2}{2}}|^2) &= \mathbb{E}_{\tilde{\mathbb{P}}}(|\mu(e^{i\frac{t}{\sqrt{n}}S_n})|^2) - e^{-t^2\sigma^2} \\
&\quad + 2e^{-\frac{\sigma^2 t^2}{2}} \Re(e^{-\frac{\sigma^2 t^2}{2}} - \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(e^{i\frac{t}{\sqrt{n}}S_n})) \\
&= \mathbb{E}_{\tilde{\mathbb{P}}}(\mu \otimes \mu(e^{i\frac{t}{\sqrt{n}}\hat{S}_n})) - e^{-\frac{t^2\hat{\sigma}^2}{2}} + (e^{-\frac{t^2\hat{\sigma}^2}{2}} - e^{-t^2\sigma^2}) \\
&\quad + 2e^{-\frac{\sigma^2 t^2}{2}} \Re(e^{-\frac{\sigma^2 t^2}{2}} - \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(e^{i\frac{t}{\sqrt{n}}S_n})).
\end{aligned}$$

By the two first assumptions, this term goes to $e^{-\frac{t^2\hat{\sigma}^2}{2}} - e^{-t^2\sigma^2}$ as n goes to infinity. But $\mathbb{E}_{\tilde{\mathbb{P}}}(|\mu(e^{i\frac{t}{\sqrt{n}}S_{n,\underline{\omega}}}) - e^{-\frac{\sigma^2 t^2}{2}}|^2)$ goes to 0 thanks to the third assumption and the dominated convergence theorem. This shows that $e^{-\frac{t^2\hat{\sigma}^2}{2}} = e^{-t^2\sigma^2}$. □

In the following, we will consider the situation where $X = [0, 1]$ and all maps preserve the Lebesgue measure m. For technical convenience, we will also assume that Ω is a finite set.

Example 7.3 Suppose that all maps T_ω are given by $T_\omega x = \beta_\omega x \bmod 1$, where $\beta_\omega > 1$ is an integer. The transfer operator of this system clearly satisfies a Lasota–Yorke on the space BV, and is random-covering, so that assumption (1) and (3) follows automatically for any $\varphi \in \mathrm{BV}$. On the other hand, the auxiliary two-dimensional system has a spectral gap on the quasi-Hölder space $V_1(X^2)$ and is also random covering. Since $\hat{\varphi}$ belongs to $V_1(X^2)$, assumption (2) follows by the above discussion and the quenched CLT holds.

Example 7.4 There exist piecewise non-linear expanding maps which preserves Lebesgue. Such a class of examples is provided by the Lorenz-like maps considered in the paper [39]: these maps have both a neutral parabolic fixed point and a point where the derivative goes to infinity. The coexistence of these two behaviors allows the possibility for the map to preserve Lebesgue measure while being non-linear. Suppose that $\Omega = \{0, 1\}$, T_0 is the doubling map $T_0 x = 2x \bmod 1$, and that T_1 is one of the maps considered in [39]. We will prove that there exists $0 \leq p^\star < 1$ such that if T_0 is iterated with probability p with $p > p_\star$, then the quenched CLT holds for any observable φ Lipschitz.

Since the annealed transfer operator P can be written as $pP_0 + (1-p)P_1$, where P_0, resp. P_1, is the transfer operator of T_0, resp. T_1 (and similarly $\hat{P} = p\hat{P}_0 + (1-p)\hat{P}_1$ for the auxiliary system), it is sufficient to prove that P_0 and $\hat{P}_0$ have a spectral gap on Banach spaces $\mathcal{B}$ and $\hat{\mathcal{B}}$, while P_1 and $\hat{P}_1$ act continuously on these spaces, and that $\varphi \in \mathcal{B}$ and $\hat{\varphi} \in \hat{\mathcal{B}}$. We will use quasi-Hölder spaces and will take $\mathcal{B} = V_\alpha(X)$ and $\hat{\mathcal{B}} = V_\alpha(X^2)$ for a convenient choice of α. Clearly, the transfer operator of T_0 and $\hat{T}_0$ have a spectral gap on these spaces, and φ (resp. $\hat{\varphi}$) belongs to $\mathcal{B}$ (resp. $\hat{\mathcal{B}}$) whenever φ is Lipschitz. To prove the continuity of P_1 and $\hat{P}_1$, we will use the following general result.

Proposition 7.5 *Let M be a compact subset of $\mathbb{R}^d$ with $m_d(M) = 1$, where m_d denotes the Lebesgue measure on $\mathbb{R}^d$, and $T : M \to M$ be a non-singular map. Define $g(x) = \frac{1}{|\det DT(x)|}$, and assume there exist a finite family of disjoint open set $\{U_i\}_i$ included in M, a constant $C > 0$ and $0 < \alpha \leq 1$ with*

1. *$m_d(\cup_i U_i) = 1$,*
2. *$T : U_i \to TU_i$ is a C^1-diffeomorphism,*
3. *$d(Tx, Ty) \geq d(x, y)$ for all i and all $x, y \in U_i$,*
4. *$|g(x) - g(y)| \leq Cd(x, y)^\alpha$, for all i and all $x, y \in U_i$,*
5. *$m_d(B_\epsilon(\partial TU_i)) \leq C\epsilon^\alpha$ for all i and all $\epsilon > 0$.*

Then the transfer operator of T acts continuously on $V_\alpha(M)$.

The map with parameter $\gamma > 1$ considered in [39] satisfies these assumptions for $\alpha = \min\{1, \gamma - 1\}$, so that the quenched CLT holds when $p^\star$ is close enough to 1.

Proof (Proof of Proposition 7.5) We denote by $T_i^{-1} : TU_i \to U_i$ the inverse branch of T restricted to U_i.

The transfer operator P of T reads as

$$Pf(x) = \sum_i (gf) \circ T_i^{-1} \mathbb{1}_{TU_i}(x).$$

Following Saussol [2], we have for all $\epsilon > 0$ and $x \in \mathbb{R}^d$:

$$\operatorname{osc}(Pf, B_\epsilon(x)) \leq \sum_i R_i^{(1)}(x) \mathbb{1}_{TU_i}(x) + 2 \sum_i R_i^{(2)}(x),$$

where $R_i^{(1)}(x) = \operatorname{osc}(gf, T^{-1}B_\epsilon(x) \cap U_i)$ and $R_i^{(2)}(x) = (\operatorname{ess\,sup}_{T^{-1}B_\epsilon(x)\cap U_i} |gf|)$ $\mathbb{1}_{B_\epsilon(\partial TU_i)}(x)$.

Using Proposition 3.2 (iii) in [2], we get

$$\begin{aligned} R_i^{(1)}(x) \leq{}& \operatorname{osc}(f, T^{-1}B_\epsilon(x) \cap U_i) \operatorname*{ess\,sup}_{T^{-1}B_\epsilon(x)\cap U_i} g \\ &+ \operatorname{osc}(g, T^{-1}B_\epsilon(x) \cap U_i) \operatorname*{ess\,inf}_{T^{-1}B_\epsilon(x)\cap U_i} |f|. \end{aligned}$$

By assumption (3), we have $T^{-1}B_\epsilon(x) \cap U_i \subset B_\epsilon(T_i^{-1}x)$, while by assumption (4), $\operatorname{osc}(g, T^{-1}B_\epsilon(x) \cap U_i) \leq C\epsilon^\alpha$ and $\operatorname{ess\,sup}_{T^{-1}B_\epsilon(x)\cap U_i} g \leq g(T_i^{-1}x) + C\epsilon^\alpha$.

This shows $R_i^{(1)}(x) \leq g(T_i^{-1}x)\operatorname{osc}(f, B_\epsilon(T_i^{-1}x) + C\epsilon^\alpha \|f\|_{\sup}$, whence

$$\int \sum_i R_i^{(1)}(x) \mathbb{1}_{TU_i}(x) dx \leq \int P(\operatorname{osc}(f, B_\epsilon(.))(x) dx + C\|f\|_\infty \epsilon^\alpha \sum_i m_d(TU_i).$$

Since the sum is finite, this gives $\int \sum_i R_i^{(1)}(x) \mathbb{1}_{TU_i}(x) dx \leq \epsilon^\alpha(|f|_\alpha + C\|f\|_{\sup}) \leq C\epsilon^\alpha \|f\|_\alpha$.

We turn now to the estimate of $R_i^{(2)}$: one has $R_i^{(2)}(x) \leq \|g\|_{\sup}\|f\|_{\sup} \mathbb{1}_{B_\epsilon(\partial TU_i)}(x)$, so that using assumption (4), $\int \sum_i R_i^{(2)} dx \leq C\|f\|_{\sup} \sum_i m_d(B_\epsilon(\partial TU_i)) \leq C\epsilon^\alpha \|f\|_{\sup}$. This shows that $|Pf|_\alpha \leq C\|f\|_\alpha$ and concludes the proof. □

8 Concentration inequalities

A function $K : X^n \to \mathbb{R}$, where (X, d) is a metric space, is separately Lipschitz if, for all i, there exists a constant $\operatorname{Lip}_i(K)$ with

$$\begin{aligned} &|K(x_0, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{n-1}) - K(x_0, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_{n-1})| \\ &\quad \leq \operatorname{Lip}_i(K) d(x_i, x_i') \end{aligned}$$

for all points $x_0, \ldots, x_{n-1}, x_i'$ in X.

Let $(\Omega, \mathbb{P}, T)$ be a finite random Lasota–Yorke system on the unit interval $X = [0, 1]$, such that $\lambda(T_\omega) > 1$ for all $\omega \in \Omega$. We assume that $(\Omega, \mathbb{P}, T)$ satisfies the

random covering property, and we denote by μ its unique absolutely continuous stationary measure. Its density h belongs to BV, and is uniformly bounded away from 0.

Theorem 8.1 *There exists a constant $C \geq 0$, depending only on $(\Omega, \mathbb{P}, T)$, such that for any $n \geq 1$ and any separately Lipschitz function $K : X^n \to \mathbb{R}$, one has*

$$\mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(e^{K(x,T^1_{\underline{\omega}}x,\ldots,T^{n-1}_{\underline{\omega}}x)-\mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(K(x,T^1_{\underline{\omega}}x,\ldots,T^{n-1}_{\underline{\omega}}x))}) \leq e^{C\sum_{i=0}^{n-1}\mathrm{Lip}_i^2(K)}$$

This leads to a large deviation estimate, namely that for all $t > 0$, one has

$$\mu \otimes \tilde{\mathbb{P}}(\{(x, \underline{\omega}) \,/\, K(x, T^1_{\underline{\omega}}x, \ldots, T^{n-1}_{\underline{\omega}}x) - m > t\}) \leq e^{-\frac{t^2}{4C\sum_{i=0}^{n-1}\mathrm{Lip}_i^2(K)}},$$

where $m = \mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(K(x, T^1_{\underline{\omega}}x, \ldots, T^{n-1}_{\underline{\omega}}x))$.

For the proof, we will use McDiarmid's bounded differences method [64,65], as in [10,11], conveniently adapted to the random context.

We will denote by P the annealed transfer operator with respect to the Lebesgue measure m, and by L the annealed transfer operator with respect to the stationary measure μ. Recall that L acts on functions in $L^1(\mu)$ by $L(f) = \frac{P(fh)}{h}$, whence

$$Lf(x) = \sum_{\omega\in\Omega} p_\omega \sum_{T_\omega y = x} \frac{h(y)f(y)}{h(x)|T'_\omega(y)|}.$$

Since h belongs to BV, together with $\frac{1}{h}$, L acts on BV and has a spectral gap.

Recall the construction of the symbolic system $(X^{\mathbb{N}}, \mathcal{F}, \sigma, \mu_c)$ and of the decreasing filtration $\{\mathcal{F}_p\}_{p\geq 0}$ of σ-algebras. We extend K as a function on $X^{\mathbb{N}}$, depending only on the n first coordinates. One has obviously $\mathbb{E}_{\mu_c}(K) = \mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(K(x, T^1_{\underline{\omega}}x, \ldots, T^{n-1}_{\underline{\omega}}x))$ and

$$\mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(e^{K(x,T^1_{\underline{\omega}}x,\ldots,T^{n-1}_{\underline{\omega}}x)-\mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(K(x,T^1_{\underline{\omega}}x,\ldots,T^{n-1}_{\underline{\omega}}x))}) = \mathbb{E}_{\mu_c}(e^{K-\mathbb{E}_{\mu_c}(K)}),$$

since $\Phi : X \times \tilde{\Omega} \to X^{\mathbb{N}}$ is a factor map. We define $K_p = \mathbb{E}_{\mu_c}(K|\mathcal{F}_p)$, and $D_p = K_p - K_{p+1}$. One has the following:

Lemma 8.2 *The dynamical system $(X^{\mathbb{N}}, \mathcal{F}, \sigma, \mu_c)$ is exact.*

Proof This follows from exactness of the skew-product system $(X \times \tilde{\Omega}, S, \mu \otimes \tilde{\mathbb{P}})$, see theorem 5.1 in [20], and the fact that $\Phi : X \times \tilde{\Omega} \to X^{\mathbb{N}}$ is a factor map. See also theorem 4.1 in [66]. □

This implies that $\mathcal{F}_\infty := \bigcap_{p\geq 0} \mathcal{F}_p = \bigcap_{p\geq 0} \sigma^{-p}\mathcal{F}$ is μ_c-trivial, from which we deduce, by Doob's convergence theorem, that K_p goes to $\mathbb{E}_{\mu_c}(K)\mu_c$-as when p goes to infinity, whence $K - \mathbb{E}_{\mu_c}(K) = \sum_{p\geq 0} D_p$.

From Azuma–Hoeffding's inequality (see lemma 4.1 in [64] and its proof for the bound of the exponential moment), we deduce that there exists some $C \geq 0$ such that for all $P \geq 0$,

$$\mathbb{E}_{\mu_c}(e^{\sum_{p=0}^{P} D_p}) \leq e^{C \sum_{p=0}^{P} \sup |D_p|^2}.$$

It remains to bound D_p:

Proposition 8.3 *There exists $\rho < 1$ and $C \geq 0$, depending only on $(\Omega, \mathbb{P}, T)$, such that for all p, one has*

$$|D_p| \leq C \sum_{j=0}^{p} \rho^{p-j} \mathrm{Lip}_i(K).$$

This proposition, together with the Cauchy–Schwarz inequality, implies immediately the desired concentration inequality, in the same manner as in [10]. The following lemma leads immediately to the result, using the Lipschitz condition on K:

Lemma 8.4 *There exists $\rho < 1$ and $C \geq 0$, depending only on $(\Omega, \mathbb{P}, T)$, such that for all p and $x_p, \ldots$, one has*

$$\left| K_p(x_p, \ldots) - \int_{\tilde{\Omega}} \int_X K(y, T^1_{\underline{\omega}} y, \ldots, T^{p-1}_{\underline{\omega}} y, x_p, \ldots) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega}) \right|$$
$$\leq C \sum_{j=0}^{p-1} \mathrm{Lip}_j(K) \rho^{p-j}.$$

The rest of this section is devoted to the proof of this lemma. For a sequence $\underline{\omega} \in \tilde{\Omega}$, we denote $g^{(p)}_{\underline{\omega}}(y) = \frac{h(y)}{h(T^p_{\underline{\omega}} y)} \frac{1}{|(T^p_{\underline{\omega}})'(y)|}$. We have

$$K_p(x_p, \ldots) = \sum_{\underline{\omega} \in \Omega^p} p^p_{\underline{\omega}} \sum_{T^p_{\underline{\omega}} y = x} g^{(p)}_{\underline{\omega}}(y) K(y, T^1_{\underline{\omega}} y, \ldots, T^{p-1}_{\underline{\omega}} y, x_p \ldots).$$

We fix a $x_\star \in X$, and we decompose K_p as

$$K_p(x_p, \ldots) = K(x_\star, \ldots, x_\star, x_p, \ldots) + \sum_{i=0}^{p-1} \sum_{\underline{\omega} \in \Omega^p} p^p_{\underline{\omega}} \sum_{T^p_{\underline{\omega}} y = x} g^{(p)}_{\underline{\omega}}(y) H_i(y, \ldots, T^i_{\underline{\omega}} y),$$

where $H_i(y_0, \ldots, y_i) = K(y_0, \ldots, y_i, x_\star, \ldots, x_\star, x_p, \ldots) - K(y_0, \ldots, y_{i-1}, x_\star, \ldots, x_\star, x_p, \ldots)$.

A simple computation then shows that $K_p(x_p, \ldots) = K(x_\star, \ldots, x_\star, x_p, \ldots) + \sum_{i=0}^{p-1} L^{p-i} f_i(x_p)$, with

$$f_i(y) = \sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}} \sum_{T^i_{\underline{\omega}}z=y} g^{(i)}_{\underline{\omega}}(z) H_i(z, \ldots, T^i_{\underline{\omega}}z).$$

From the spectral gap of L, we deduce that there exists $C \geq 0$ and $\rho < 1$ depending only on the system, such that $\|L^{p-i} f_i - \int_X f_i d\mu\|_{\mathrm{BV}} \leq C\rho^{p-i}\|f_i\|_{\mathrm{BV}}$. On one hand, since the BV-norm dominates the supremum norm, one has $|L^{p-i} f_i(x_p) - \int_X f_i d\mu| \leq C\rho^{p-i}\|f_i\|_{\mathrm{BV}}$. On the other hand, one has easily $\int_X f_i d\mu = \int_{\tilde{\Omega}} \int_X H_i(y, \ldots, T^i_{\underline{\omega}}y) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega})$, from which it follows, summing all the relations, that

$$\left| K_p(x_p, \ldots) - \int_{\tilde{\Omega}} \int_X K(y, T^1_{\underline{\omega}}y, \ldots, T^{n-1}_{\underline{\omega}}y, x_p, \ldots) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega}) \right|$$
$$\leq C \sum_{i=0}^{p-1} \rho^{p-i} \|f_i\|_{\mathrm{BV}}.$$

It remains to estimate $\|f_i\|_{\mathrm{BV}} \leq \|f_i\|_{\sup} + \mathrm{Var}(f_i)$. For this, we'll need a technical lemma. For $\omega \in \Omega$, we denote by $\mathcal{A}_\omega$ the partition of monotonicity of T_ω, and for $\underline{\omega} \in \Omega^{\mathbb{N}}$, we define $\mathcal{A}^{n-1}_{\underline{\omega}} = \bigvee_{k=0}^{n-1} (T^k_{\underline{\omega}})^{-1}(\mathcal{A}_{\omega_{k+1}})$, which is the partition of monotonicity of $T^n_{\underline{\omega}}$.

If T is a Lasota–Yorke map of the interval, with partition of monotonicity $\mathcal{A}$, we define its distortion $\mathrm{Dist}(T)$ as the least constant C such that $|T'(x) - T'(y)| \leq C|T'(x)||Tx - Ty|$ for all $x, y \in I$ and $I \in \mathcal{A}$.

Lemma 8.5 *There exists $\lambda > 1$ and $C \geq 0$ so that, for all $\underline{\omega} \in \Omega^{\mathbb{N}}$ and $n \geq 0$:*

1. $\lambda(T^n_{\underline{\omega}}) \geq \lambda^n$,
2. $\mathrm{Dist}(T^n_{\underline{\omega}}) \leq C$,
3. $\sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} \sum_{I\in\mathcal{A}^{n-1}_{\underline{\omega}}} \sup_I \frac{1}{|(T^n_{\underline{\omega}})'|} \leq C$,
4. $\sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} \sum_{I\in\mathcal{A}^{n-1}_{\underline{\omega}}} \mathrm{Var}_I(\frac{1}{|(T^n_{\underline{\omega}})'|}) \leq C$.

Proof 1. is obvious, since Ω is a finite set.

2. This is a classical computation. It follows from (1) and the chain rule.
3. This an easy adaptation of lemma II.4 in [11]. For any $\underline{\omega} \in \Omega^n$, and $I \in \mathcal{A}^{n-1}_{\underline{\omega}}$, there exists a least integer $p = p_{\underline{\omega},I}$ such that $T^p_{\underline{\omega}}(I) \cap \partial\mathcal{A}_{\omega_{p+1}} \neq \emptyset$. We denote by $\mathcal{A}^{n-1,p}_{\underline{\omega}}$ the set of all $I \in \mathcal{A}^{n-1}_{\underline{\omega}}$ for which we have $p = p_{\underline{\omega},I}$. We define $\partial = \cup_{\omega\in\Omega}\partial\mathcal{A}_\omega$. Fix $I \in \mathcal{A}^{n-1,p}_{\underline{\omega}}$. There exists $a \in \partial I$ such that $b = T^p_{\underline{\omega}}a \in \partial$. From (2), we deduce the existence of a constant C, depending only on the system, such that, for any $x \in I$,

$$|(T^n_{\underline{\omega}})'(x)| \geq C|(T^n_{\underline{\omega}})'(a)| = C|(T_{\omega_n} \circ \cdots \circ T_{\omega_{p+1}})'(b)||(T^p_{\underline{\omega}})'(a)|$$
$$\geq C\lambda^{n-p}|(T^p_{\underline{\omega}})'(a)|.$$

One has then $\sup_I \frac{1}{|(T^n_{\underline{\omega}})'|} \le C^{-1}\lambda^{-(n-p)}\frac{1}{|(T^p_{\underline{\omega}})'(a)|}$. Since a pre-image by $T^p_{\underline{\omega}}$ of an element $b \in \partial$ can only belong to at most two different $I \in \mathcal{A}^{n-1}_{\underline{\omega}}$, it follows

$$\sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} \sum_{I\in\mathcal{A}^{n-1}_{\underline{\omega}}} \sup_I 8\frac{1}{|(T^n_{\underline{\omega}})'|} \le 2C^{-1}\sum_{p=0}^{n-1}\lambda^{-(n-p)}\sum_{b\in\partial}\sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} \sum_{T^p_{\underline{\omega}}a=b}\frac{1}{|(T^p_{\underline{\omega}})'(a)|}$$
$$= 2C^{-1}\sum_{p=0}^{n-1}\lambda^{-(n-p)}\sum_{b\in\partial} P^p\mathbb{1}(b).$$

This quantity is bounded, since P is power bounded, and ∂ is a finite set.

4. It follows from the three previous points, and the definition of the total variation. □

Since $L^i\mathbb{1} = \mathbb{1}$, one has $\|f_i\|_{\sup} \le \|H_i\|_{\sup} \le \mathrm{Lip}_i(K)$. The crucial point lies in the estimate of the variation of f_i. We first note that

$$\mathrm{Var}(f_i) \le \mathrm{Var}\left(\frac{1}{h}\right)\|hf_i\|_{\sup} + \left\|\frac{1}{h}\right\|_{\sup}\mathrm{Var}(hf_i).$$

Since $\|hf_i\|_{\sup} \le \mathrm{Lip}_i(K)\|P^i h\|_{\sup} \le C\mathrm{Lip}_i(K)$, one has just to estimate $\mathrm{Var}(hf_i)$.

For $\underline{\omega} \in \Omega^i$ and $I \in \mathcal{A}^{i-1}_{\underline{\omega}}$, we denote by $S_{i,I,\underline{\omega}}$ the inverse branch of $T^i_{\underline{\omega}}$ restricted to I. We define also $H_{i,\underline{\omega}}(z) = H_i(z,\dots,T^i_{\underline{\omega}}z)$.

Then, we can write

$$hf_i = \sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}} \sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}}\left(\frac{hH_{i,\underline{\omega}}}{|(T^i_{\underline{\omega}})'|}\right)\circ S_{i,I,\underline{\omega}}\,\mathbb{1}_{T^i_{\underline{\omega}}(I)}.$$

It follows that

$$\mathrm{Var}(hf_i) \le \sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}}\left(\sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}}\mathrm{Var}_I\left(\frac{hH_{i,\underline{\omega}}}{|(T^i_{\underline{\omega}})'|}\right) + 2\sum_{a\in\partial\mathcal{A}^{i-1}_{\underline{\omega}}}\frac{|h(a)||H_{i,\underline{\omega}}(a)|}{|(T^i_{\underline{\omega}})'(a)|}\right)$$
$$\le \sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}}(\mathrm{I}_{\underline{\omega},i} + \mathrm{II}_{\underline{\omega},i} + \mathrm{III}_{\underline{\omega},i} + \mathrm{IV}_{\underline{\omega},i}),$$

where

$$\mathrm{I}_{\underline{\omega},i} = \sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}}\mathrm{Var}_I(h)\sup_I\frac{1}{|(T^i_{\underline{\omega}})'|}\sup_I|H_{i,\underline{\omega}}|,$$

$$\mathrm{II}_{\underline{\omega},i} = \sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}}\sup_I h\,\mathrm{Var}_I\left(\frac{1}{|(T^i_{\underline{\omega}})'|}\right)\sup_I|H_{i,\underline{\omega}}|,$$

$$\mathrm{III}_{\underline{\omega},i} = \sum_{I\in\mathcal{A}_{\underline{\omega}}^{i-1}} \sup_I h \sup_I \frac{1}{|(T_{\underline{\omega}}^i)'|} \mathrm{Var}_I(H_{i,\underline{\omega}}),$$

$$\mathrm{IV}_{\underline{\omega},i} = 2 \sum_{a\in\partial\mathcal{A}_{\underline{\omega}}^{i-1}} \frac{|h(a)||H_{i,\underline{\omega}}(a)|}{|(T_{\underline{\omega}}^i)'(a)|}.$$

Using the Lipschitz condition for K, one gets $\mathrm{I}_{\underline{\omega},i} \le C\mathrm{Lip}_i(K) \sum_{I\in\mathcal{A}_{\underline{\omega}}^{i-1}} \sup_I \frac{1}{|(T_{\underline{\omega}}^i)'|}$, which gives, by Lemma 8.5, $\sum_{\underline{\omega}\in\Omega^i} p_{\underline{\omega}}^i \mathrm{I}_{\underline{\omega},i} \le C\mathrm{Lip}_i(K)$. The same argument applies to prove that $\sum_{\underline{\omega}\in\Omega^i} p_{\underline{\omega}}^i \mathrm{II}_{\underline{\omega},i} \le C\mathrm{Lip}_i(K)$.

We turn now to the estimate of $\mathrm{III}_{\underline{\omega},i}$. Let $y_0 < \cdots < y_l$ be a sequence of points of I. In order to estimate $\sum_{j=0}^{l-1} |H_{i,\underline{\omega}}(y_{j+1}) - H_{i,\underline{\omega}}(y_j)|$, we split $H_{i,\underline{\omega}}$ into two terms in an obvious way, and we deal with the first one, the second being completely similar. We have

$$\begin{aligned}
&\sum_{j=0}^{l-1}\sum_{k=0}^{i} |K(y_{j+1},\ldots,T_{\underline{\omega}}^k y_{j-1}, T_{\underline{\omega}}^{k+1} y_j,\ldots,T_{\underline{\omega}}^i y_j,\ldots)\\
&\quad -K(y_{j+1},\ldots,T_{\underline{\omega}}^{k-1} y_{j+1}, T_{\underline{\omega}}^k y_j,\ldots,T_{\underline{\omega}}^i y_j,\ldots)|\\
&\le \sum_{j=0}^{l-1}\sum_{k=0}^{i} \mathrm{Lip}_k(K)\left|T_{\underline{\omega}}^k y_{j+1} - T_{\underline{\omega}}^k y_j\right| = \sum_{k=0}^{i} \mathrm{Lip}_k(K) m(T_{\underline{\omega}}^k(I)).
\end{aligned}$$

Since $I \in \mathcal{A}_{\underline{\omega}}^{i-1}$, $T_{\underline{\omega}}^k(I)$ is included in an interval of monotonicity of $T_{\omega_i}\circ\cdots\circ T_{\omega_{k+1}}$, and hence its length is less than $(\lambda(T_{\omega_i}\circ\cdots\circ T_{\omega_{k+1}}))^{-1} \le \lambda^{-(i-k)}$. Therefore, one has $\mathrm{Var}_I(H_{i,\underline{\omega}}) \le \sum_{k=0}^i \lambda^{-(i-k)}\mathrm{Lip}_k(K)$. An application of Lemma 8.5 shows that $\sum_{\underline{\omega}\in\Omega^i} p_{\underline{\omega}}^i \mathrm{III}_{\underline{\omega},i} \le C\sum_{k=0}^i \lambda^{-(i-k)}\mathrm{Lip}_k(K)$.

Using again Lemma 8.5 and Lipschitz condition on K, we can bound the last term by $\sum_{\underline{\omega}\in\Omega^i} p_{\underline{\omega}}^i \mathrm{IV}_{\underline{\omega},i} \le C\mathrm{Lip}_i(K)$.

Finally, putting together all the estimates, we find that $\mathrm{Var}(hf_i) \le C\sum_{k=0}^i \lambda^{-(i-k)}$ $\mathrm{Lip}_k(K)$, which gives $\mathrm{Var}(f_i) \le C\sum_{k=0}^i \lambda^{-(i-k)}\mathrm{Lip}_k(K)$, and the same estimate for $\|f_i\|_{\mathrm{BV}}$.

We then have

$$\begin{aligned}
&\left|K_p(x_p,\ldots) - \int_{\tilde{\Omega}}\int_X K(y, T_{\underline{\omega}}^1 y,\ldots,T_{\underline{\omega}}^{p-1} y, x_p,\ldots) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega})\right|\\
&\le C\sum_{i=0}^{p-1} \rho^{p-i} \sum_{k=0}^{i} \lambda^{-(i-k)}\mathrm{Lip}_k(K).
\end{aligned}$$

A simple calculation shows that this term is less than $C\sum_{k=0}^{p-1}(\rho')^{p-k}\mathrm{Lip}_k(K)$, for $\max(\rho,\lambda^{-1}) < \rho' < 1$. This concludes the proof. □

Concentration inequalities have several statistical applications concerning the empirical measure, the shadowing, the integrated periodogram, the correlation dimension, the kernel density estimation, the almost-sure CLT, ... We describe here an application to the rate of convergence of the empirical measure to the stationary measure, and refer the reader to [6,10,11,67] for others possibilities. We also mention the work of Maldonado [40], where concentration inequalities are proved in a random context. He considers the so-called observational noise, where the randomness doesn't affect the dynamics, but only the observations, so the setup is somewhat different from ours, but once an annealed concentration inequality is established, all consequences are derived in a similar way.

The empirical measure is the random measure defined by

$$\mathcal{E}_n(x, \underline{\omega}) = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{T_{\underline{\omega}}^j x}.$$

Since the skew-product system $(X \times \tilde{\Omega}, S, \mu \otimes \tilde{\mathbb{P}})$ is ergodic, it follows from Birkhoff's theorem that $\mathcal{E}_n(x, \underline{\omega})$ converges weakly to the stationary measure μ, for $\mu \otimes \tilde{\mathbb{P}}$-ae $(x, \underline{\omega})$. For statistical purposes, it proves useful to estimate the speed of this convergence. We introduce the Kantorovitch distance κ on the space of probability measures on $[0, 1]$. For any ν_1, ν_2 probabilities measure on the unit interval, their Kantorovitch distance $\kappa(\nu_1, \nu_2)$ is equal to

$$\kappa(\nu_1, \nu_2) = \int_0^1 |F_{\nu_1}(t) - F_{\nu_2}(t)| dt,$$

where $F_\nu(t) = \nu([0, t])$ is the distribution function of ν. We show the following:

Proposition 8.6 *The exists $t_0 > 0$ and $C > 0$ such that for all $t > t_0$ and $n \geq 1$:*

$$\mu \otimes \tilde{\mathbb{P}} \left(\left\{ (x, \underline{\omega}) \,/\, \kappa(\mathcal{E}_n(x, \underline{\omega}), \mu) > \frac{t}{\sqrt{n}} \right\} \right) \leq e^{-Ct^2}.$$

Proof We follow closely the proof of Theorem III.1 in [11]. For $t \in [0, 1]$, define the function of n variables

$$K_n(x_0, \ldots, x_{n-1}) = \int_0^1 |F_{n,t}(x_0, \ldots, x_{n-1}) - F_\mu(t)| dt,$$

where $F_{n,t}$ is given by

$$F_{n,t}(x_0, \ldots, x_{n-1}) = \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{[0,t]}(x_k).$$

We clearly have $\kappa(\mathcal{E}_n(x,\underline{\omega}),\mu) = K_n(x,\ldots,T_{\underline{\omega}}^{n-1}x)$, and $\mathrm{Lip}_j(K_n) \le \frac{1}{n}$ for any $0 \le j \le n-1$. We derive immediately from the exponential concentration inequality (see the remark just below Theorem 8.1) that

$$\mu \otimes \tilde{\mathbb{P}}\left(\left\{(x,\underline{\omega}) \,/\, \kappa(\mathcal{E}_n(x,\underline{\omega}),\mu) - \mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(\kappa(\mathcal{E}_n(.),\mu)) > \frac{t}{\sqrt{n}}\right\}\right) \le e^{-Ct^2}.$$

To conclude, it is then sufficient to prove that $\mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(\kappa(\mathcal{E}_n(.),\mu))$ is of order $\frac{1}{\sqrt{n}}$.

Using Schwartz inequality, we have

$$\begin{aligned}\mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(\kappa(\mathcal{E}_n(.),\mu)) &= \int_0^1 \left(\int_{X\times\tilde{\Omega}} |F_{n,t}(x,\ldots,T_{\underline{\omega}}^{n-1}x) - F_\mu(t)| d\mu(x) d\tilde{\mathbb{P}}(\underline{\omega})\right) dt \\ &\le \left[\int_0^1 \left(\int_{X\times\tilde{\Omega}} |F_{n,t}(x,\ldots,T_{\underline{\omega}}^{n-1}x) - F_\mu(t)|^2 d\mu(x) d\tilde{\mathbb{P}}(\underline{\omega})\right) dt\right]^{\frac{1}{2}}.\end{aligned}$$

Expanding the square and using the invariance of $\mu \otimes \tilde{\mathbb{P}}$ by the skew-product, we obtain

$$\begin{aligned}\int_{X\times\tilde{\Omega}} |F_{n,t}(x,\ldots,T_{\underline{\omega}}^{n-1}x) - F_\mu(t)|^2 d\mu(x) d\tilde{\mathbb{P}}(\underline{\omega}) &= \frac{1}{n}\int_0^1 (f_t - F_\mu(t))^2 d\mu \\ + \frac{2}{n}\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\int_X (f_t - F_\mu(t))(U^k f_t - F_\mu(t)) d\mu,\end{aligned}$$

where f_t is the characteristic function of $[0,t]$ and $U^k f_t(x) = \int_{\tilde{\Omega}} f_t(T_{\underline{\omega}}^k x) d\tilde{\mathbb{P}}(\underline{\omega})$ as usual.

Since $F_\mu(t) = \int_X f_t d\mu$ and f_t is bounded independently of t in BV, we can use exponential decay of annealed correlations to get $\int_X (f_t - F_\mu(t))(U^k f_t - F_\mu(t)) d\mu = \mathcal{O}(\lambda^k)$, where $\lambda < 1$, independently of t. This shows $\int_{X\times\tilde{\Omega}} |F_{n,t}(x,\ldots,T_{\underline{\omega}}^{n-1}x) - F_\mu(t)|^2 d\mu(x) d\tilde{\mathbb{P}}(\underline{\omega}) = \mathcal{O}(n^{-1})$ and after integration over t, we finally get $\mathbb{E}_{\mu\otimes\tilde{\mathbb{P}}}(\kappa(\mathcal{E}_n(.),\mu)) = \mathcal{O}(n^{-\frac{1}{2}})$. □

Acknowledgments RA was partially supported by Conseil Régional Provence-Alpes-Côte d'Azur. RA and SV were supported by the ANR- Project *Perturbations* and by the PICS (Projet International de Coopération Scientifique), Propriétés statistiques des systèmes dynamiques deterministes et aléatoires, with the University of Houston, n. PICS05968. SV thanks the University of Houston for supporting his visits during the completion of this work. MN was supported by the French CNRS with a *poste d'accueil* position at the Center of Theoretical Physics in Luminy. MN was partially supported by NSF grant DMS 1101315. RA wishes to thank Carlangelo Liverani for discussions and encouragements.

References

1. Lasota, A., Yorke, J.-A.: On the existence of invariant measures for piecewise monotonic transformations. Trans. Am. Math. Soc. **186**, 481–488 (1973)
2. Saussol, B.: Absolutely continuous invariant measures for multidimensional expanding maps. Isr. J. Math. **116**, 223–248 (2000)
3. Guivarc'h, Y., Hardy, J.: Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov. Ann. de l'I.H.P., section B, **24**, 73–98 (1988)
4. Hofbauer, F., Keller, G.: Ergodic properties of invariant measures for piecewise monotonic transformations. Mathematische Zeitschrift **180**, 119–140 (1982)
5. Rousseau-Egele, J.: Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux. Ann. Probab. **11**, 772–788 (1983)
6. Chazottes, J.-R., Collet, P.: Almost sure central limit theorems and Erdös–Rényi type law for expanding maps of the interval. Ergod. Theory Dyn. Syst. **25**, 419–441 (2005)
7. Denker, M., Nicol, M.: Erdös–Rényi laws for dynamical systems. J. Lond. Math. Soc. **87**, 497–508 (2012)
8. Chernov, N., Kleinbock, D.: Dynamical Borel–Cantelli lemmas for Gibbs measures. Isr. J. Math. **122**, 1–27 (2001)
9. Kim, D.: The dynamical Borel–Cantelli lemma for interval maps. Discret. Contin. Dyn. Syst. **17**, 891–900 (2007)
10. Chazottes, J.-R., Gouëzel, S.: Optimal concentration inequalities for dynamical systems. Commun. Math. Phys. **316**, 843–889 (2012)
11. Collet, P., Martinez, S., Schmitt, B.: Exponential inequalities for dynamical measures of expanding maps of the interval. Probab. Theory Relat. Fields **123**, 301–322 (2002)
12. Baladi, V.: Positive Transfer Operators and Decay of Correlations, vol. 16. World Scientific, Singapore (2000)
13. Boyarsky, A., Göra, P.: Laws of Chaos. Invariant Measures and Dynamical Systems in One Dimension. Probability and Its Applications. Birkhauser, Boston (1997)
14. Kifer, Y.: Ergodic Theory for Random Transformations. Birkhauser, Boston (1986)
15. Arnold, L.: Random Dynamical Systems. Springer, Berlin (1998)
16. Kifer, Y.: Random Perturbations of Dynamical Systems. Birkhauser, Boston (1988)
17. Kobre, E., Young, L.-S.: Extended systems with deterministic local dynamics and random jumps. Commun. Math. Phys. **275**, 709–720 (2007)
18. Tümel, F.: Random walks on a lattice with deterministic local dynamics. Ph.D thesis, University of Houston (2012)
19. Pelikan, S.: Invariant densities for random maps of the interval. Trans. Am. Math. Soc. **281**, 813–825 (1984)
20. Morita, T.: Random iteration of one-dimensional transformations. Osaka J. Math. **22**, 489–518 (1985)
21. Hsieh, L.-Y.S.: Ergodic theory of multidimensional random dynamical systems. Ph.D. Thesis, University of Victoria (2008)
22. Bahsoun, W., Göra, P.: Position dependent random maps in one and higher dimensions. Studia Math. **166**, 271–286 (2005)
23. Baladi, V.: Correlation spectrum of quenched and annealed equilibrium states for random expanding maps. Commun. Math. Phys. **186**, 671–700 (1997)
24. Baladi, V., Young, L.-S.: On the spectra of randomly perturbed expanding maps. Commun. Math. Phys. **156**, 355–385 (1993)
25. Ishitani, H.: Central limit theorems for the random iterations of 1-dimensional transformations (dynamics of complex systems). RIMS Kokyuroku, Kyoto Univ. **1404**, 21–31 (2004)
26. Baladi, V., Kondah, A., Schmitt, B.: Random correlations for small perturbations of expanding maps. Random Comput. Dyn. **4**, 179–204 (1996)
27. Buzzi, J.: Exponential decay of correlations for random Lasota–Yorke maps. Commun. Math. Phys. **208**, 25–54 (1999)
28. Kifer, Y.: Thermodynamic formalism for random transformations revisited. Stoch. Dyn. **08**, 77–102 (2008)
29. Kifer, Y.: Limit theorems for random transformations and processes in random environments. Trans. Am. Math. Soc. **350**, 1481–1518 (1998)

30. Ayyer, A., Stenlund, M.: Exponential decay of correlations for randomly chosen hyperbolic toral automorphisms, vol. 17. Chaos (2007)
31. Ayyer, A., Liverani, C., Stenlund, M.: Quenched CLT for random toral automorphism. Discret. Contin. Dyn. Syst. **24**, 331–348 (2009)
32. Hennion, H., Hervé, L.: Central limit theorems for iterated random Lipschitz mappings. Ann. Probab. **32**, 1934–1984 (2004)
33. Lacey, M.T., Philipp, W.: A note on the almost sure central limit theorem. Stat. Probab. Lett. **9**, 201–205 (1990)
34. Gordin, M.I.: The central limit theorem for stationary processes. Soviet. Math. Dokl. **10**, 1174–1176 (1969)
35. Hall, P., Heyde, C.C.: Martingale Limit Theory and Its Application. Probability and Mathematical Statistics. Academic Press, New York (1980)
36. Haydn, N., Nicol, M., Vaienti, S., Zhang, L.: Central limit theorems for the shrinking target problem. J. Stat. Phys. (2013). doi:10.1007/s10955-013-0860-3
37. Erdös, P., Rényi, A.: On a new law of large numbers. J. Anal. Math. **23**, 103–111 (1970)
38. Jaerisch, J., Kesseböhmer, M., Stratmann, B.: A Fréchet law and an Erdös–Philipp law for maximal cuspidal windings. Ergod. Theory Dyn. Syst. **33**(4), 1008–1028 (2013)
39. Cristadoro, G., Haydn, N., Marie, P., Vaienti, S.: Statistical properties of intermittent maps with unbounded derivative. Nonlinearity **23**, 1071–1096 (2010)
40. Maldonado, C.: Fluctuation bounds for chaos plus noise in dynamical systems. J. Stat. Phys. **148**, 548–564 (2012)
41. Ionescu-Tulcea, C.T., Marinescu, G.: Théorie ergodique pour des classes d'opérations non complètement continues. Ann. Math. **52**, 140–147 (1950)
42. Hennion, H.: Sur un théorème spectral et son application aux noyaux Lipschitziens. Proc. AMS **118**, 627–634 (1993)
43. Keller, G.: Generalized bounded variation and applications to piecewise monotonic transformations. Z. Wahr. verw. Geb. **69**, 461–478 (1985)
44. Liverani, C.: Multidimensional expanding maps with singularities: a pedestrian approach. Ergod. Theory Dyn. Syst. **33**, 168–182 (2013)
45. Boyarsky, A., Göra, P.: Absolutely continuous invariant measures for piecewise expanding C^2 transformations in R^N. Isr. J. Math. **67**, 272–286 (1989)
46. Thomine, D.: A spectral gap for transfer operators of piecewise expanding maps. Discret. Contin. Dyn. Syst. **30**, 917–944 (2011)
47. Tsujii, M.: Absolutely continuous invariant measures for expanding piecewise linear maps. Invent. Math. **143**, 349–373 (2001)
48. Alves, J.F., Freitas, J.M., Luzzato, S., Vaienti, S.: From rates of mixing to recurrence times via large deviations. Adv. Math. **228**(2), 1203–1236 (2011)
49. Hennion, H., Hervé, L.: Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasicompactness. Lecture Notes in Mathematics, vol. 1766. Springer, Berlin (2001)
50. Durrett, R.: Probability: Theory and Examples, 4th edn. Cambridge Series in Statistical and Probabilistic Mathematics (2010)
51. Dembo, A., Zeitouni, O.: Large Deviations, Techniques and Applications. Applications of Mathematics, vol. 38, 2nd edn. Springer, Berlin (1998)
52. Ellis, R.S.: Entropy, Large Deviations and Statistical Mechanics. Springer, New York (1985)
53. Aimino, R., Vaienti, S.: A note on the large deviations for piecewise expanding multidimensional maps. In: Ugalde, E., Salazar, G. (eds.) Nonlinear Dynamics: New Directions, Theoretical Aspects, vol.1, 11 p. Series Mathematical Method and Modeling. Springer, Berlin. http://arxiv.org/abs/1110.5488
54. Breiman, L.: Probability. Addison-Wesley, Reading (1968)
55. Morita, T.: A generalized local limit theorem for Lasota–Yorke transformations. Osaka J. Math. **26**, 579–595 (1989)
56. Broise, A.: Etudes spectrales d'opérateurs de transfert et applications. Astérisque **238** (1996)
57. Aaronson, J., Denker, M., Sarig, O., Zweimuller, R.: Aperiodicity of cocycles and conditional local limit theorems. Stoch. Dyn. **4**, 31–62 (2004)
58. Aaronson, J., Denker, M.: Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps. Stoch. Dyn. **1**, 193–237 (2001)
59. Melbourne, I., Nicol, M.: A vector-valued almost sure invariance principle for hyperbolic dynamical systems. Ann. Probab. **37**, 478–505 (2009)

60. Gouëzel, S.: Almost sure invariance principle for dynamical systems by spectral methods. Ann. Probab. **38**, 1639–1671 (2010)
61. Liverani, C.: Central Limit Theorem for Deterministic Systems. Pitman Research Notes in Mathematics Series, vol. 362, pp. 56–75 (1996)
62. Billingsley, P.: The Lindeberg–Lévy theorem for martingales. Proc. AMS **12**, 788–792 (1961)
63. Sprindzuk, V.G.: Metric Theory of Diophantine Approximations. V. H. Winston and Sons, Washington, D.C. (1979) (Translated from the Russian and edited by Richard A. Silverman, With a foreword by Donald J. Newman, Scripta Series in Mathematics. MR MR548467 (80k:10048))
64. McDiarmid, C.: On the Method of Bounded Differences, in Surveys in Combinatorics. London Mathematical Society Lecture Note Series, vol. 141, pp. 148–188. Cambridge Univ. Press, London (1989)
65. McDiarmid, C.: Concentration, in Probabilistic Methods for Algorithmic Discrete Mathematics. Algorithms and Combinatorics, vol. 16, pp. 195–248. Springer, Berlin (1998)
66. Morita, T.: Deterministic version lemmas in ergodic theory of random dynamical systems. Hiroshima Math. J. **18**, 15–29 (1988)
67. Chazottes, J.-R., Collet, P., Schmitt, B.: Statistical consequences of the Devroye inequality for processes. Applications to a class of non-uniformly hyperbolic dynamical systems. Nonlinearity **18**, 2341–2364 (2005)