

# A generalized central limit theorem in infinite ergodic theory

Damien Thomine

Received: 28 September 2011 / Revised: 1 February 2013 / Published online: 27 February 2013  
© Springer-Verlag Berlin Heidelberg 2013

**Abstract** We prove a generalized central limit theorem for dynamical systems with an infinite ergodic measure which induce a Gibbs–Markov map on some subset, provided the return time to this subset has regularly varying tails. We adapt a method designed by Csáki and Földes for observables of random walks to show that the partial sums of some functions of the system—the return time and the observable—are asymptotically independent. Some applications to random walks and Pomeau–Manneville maps are discussed.

**Mathematics Subject Classification** 37A40 Non-singular (and infinite-measure preserving) transformations · 37A50 Relations with probability theory and stochastic processes · 37D25 Nonuniformly hyperbolic systems (Lyapunov exponents, Pesin theory, etc.) · 60F05 Central limit and other weak theorems

Although there is no simple equivalent of Birkhoff’s ergodic theorem for dynamical systems endowed with an infinite ergodic measure, with stronger hypotheses on the system, and after some nice renormalization, it is well-known that the Birkhoff sums of any given integrable function converge weakly to a Mittag–Leffler distribution [1]. One of the nicer settings for ergodic theory with infinite measure is the theory of Gibbs–Markov systems, which encompasses recurrent random walks and hyperbolic maps with a neutral fixed point (at least in one dimension). This theory is well-developed, and recent works include [11] and [15], which give precise asymptotics on the speed of decay of correlations for Gibbs–Markov systems by using operator algebra versions of the renewal equations used in probability theory.

---

D. Thomine (✉)  
IRMAR, Rennes, France  
e-mail: damien.thomine@univ-rennes1.fr

A problem arises when one works with functions with zero average: with the usual renormalization, the distributional limit is then trivial, which means that such a result fails to give an accurate picture of the asymptotic behavior of the Birkhoff sums. When working with probability measures, one typically switches from a law of large numbers to a central limit theorem. Much less is known for systems with infinite measures. Even in this setting of Gibbs–Markov systems, the known bounds for the sums of observables with zero average are of the same order that the bounds for functions of non-zero average. One exception to this claim is Theorem 1.3 in [15], but the nature of this result is outside the scope of our paper, which deal with distributional and almost sure limits.

An equivalent of the central limit theorem for the observables of a simple random walk was found by Dobushin in 1955 [6]. Up to our knowledge, the most advanced work on this subject was written by Csáki and Földes in 1998 and 2000 [4, 5] and deals with more general one and two-dimensional random walks. In this paper we extend the method of the latter to study the more general setting of transformations which induce a mixing Gibbs–Markov map on some part of the system.

**How to approach the subject:** One does not work easily with infinite measures. A standard tactic is to fix some nice part of the system, with finite measure, and to study the induced transformation on this subset. Then, one can hope to get a central limit theorem for the induced system. If  $(S_n)$  is a simple random walk on  $\mathbb{Z}$  and  $f : \mathbb{Z} \rightarrow \mathbb{R}$  is an observable, this means for instance that we look at the excursions from 0, and we get roughly:

$$\sum_{k=0}^{N-1} f(S_k) \sim \sum_{k=0}^{\xi_{N-1}} \sum_{\ell \text{ in the } k\text{th excursion}} f(S_\ell),$$

where  $\xi_N := \text{Card}\{k \leq N : S_k = 0\}$  is the local time in 0. Using the central limit theorem, we know that this sum behaves like  $\sqrt{\xi_N}$  times a Gaussian random variable. Moreover,  $\xi_N$  behaves like  $\sqrt{N}$  times the absolute value of a Gaussian random variable (this is a special case of Mittag–Leffler distribution). It was shown by Dobrushin [6] that the whole sum behaves like  $N^{\frac{1}{4}}$  times  $\mathcal{N}|\mathcal{N}'|^{\frac{1}{2}}$ , where  $\mathcal{N}$  and  $\mathcal{N}'$  are independent Gaussian random variables.

The most surprising feature here is the independence of  $\mathcal{N}$  and  $\mathcal{N}'$ . One of those random variable corresponds to the limit of the sum of  $f$  along the excursions, while the other corresponds to the limit of the local time in 0. In general, the sum along the excursions and the length of the excursions are correlated. However, it was shown that they are asymptotically independent under mild assumptions on the return time in 0 [5], and for mere distributional reasons.

The argument used by Csáki and Földes to prove this asymptotic independence relies heavily on the Markov property of the random walks (i.e., on the independence of the excursions). A good part of this paper is devoted to adapting the Gibbs–Markov maps into this setting. To this end, we use explicitly the structure of Gibbs–Markov maps; the problem of the extension of our results to more general classes of dynamical systems, such as AFN maps, is open.

**Structure of the article:** This paper is divided into four sections. In Sect. 1, we explain the setting we choose and present our main results, Theorem 1.7, Theorem 1.11 and Theorem 1.12. The choice of Gibbs–Markov maps is motivated by their generality (they appear naturally when one works with random walks or non-uniformly expanding maps) and their structure (one can express some kind of decorrelation for the return times, which are usually very non-integrable and do not belong to any good function space). Section 2 presents two examples to which we can apply our results: we first get back and slightly generalize the results of [4] and [5] for random walks, and then move on to study the Pomeau–Manneville maps, a standard example of transformations with an indifferent fixed point.

In Sect. 3 we prove Theorem 1.7, and in Sect. 4.1 we prove Theorem 1.11, as well as an almost everywhere bound in Sect. 4.2 and Theorem 1.12 in Sect. 4.3.

## 1 Setting and results

We use the first subsection of this section to recall some basic definitions and properties of Gibbs–Markov dynamical systems. The two next subsections will present the two main results of our article, respectively Theorems 1.7 and 1.11.

### 1.1 Gibbs–Markov maps

The main limit theorems of this paper shall be established for Gibbs–Markov maps, or for transformations which induce a Gibbs–Markov map. We recall here some basic definitions and properties; a more in-depth introduction can be found for instance in [1, Chapter 4.7].

**Definition 1.1** (*Gibbs–Markov maps*) Let  $(\Omega, d, \mathcal{B}, \mu)$  be a measured, metric, bounded Polish space, where  $\mu$  is a probability measure. A non-singular, measurable map  $T : \Omega \rightarrow \Omega$  is said to be a Markov map if  $\mu$  is  $T$ -invariant and if there exists a partition  $\pi$  of  $\Omega$  in sets of positive measure such that:

- for all  $a$  in  $\pi$ , the image of  $a$  by  $T$  is a union of elements of  $\pi$  (up to a set of null measure);
- for all  $a$  in  $\pi$ , the map  $T$  is an isomorphism from  $a$  onto its image;
- the completion for  $\mu$  of the  $\sigma$ -algebra  $\bigvee_{n \in \mathbb{N}} T^{-n}\pi$  is  $\mathcal{B}$ .

Furthermore, such a map is said to be Gibbs–Markov if it also has the following properties:

- $\inf_{a \in \pi} \mu(Ta) > 0$ ;
- it is locally uniformly expanding: there exists  $\lambda > 1$  such that, for all  $a$  in  $\pi$  and  $x, y$  in  $a$ , we have  $d(Tx, Ty) \geq \lambda d(x, y)$ ;
- it has a Lipschitz distortion: there exists a constant  $C$  such that, for all  $a$  in  $\pi$ , for almost every  $x$  and  $y$  in  $a$ :

$$\left| \frac{d\mu}{d\mu \circ T}(x) - \frac{d\mu}{d\mu \circ T}(y) \right| \leq C d(Tx, Ty) \frac{d\mu}{d\mu \circ T}(x). \quad (1.1)$$

A Gibbs–Markov map is said to be mixing if, for any Borel sets  $A$  and  $B$ ,

$$\lim_{n \rightarrow +\infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B).$$

A function  $f$  on  $\Omega$  is said to be a coboundary if there exists a measurable function  $g$  such that  $f = g - g \circ T$ .

For any points  $x$  and  $y$ , let us denote by  $s(x, y)$  the smallest time  $n \geq 0$  at which the points  $T^n x$  and  $T^n y$  do not belong to the same element of the partition  $\pi$ . Then, for any  $\kappa > 1$ , one can define a metric  $d_\kappa$  on  $\Omega$  by  $d_\kappa(x, y) := \kappa^{-s(x,y)}$ . The dynamical system  $(\Omega, d_\kappa, \mathcal{B}, \mu, T)$  is also Gibbs–Markov if  $\kappa$  belongs to  $(1, \lambda]$ .

We shall denote by  $g(x) := \frac{d\mu}{d\mu \circ T}(x)$  the inverse of the Jacobian of  $T$  at point  $x$ , and by  $g^{(n)}(x) := g(x) \cdots g(T^{n-1}x)$  the inverse of the Jacobian of  $T^n$ . Thus, the bounded distortion property reads  $|g(x) - g(y)| \leq Cd(Tx, Ty)g(x)$  for all  $a$  in  $\pi$  and almost every  $x$  and  $y$  in  $a$ .

Naturally, we introduce the transfer operator for this dynamics: for every function  $f$  in  $\mathbb{L}^1(\Omega, \mu)$ , we put  $\mathcal{L}f(x) := \sum_{Ty=x} g(y)f(y)$ . Its iterates can be written in terms of  $g^{(n)}$  in the following way:  $\mathcal{L}^n f(x) = \sum_{T^n y=x} g^{(n)}(y)f(y)$ . A cylinder is a set  $\bar{a} = [a_0, \dots, a_{n-1}]$  such that  $a_i$  belongs to  $\pi$  for all  $0 \leq i \leq n - 1$  and  $\bar{a} = \{x \in \Omega : \forall 0 \leq i \leq n - 1, T^i x \in a_i\}$ . A Gibbs–Markov map then satisfies a stronger distortion property.

**Lemma 1.2** (Distortion Lemma) *Let  $(\Omega, d, \mu, T)$  be a Gibbs–Markov map. Then, there exists a constant  $C$  such that, for almost every  $x$  and  $y$  in  $\Omega$ , for all  $n \leq s(x, y)$ ,*

$$\left| g^{(n)}(x) - g^{(n)}(y) \right| \leq Cd(T^n x, T^n y)g^{(n)}(x), \tag{1.2}$$

and, for all cylinder of  $\bar{a}$  of length  $n$  and all  $x$  in  $\bar{a}$ :

$$g^{(n)}(x) \leq C\mu(\bar{a}). \tag{1.3}$$

For any subset  $\omega$  of  $\Omega$ , we denote by  $|\cdot|_{\text{Lip}(\omega)}$  the Lipschitz semi-norm on  $\omega$ : for any function from  $\omega$  to a metric space  $(\Omega', d')$ , it is defined by:

$$|f|_{\text{Lip}(\omega)} := \inf \{C > 0 : \forall x \in \omega, \forall y \in \omega, d'(f(x), f(y)) \leq Cd(x, y)\}.$$

We shall denote by  $\text{Lip}^\infty$  the set of functions  $f$  from  $\Omega$  to  $\mathbb{R}$  such that  $\|f\|_{\text{Lip}^\infty} := \|f\|_{\infty} + \sup_{a \in \pi} |f|_{\text{Lip}(a)}$  is finite. Let  $h$  be in  $(0, 1)$ . If  $\kappa$  is chosen close enough to 1, then any  $h$ -Hölder function (for the initial metric  $d$ ) is Lipschitz for the metric  $d_\kappa$ . Hence, any result stated for Lipschitz functions actually holds for Hölder functions.

The transfer operator of a Gibbs–Markov system acting on  $\text{Lip}^\infty$  has a spectral gap, which entails many important results. We need to use three of them. The first is the exponential decay of correlations for observables in  $\text{Lip}^\infty$ :

**Proposition 1.3** (Exponential decay of correlations) *If  $(\Omega, d, \mu, T)$  is a mixing Gibbs–Markov map, then there exist two real numbers  $C$  and  $\rho > 1$  such that, for all  $f \in \text{Lip}^\infty$  and  $g \in \mathbb{L}^1$  and for every integer  $n$ :*

$$\left| \int f \cdot g \circ T^n \, d\mu - \int f \, d\mu \cdot \int g \, d\mu \right| \leq C\rho^{-n} \|f\|_{\text{Lip}^\infty} \|g\|_{\mathbb{L}^1}. \tag{1.4}$$

The second proposition is a Central Limit Theorem for smooth observables, which can be proved either by spectral perturbation of the transfer operator (Theorem 3.7 in [9]) or martingale methods:

**Proposition 1.4** (Central Limit Theorem) *Let  $(\Omega, d, \mu, T)$  be a mixing Gibbs–Markov map. Let  $f \in \mathbb{L}^2(\Omega)$  be such that  $\int_\Omega f \, d\mu = 0$  and  $\sum_{a \in \pi} \mu(a) |f|_{\text{Lip}(a)} < +\infty$ . Then:*

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T^k \rightarrow \sigma(f)\mathcal{N},$$

where the convergence is in distribution,  $\mathcal{N}$  is a standard Gaussian random variable, and:

$$\sigma(f)^2 = \int_\Omega f^2 \, d\mu + 2 \sum_{n=1}^{+\infty} \int_\Omega f \cdot f \circ T^n \, d\mu.$$

The third proposition is a consequence of Rosenthal inequality. We have not been able to locate this version of Rosenthal inequality in the literature, but it is folklore. Proposition 4.1 in [10] states an inequality which is almost as strong (and actually strong enough for our purposes), and our version can be recovered from the martingale and weakly dependent cases, for which there is an abundant literature.

**Proposition 1.5** (Rosenthal inequality) *Let  $(\Omega, d, \mu, T)$  be an ergodic Gibbs–Markov map. Let  $p > 2$ . Let  $f \in \mathbb{L}^p(\Omega)$  be such that  $\int_\Omega f \, d\mu = 0$  and  $\sum_{a \in \pi} \mu(a) |f|_{\text{Lip}(a)} < +\infty$ . Then there exists a constant  $C$ , depending only on  $p$  and  $f$ , such that for all  $n$ :*

$$\int_\Omega \left| \sum_{k=0}^{n-1} f \circ T^k \right|^p \, d\mu \leq Cn^{\frac{p}{2}}.$$

The last inequality can be used in conjunction with a result by Serfling (Theorem B in [18]) to get estimates on the supremum of the partial sums of the process  $(f \circ T^n)$ .

### 1.2 Regular variation

A common assumption in infinite ergodic theory is that, the return time to some subset of the whole space has regularly varying tails. It is not surprising that the theory of functions with regular variation appears in this setting. We present in this subsection a few definitions and technical results.

A function  $\psi : \mathbb{R}_+ \mapsto \mathbb{R}_+$  is said to have regular variation of index  $\beta \in \mathbb{R}$  at infinity if, for all positive  $x$ :

$$\lim_{y \rightarrow +\infty} \frac{\psi(xy)}{\psi(y)} = x^\beta.$$

In addition, if  $\beta$  is nonnegative and  $\psi$  is nondecreasing, unbounded and càglàd (left-continuous with right limits) function with regular variation of index  $\beta$ , we define its generalized inverse  $\psi^*$  by:

$$\psi^*(x) = \sup \{t \geq 0 : \psi(t) \leq x\}. \tag{1.5}$$

With the assumption that  $\psi$  is càglàd and with this definition of its generalized inverse, one gets  $\psi^*(\psi(x)) \geq x$  for all  $x \geq 0$ , and  $\psi(\psi^*(y)) \leq y$  for all  $y \geq 0$ .

Moreover, for any  $y > \psi^*(x)$ , one has  $\psi(y) \geq x$ . If we assume that  $\beta > 0$ , it follows that  $\psi^*$  is regularly varying with index  $1/\beta$  (see e.g. Theorem 1.5.12 in [2]), and hence for all  $C > 1$ , for all large enough  $x$ , one has  $\psi^*(Cx) > C^{1/(2\beta)}\psi^*(x) > \psi^*(x)$  so that  $\psi(\psi^*(Cx)) \geq x$ . The case  $\beta = 0$  can be dealt with in a similar fashion. This is a way to simplify the bounds involving expressions such as  $\psi \circ \psi^*$  or  $\psi^* \circ \psi$ , which will appear in many of our proofs.

We shall also use the following technical lemma:

**Lemma 1.6** *Let  $\psi$  be a positive, càglàd, unbounded and nondecreasing function with regular variation of index  $\beta \in [0, 1)$  at infinity. Let  $\gamma^* > \gamma > 0$  and  $\kappa > 0$ . Then there exists a constant  $C$  such that, for all positive  $x$ ,*

$$x^\kappa \psi^*(x^\gamma) \leq C \psi^*(x^{\gamma^* + \kappa\beta}). \tag{1.6}$$

*Proof* Let  $\gamma < \gamma' < \gamma^*$ . Let us put  $\varepsilon := \frac{\gamma' - \gamma}{\kappa}$ . Then, as a consequence of Potter’s theorem (see e.g. Theorem 1.5.6 in [2]), there exists a constant  $C$  such that:

$$\psi(x^\kappa \psi^*(x^\gamma)) \leq C x^{\kappa(\beta + \varepsilon)} \psi(\psi^*(x^\gamma)) \leq C x^{\gamma' + \kappa\beta}.$$

Taking the inverse  $\psi^*$ , we obtain:

$$x^\kappa \psi^*(x^\gamma) \leq \psi^*(C x^{\gamma' + \kappa\beta}) \leq C \psi^*(x^{\gamma^* + \kappa\beta}).$$

□

In this article we will work with random variables with regularly varying tails. Let  $\varphi$  be a real-valued, nonnegative random variable. We will routinely use the following conditions:

$$\forall x > 0, \mathbb{P}(\varphi \geq x) \leq 1/\psi(x), \tag{1.7}$$

where the function  $\psi$  is nondecreasing, unbounded, càglàd, and has regular variation of index  $\beta \in [0, 1]$  at infinity.

We will sometimes need a stronger assumption on the tail of  $\varphi$ , such as:

$$\forall x > 0, \mathbb{P}(\varphi \geq x) = 1/\psi(x), \tag{1.8}$$

where the function  $\psi$  has regular variation of index  $\beta \in [0, 1]$  at infinity (by construction, such a function  $\psi$  is automatically nondecreasing, unbounded, and càglàd).

### 1.3 An independence result

Our first result will be expressed in terms of coupling. Let us recall the definition of a coupling, and a few useful facts. Let  $X$  and  $Y$  be two random variables taking their values respectively in some Polish spaces  $\mathcal{X}$  and  $\mathcal{Y}$ . We call a coupling between  $X$  and  $Y$  a random variable taking its value in  $\mathcal{X} \times \mathcal{Y}$  whose first marginal (its projection onto  $\mathcal{X}$ ) has the same law as  $X$ , and whose second marginal (its projection onto  $\mathcal{Y}$ ) has the same law as  $Y$ . Given three random variables  $X, Y$  and  $Z$ , a coupling between  $X$  and  $Y$ , and a coupling between  $Y$  and  $Z$ , one can canonically couple  $X$  and  $Z$  by first choosing  $Y$ , and then choosing simultaneously  $X$  and  $Z$  conditionally to the value of  $Y$ . Let us stress the fact that the existence of a specific coupling between two random variables depends only on their distribution, not on the probability space on which they are defined (as long as it is Polish).

Let us detail the different assumptions we shall use on the observables, i.e., on the functions on  $\Omega$  we study. Let  $(X, \varphi)$  be a pair of functions on  $\Omega$ , taking their values in  $\mathbb{R}$  and  $\mathbb{R}_+$  respectively. We shall think of  $X$  as a random variable with nice tails, while  $\varphi$  shall be heavy-tailed (in our applications,  $\varphi$  will be the return time for a null recurrent random walk). More precisely, we assume that  $X$  belongs to  $\mathbb{L}^p(\Omega, \mu)$  for some  $p > 2$ , and that  $\varphi$  satisfies either condition (1.7) or condition (1.8).

Moreover, we assume that both  $X$  and  $\varphi$  have some smoothness. We shall use in this article the following standard assumption, already used in Propositions 1.4 and 1.5:

$$\sum_{a \in \pi} \mu(a) |X|_{\text{Lip}(a)} + \sum_{a \in \pi} \mu(a) |\varphi|_{\text{Lip}(a)} < +\infty, \tag{1.9}$$

which implies some kind of exponential decay of correlations. This condition can however be relaxed, as our machinery works well with (fast enough) polynomial decay of correlations. This shall be the object of a future publication. Given these assumptions, we can state the first of our main results. We stress the fact that it is only valid for  $\beta$  in  $[0, 1)$ .

**Theorem 1.7** *Using the terminology of Sect. 1.3, let  $(\Omega, d, \mu, T)$  be a mixing Gibbs–Markov map. Let  $X$  and  $\varphi$  be measurable functions from  $\Omega$  to  $\mathbb{R}$  and to  $\mathbb{R}_+$  respectively, fulfilling the condition (1.9). We put  $(X_i, \varphi_i) := (X \circ T^i, \varphi \circ T^i)$ . Let  $(\tilde{X}_i)_{i \in \mathbb{N}}$  and  $(\tilde{\varphi}_i)_{i \in \mathbb{N}}$  be copies of the processes  $(X_i)_{i \in \mathbb{N}}$  and  $(\varphi_i)_{i \in \mathbb{N}}$  respectively, such that  $(\tilde{X}_i)_{i \in \mathbb{N}}$  and  $(\tilde{\varphi}_i)_{i \in \mathbb{N}}$  are mutually independent.*

*Assume that  $X$  belongs to  $\mathbb{L}^p(\Omega, \mu)$  for some  $p > 2$ . Assume that  $\varphi$  satisfies the condition (1.7) for some  $\beta \in [0, 1)$ . Then there exist  $r \in (0, 1)$  and a coupling between  $(X_i, \varphi_i)_{i \in \mathbb{N}}$  and  $(\tilde{X}_i, \tilde{\varphi}_i)_{i \in \mathbb{N}}$  such that, almost surely, for all large enough integer  $N$ ,*

$$\left| \sum_{i=0}^{N-1} X_i - \sum_{i=0}^{N-1} \tilde{X}_i \right| \leq N^{\frac{r}{2}},$$

$$\left| \sum_{i=0}^{N-1} \varphi_i - \sum_{i=0}^{N-1} \tilde{\varphi}_i \right| \leq \psi^*(N^r).$$

Assume that  $\varphi$  satisfies the stronger condition (1.8), and that  $X$  is not a coboundary. Then, the typical size of  $\sum_{i=0}^{N-1} X_i$  is  $\sqrt{N}$  (since, by Proposition 1.4, there is a central limit theorem for this sequence under our assumptions), and the sum  $\sum_{i=0}^{N-1} \varphi_i$  is roughly of magnitude  $\psi^*(N)$ . Thus, this theorem states that for some coupling, the distance between the original process  $(X_i, \varphi_i)$  and the process  $(\tilde{X}_i, \tilde{\varphi}_i)$  is small in comparison with the typical size of both these processes.

*Remark 1.8* The best situation occurs when  $X$  is in  $\mathbb{L}^p$  for all finite  $p$ . This is guaranteed if  $X$  is induced by a Lipschitz function with compact support for Pomeau–Manneville or Boole transformation, or by a function with finite support for the random walks on  $\mathbb{Z}$  or  $\mathbb{Z}^2$  (see Sect. 2.1). Then, the parameter  $r$  in Theorem 1.7 can be taken such that:

$$r < \frac{3 - \beta}{7 - 5\beta} + \varepsilon$$

for any  $\varepsilon > 0$ . With our method, the best error bound one can hope for is roughly  $3/7$  for  $\beta = 0$ . Of course, one can choose to split the control of  $\left| \sum_{i=0}^{N-1} X_i - \sum_{i=0}^{N-1} \tilde{X}_i \right|$  and  $\left| \sum_{i=0}^{N-1} \varphi_i - \sum_{i=0}^{N-1} \tilde{\varphi}_i \right|$ , and to manage them with two different parameters. Then, one can sacrifice one of the bounds to improve the other.

This shows, by the way, that our argument breaks down if  $\beta = 1$ .

### 1.4 Two limit theorems

Theorem 1.7 allows us to derive explicit distributional limits for the partial sums of observables for a class of dynamical systems. We shall now use the return time and the induced transformation on a Borel subset  $A$  of  $\Omega$ .

Let  $(\Omega, \mu, T)$  be a conservative and ergodic dynamical system, where  $\Omega$  is a Polish space and  $\mu$  is an infinite, nonnegative  $T$ -invariant measure. For any Borel set  $A$  such that  $\mu(A) > 0$ , we denote by  $\varphi_A$  the return time in  $A$ , i.e.,  $\varphi_A(x) := \inf \{i > 0 : T^i x \in A\}$  for all  $x$  in  $\Omega$ , and we put  $T_A := T^{\varphi_A}$ . Since the transformation  $T$  is conservative and ergodic, the return time is finite almost everywhere in  $A$  (see e.g. Proposition 1.2.2 in [1]), and  $(A, \mu|_A, T_A)$  is an ergodic dynamical system.

**Definition 1.9** We say that a dynamical system  $(\Omega, \mu, T)$  induces a Gibbs–Markov map on a Borel set  $A$  if  $0 < \mu(A) < +\infty$  and  $(A, d, \mu|_A/\mu(A), T_A)$ , for some metric  $d$  on  $A$  and with respect to some partition  $\pi$ , is a Gibbs–Markov map, and if  $\varphi_A$  is constant on each set of the partition  $\pi$ .

Notice that, if the set  $A$  is given, we can rescale  $\mu$  so that  $\mu|_A$  is a probability measure. From now on, when we restrict such a system to its induced system on some Borel set  $A$ , we shall always assume that  $\mu(A) = 1$ . We denote by  $\psi$  the inverse of the tail of the random variable  $\varphi_A$  under  $\mu|_A$ , that is, for all  $x \geq 0$ :

$$\psi(x) := \frac{1}{\mu|_A(\varphi_A \geq x)}.$$



Let  $A$  be a Borel subset of  $\Omega$  with positive and finite measure. For any measurable real-valued function  $f$  on  $\Omega$ , we denote by  $X_f$  the function on  $A$  defined by:

$$X_f(x) := \sum_{i=0}^{\varphi_A(x)-1} f(T^i x).$$

Since  $f$  is integrable,  $X_f$  is well-defined  $\mu|_A$ -almost everywhere and is also integrable.

If  $f$  is a coboundary for  $T$ , then one can check that  $X_f$  is a coboundary for  $T_A$ . Conversely, if  $X_f$  is a coboundary for  $T_A$ , then  $f$  is a coboundary for  $T$ . This is because if we have  $X_f = \tilde{g} - \tilde{g} \circ T$  and we put:

$$g(x) := \sum_{k=0}^{\varphi_A(x)-1} f(T^k x) + \tilde{g}(T^{\varphi_A(x)} x),$$

then one can check that  $f = g - g \circ T$ .

Before stating our main result, let us introduce the Mittag-Leffler distributions, which naturally appear when one deals with the distributional limit of local times.

**Definition 1.10** (*Mittag-Leffler distribution*) Let  $\beta$  be in  $[0, 1]$ . A real-valued, non-negative random variable  $Y_\beta$  is said to have a standard Mittag-Leffler distribution of order  $\beta$  if, for all  $z$  in  $\mathbb{C}$  (or all  $z$  in the open unit disc of  $\mathbb{C}$  if  $\beta = 0$ ):

$$\mathbb{E}(e^{zY_\beta}) = \sum_{n=0}^{+\infty} \frac{\Gamma(1 + \beta)^n}{\Gamma(1 + n\beta)} z^n. \tag{1.10}$$

We will also denote by  $\text{sinc}(x) := \sin(x)/x$  the cardinal sine. We may now state our second main result:

**Theorem 1.11** *Using the terminology of Sects. 1.3 and 1.4, let  $(\Omega, \mu, T)$  be a dynamical system which induces a mixing Gibbs-Markov map on a Borel set  $A$ . Assume that the function  $\psi$  associated with  $\varphi_A$  is regularly varying with index  $\beta \in [0, 1)$ .*

*Let  $f$  be in  $\mathbb{L}^1(\Omega, \mu)$ . Assume that  $\int_\Omega f \, d\mu = 0$ , that the random variable  $X_{|f|}$  belongs to  $\mathbb{L}^p(A, \mu|_A)$  for some  $p > 2$  and that:*

$$\sum_{a \in \pi} \mu|_A(a) |X_f|_{\text{Lip}(a)} < +\infty. \tag{1.11}$$

*Then, for any probability measure  $\nu \ll \mu$ :*

$$\frac{1}{\sqrt{\text{sinc}(\beta\pi)\psi(N)}} \sum_{i=0}^{N-1} f \circ T^i \rightarrow \sigma(f)\sqrt{Y_\beta}\mathcal{N}, \tag{1.12}$$

*where the convergence is in distribution when the left-hand side is seen as a random variable from  $(\Omega, \nu)$  to  $\mathbb{R}$ , where  $Y_\beta$  and  $\mathcal{N}$  are independent,  $Y_\beta$  is a standard Mittag-Leffler distribution of order  $\beta$  and  $\mathcal{N}$  is a standard Gaussian random variable, and where:*

$$\sigma(f)^2 = \int_A X_f^2 \, d\mu + 2 \sum_{i=1}^{+\infty} \int_A X_f \cdot X_f \circ T_A^i \, d\mu. \tag{1.13}$$

Moreover,  $\sigma(f) = 0$  if and only if  $f$  is a coboundary.

We will also prove a version of this theorem for the case  $\beta = 1$ , which is somewhat degenerated. Its statement is slightly different, and its proof relies on a very different mechanism (since Theorem 1.7 only holds for  $\beta < 1$ ). However, as it does not involve Theorem 1.7, we never use the Gibbs–Markov structure, but only spectral properties of the transfer operator. The conclusion stays true in a wider setting, including non-markovian nonuniformly expanding interval maps (or AFN maps) [16], but for simplicity we will only formulate the result for Gibbs–Markov maps.

**Theorem 1.12** *Assume that the hypotheses on the system and on the observables  $f$  of Theorem 1.11 hold, except that the function  $\psi$  associated with  $\varphi_A$  is regularly varying with index 1. Then, for any probability measure  $\nu \ll \mu$ :*

$$\frac{1}{\sqrt{N \sum_{k=0}^{N-1} \frac{1}{\psi(k)}}} \sum_{i=0}^{N-1} f \circ T^i \rightarrow \sigma(f)\mathcal{N}, \tag{1.14}$$

where the convergence is in distribution when the left-hand side is seen as a random variable from  $(\Omega, \nu)$  to  $\mathbb{R}$ , where  $\mathcal{N}$  is a standard Gaussian random variable, and where the variance  $\sigma(f)^2$  is given by Eq. (1.13). Moreover,  $\sigma(f) = 0$  if and only if  $f$  is a coboundary.

## 2 Applications

The two main examples we work with are random walks (Sect. 2.1) and Pomeau–Manneville maps (Sect. 2.2). Random walks are interesting for themselves, and we shall prove generalizations of some results from [4] and [5]. The Pomeau–Manneville maps are archetypal expanding interval maps with an indifferent fixed point; the integrability and smoothness conditions of our main theorems will play an explicit role (as opposed to random walks, for which the smoothness condition is somewhat hidden). We only give applications of Theorem 1.11; the corresponding statement derived from Theorem 1.12 are irrelevant for random walks, and are left to the reader for the Pomeau–Manneville maps.

### 2.1 Random walks

With this first example we shall show a way to get some of the results of [5] for random walks. Let  $\Omega_d := (\mathbb{Z}^d)^{\mathbb{N}}$ , with  $d \leq 2$ , be the space of trajectories of a random walk on  $\mathbb{Z}^d$ . We endow it with the product topology and the distance  $d(\sigma, \bar{\sigma}) := \sum_{N=0}^{+\infty} 2^{-N} \mathbf{1}_{\sigma_N \neq \bar{\sigma}_N}$ . We define on  $\Omega_d$  the one-sided shift  $T$  by putting  $(T\sigma)_N := \sigma_{N+1}$  for all  $N$ . Let  $(p_k)_{k \in \mathbb{Z}^d}$  be a probability measure on  $\mathbb{Z}^d$ .

Let  $\mu$  be the sigma-finite measure on  $\Omega_d$  such that, for any finite sequence  $(\bar{\sigma}_0, \dots, \bar{\sigma}_N)$  of points in  $\mathbb{Z}^d$ , we have:

$$\mu(\{\sigma \in \Omega_d : \sigma_0 = \bar{\sigma}_0, \dots, \sigma_N = \bar{\sigma}_N\}) = \prod_{i=0}^{N-1} p_{\bar{\sigma}_{i+1} - \bar{\sigma}_i}.$$

Then,  $T$  preserves the measure  $\mu$ .

Let  $A$  be the set  $\{0\} \times (\mathbb{Z}^d)^{\mathbb{N}^*}$ , or in other words the space of trajectories which start from 0. We now define the truncated Green function.

**Definition 2.1** (*Truncated Green function*) For any nonnegative integer  $N$ , we put:

$$g(N) := \sum_{i=0}^N \mu(A \cap T^{-i} A).$$

The function  $g$  is called the truncated Green function. If  $(\sigma_N)$  is a random walk on  $\mathbb{Z}^d$ , starting from 0 and with transition kernel  $(p_k)$  which is not supported on any strict subgroup of  $\mathbb{Z}^d$ , then one has the equality:

$$g(N) = \sum_{i=0}^N \mathbb{P}(\sigma_i = 0).$$

It is well-known (see e.g. Proposition 1.17 in [20]) that the random walk on  $\mathbb{Z}^d$  is recurrent if and only if  $g$  converges to  $+\infty$  at infinity. Under this condition, the transformation  $T$  is measure-preserving, conservative and ergodic, and the dynamical system  $(\Omega_d, \mu, T)$  induces a mixing Gibbs–Markov map on  $A$ . Moreover, the asymptotics of  $g$  can be deduced from the asymptotics of the first return time (and conversely). Hence, we can apply Theorem 1.11.

**Proposition 2.2** *Let  $(\sigma_N)_{N \in \mathbb{N}}$  be a random walk on  $\mathbb{Z}^d$ . Assume that the truncated Green function  $g$  is regularly varying with index  $\beta \in [0, 1)$  and converges to  $+\infty$  at infinity, and that the transition kernel is not supported on any strict subgroup of  $\mathbb{Z}^d$ .*

*Let  $f$  in  $\ell^1(\mathbb{Z}^d)$  be such that  $\sum_{k \in \mathbb{Z}^d} f(k) = 0$  and, for some  $p > 2$ :*

$$\int_A \left( \sum_{i=0}^{\varphi_A(\sigma)-1} |f(\sigma_i)| \right)^p d\mu(\sigma) < +\infty. \tag{2.1}$$

*Then, for any starting distribution on  $\mathbb{Z}^d$ , one has:*

$$\frac{1}{\sqrt{g(N)}} \sum_{i=0}^{N-1} f(\sigma_i) \rightarrow \left( \int_A \left( \sum_{i=0}^{\varphi_A(\sigma)-1} f(\sigma_i) \right)^2 d\mu(\sigma) \right)^{\frac{1}{2}} \sqrt{Y_\beta \mathcal{N}}, \tag{2.2}$$

where the convergence is in distribution, where  $Y_\beta$  and  $\mathcal{N}$  are independent,  $Y_\beta$  is a standard Mittag–Leffler distribution of order  $\beta$  and  $\mathcal{N}$  is a standard Gaussian random variable.

This proposition is a slightly generalized version of Theorem 2 in [5]. Actually, since the state space of the underlying dynamical system is  $\mathbb{Z}^{\mathbb{N}}$ , the observable  $f$  and the starting distribution may depend on the following steps of the random walk, and not only on the current step. One only need to be careful that the starting distribution is still absolutely continuous, and that the observables are regular enough.

Here are some sufficient conditions on the random walk under which the truncated Green function has regular variation of order  $\beta$ . If  $d = 2$  and the transition kernel  $(p_k)$  has finite variance, then the hypotheses of this proposition hold with  $\beta = 0$  (cf. the proof of Theorem 1.1 in [12]). If  $d = 1$  and  $(p_k)$  has finite variance, then the hypotheses of this proposition hold with  $\beta = 1/2$ . We recall that  $Y_0$  is an exponential random variable, and that  $Y_{1/2}$  is the absolute value of a Gaussian random variable. More generally, if  $d = 1$  and  $(p_k)$  is in the domain of attraction of a stable law of index  $\alpha \in (1, 2]$ , then the hypotheses of this proposition hold with  $\beta = 1 - 1/\alpha$  (Proposition 6.1 in [12]).

Here is a sufficient conditions on the observable  $f \in \ell^1(\mathbb{Z}^d)$  under which the corresponding hypotheses of Proposition 2.2 are satisfied. Condition (2.1) is fulfilled for all  $p < +\infty$  when the random walk is recurrent and  $f$  has finite support. In such a case, the random walk  $(\sigma_N)_{N \in \mathbb{N}}$  starting from 0 induces a transitive Markov chain on  $\{0\} \cup \text{Supp}(f)$ . Since there are only finitely many points in  $\{0\} \cup \text{Supp}(f)$ , the return time in 0 for the induced Markov chain has an exponentially decreasing tail. Thus, the number of time  $|f(\sigma_i)|$  is non-zero in an excursion also has an exponentially decreasing tail, and so has  $\sum_{i=0}^{\varphi_A(\sigma)-1} f(\sigma_i)$ . The square of this random variable has therefore finite  $p$ -moment for all  $p < +\infty$ .

## 2.2 Pomeau–Manneville maps

As an example of a family of transformations with an ergodic, infinite invariant measure, we study the Pomeau–Manneville maps. They are maps of  $(0, 1]$  onto itself with an indifferent fixed point at 0. They have been around since at least an article of Gaspard and Wang [7], which is a good reference for the combinatorial results we use here, but we will use the formalism of Liverani et al. [13]. For each  $\alpha \geq 0$  and  $x \in (0, 1]$ , we define:

$$T_\alpha x = \begin{cases} x(1 + (2x)^\alpha), & x \in (0, 1/2]; \\ 2x - 1, & x \in (1/2, 1]. \end{cases} \quad (2.3)$$

The transformation  $T_0$  is the dyadic map; when one increases the parameter  $\alpha$ , the ergodicity of the transformation remains, but the trajectory of (Lebesgue-)almost every point spends more and more time in small neighborhoods of 0. Up to a multiplicative constant, there is still a unique, nonnegative, ergodic, invariant measure absolutely continuous with respect to Lebesgue measure; its density is locally bounded, and has a pole of order  $\alpha$  at 0 [19]. Hence, if  $\alpha \geq 1$  this measure is no longer finite. We shall

denote by  $\mu_\alpha$  the measure with these properties, and such that  $\mu_\alpha((1/2, 1]) = 1$ . For any nonnegative  $\alpha$ , the transformation  $T_\alpha$  induces a mixing Gibbs–Markov map on  $A := (1/2, 1]$ . Let  $|\cdot|_{\text{Hol}_h(B)}$  denote the  $h$ -Hölder semi-norm on a subset  $B$  of  $(0, 1]$ . We apply Theorem 1.11:

**Proposition 2.3** *Let  $\alpha > 1$ . Let  $f$  be a real function over  $(0, 1]$  which is locally Hölder-continuous with exponent  $h > 0$ , and such that:*

- $|f|(x) = O(x^{\alpha-\frac{1}{2}+\varepsilon})$  at 0 for some  $\varepsilon > 0$ ;
- $|f|_{\text{Hol}_h((0,x))} = O(x^{\alpha-1+\varepsilon'})$  at 0 for some  $\varepsilon' > 0$ ;
- $\int_{(0,1]} f \, d\mu_\alpha = 0$ .

Then there exists some  $\sigma \geq 0$  such that, for any probability measure  $\nu$  absolutely continuous with respect to Lebesgue measure,

$$\frac{1}{N^{\frac{1}{2\alpha}}} \sum_{k=0}^{N-1} f(T_\alpha^k x) \rightarrow \sigma(f) \sqrt{Y_{\frac{1}{\alpha}} \mathcal{N}},$$

where the convergence is in distribution when the left-hand side is seen as a random variable from  $((0, 1], \nu)$  to  $\mathbb{R}$ , where  $Y_{\frac{1}{\alpha}}$  and  $\mathcal{N}$  are independent,  $Y_{\frac{1}{\alpha}}$  is a standard Mittag–Leffler distribution of order  $1/\alpha$  and  $\mathcal{N}$  is a standard Gaussian random variable. Moreover,  $\sigma(f) = 0$  if and only if  $f$  is a coboundary.

The hypotheses cover the case of Hölder functions with compact support in  $(0, 1]$ .

*Proof* Let us put  $A := (1/2, 1]$ , and let us define  $A_N := \{x \in A : \varphi_A(x) = N\}$ . It is well-known that the Pomeau–Manneville map induces a mixing Gibbs–Markov map on  $A$  endowed with the partition  $(A_N)_{N \geq 1}$ . Moreover, one has (see e.g. [7]):

$$\text{Leb}(A_N) \leq \frac{C_\alpha}{N^{1+\frac{1}{\alpha}}}. \tag{2.4}$$

Let  $f$  be a function satisfying the hypotheses of Proposition 2.3. Obviously, the restriction of  $X_f$  to any  $A_N$  is Lipschitz.

**First step: Integrability condition**

We first check that  $X_{|f|}$  belongs to  $\mathbb{L}^p(A, \mu_{\alpha|A})$  for some  $p > 2$ ; since the density of  $\mu_\alpha$  with respect to the Lebesgue measure is locally bounded (and, as a consequence, bounded on  $A$ ), it is enough to prove that  $X_{|f|}$  belongs to  $\mathbb{L}^p(A, \text{Leb})$ .

Due to the combinatorial nature of the Pomeau–Manneville maps and the assumptions on  $f$  (notice that we can take  $\varepsilon < 1/2$ ), we have:

$$\sup_{A_N} X_{|f|} \leq \sum_{k=1}^N \sup_{T_\alpha^k A} |f| \leq C \sum_{k=1}^N \left(k^{-\frac{1}{\alpha}}\right)^{\alpha-\frac{1}{2}+\varepsilon} \leq CN^{\frac{1}{2\alpha}-\frac{\varepsilon}{\alpha}}.$$

By Eq. (2.4) we get:

$$\|X_{|f|}\|_{\mathbb{L}^p}^p \leq \sum_{N=1}^{+\infty} \text{Leb}(A_N) \sup_{A_N} X_{|f|}^p \leq C \sum_{N=1}^{+\infty} N^{\frac{p}{2\alpha} - \frac{p\varepsilon}{\alpha} - 1 - \frac{1}{\alpha}} = C \sum_{N=1}^{+\infty} N^{-1 - \frac{2-p(1-2\varepsilon)}{2\alpha}},$$

which shows that  $X_{|f|}$  belongs to  $\mathbb{L}^p(A, \text{Leb})$  for any  $p \in \left(2, \frac{2}{1-2\varepsilon}\right)$ .

**Second step: Smoothness condition**

Now, we have to check the smoothness condition on  $X_f$  in order to apply Theorem 1.11. As explained in Sect. 1.1, we can define for any  $\kappa > 1$  a symbolic metric  $d_\kappa$  on  $A$ , and if  $\kappa$  is close enough to 1 then  $(A, d_\kappa, \mu_{\alpha|A}, T_A)$  is Gibbs–Markov. We denote by  $|\cdot|_{\text{Lip}_\kappa}$  the Lipschitz semi-norm with respect to the symbolic metric  $d_\kappa$  for some well-chosen  $\kappa$ . Up to taking  $\kappa$  even closer to 1, we can assume that  $d^h \leq C d_\kappa$ .

The following computations show the same pattern as the estimates on  $\|X_{|f|}\|_{\mathbb{L}^p}$ . First, we have for any  $x$  and  $y$  in  $A_N$ :

$$\begin{aligned} |X_f(x) - X_f(y)| &\leq \sum_{k=0}^{N-1} \left| f(T^k x) - f(T^k y) \right| \\ &\leq \sum_{k=1}^N |f|_{\text{Hol}_h(T_\alpha A_k)} d(T^{N-k+1}x, T^{N-k+1}y)^h \\ &\leq \sum_{k=1}^N |f|_{\text{Hol}_h(T_\alpha A_k)} d(T_A x, T_A y)^h \\ &\leq C\kappa^{-1} \sum_{k=1}^N |f|_{\text{Hol}_h(T_\alpha A_k)} d_\kappa(x, y). \end{aligned}$$

Hence, if we assume that  $\varepsilon' < 1/2$ :

$$|X_f|_{\text{Lip}_\kappa(A_N)} \leq C\kappa^{-1} \sum_{k=1}^N |f|_{\text{Hol}_h(T_\alpha A_k)} \leq C_{(1)} \sum_{k=1}^N \left(k^{-\frac{1}{\alpha}}\right)^{\alpha-1+\varepsilon'} \leq C_{(2)} N^{\frac{1-\varepsilon'}{\alpha}};$$

together with Eq. (2.4), this yields:

$$\sum_{N=1}^{+\infty} \mu_{\alpha|A}(A_N) |X_f|_{\text{Lip}_\kappa(A_N)} \leq C \sum_{N=1}^{+\infty} N^{-1 - \frac{\varepsilon'}{\alpha}} < +\infty.$$

The smoothness condition in Theorem 1.11 is satisfied.

The perturbation methods used to get Propositions 1.4 and 1.5 work as well under the weaker assumption

$$\sum_{a \in \pi} \mu(a) |X_f|_{\text{Lip}(a)}^\theta < +\infty,$$

where  $\theta > 0$  is small. Actually, all of our results can be proved under this hypothesis. The smoothness condition in Proposition 2.3 can then be replaced by “ $|f|_{\text{Hol}_h((x,1])} = O(x^r)$  at 0 for some negative number  $r$ ”. However, we will prove in a future article that all of our results hold under even weaker assumptions, so that this condition is still far from optimal.

Remark 4.6 implies that our limit theorem can also be applied to a larger class of functions. For instance, let us define a function  $f$  on  $(0, 1]$  by  $f(x) := (-1)^N$  if  $x$  belongs to  $T_\alpha A_N$  ( $N \geq 1$ ), and tune the value of  $f$  on  $A_1$  such that  $X_f$  has zero average. Then  $f$  satisfies all the assumptions of Remark 4.6, and as such the conclusions of Theorem 1.11 are valid.

Following Remark 4.8, one can even find examples of functions  $f$  on  $(0, 1]$  such that  $\lim_0 |f| = +\infty$ , and for which the conclusion of Theorem 4.7 still holds.

### 3 Asymptotic independence

In this section,  $T$  shall be a Gibbs–Markov map. Let  $X$  and  $\varphi$  be two functions satisfying the assumptions of Sect. 1.3. We want to prove Theorem 1.7 which states that, under some condition depending the smoothness (Eq. 1.9) and integrability (the parameter  $p$ ) of those functions, the partial sums  $\sum_{i=0}^{N-1} X \circ T^i$  and  $\sum_{i=0}^{N-1} \varphi \circ T^i$  are asymptotically independent. The heuristic behind this claim is that the growth of  $\sum_{i=0}^{N-1} \varphi \circ T^i$  is due to a very small set (of null asymptotic density) of events on which  $\varphi \circ T^i$  takes large values, while the growth of  $\sum_{i=0}^{N-1} X \circ T^i$  is due to the accumulation of a large number of small steps. However, this argument, already formalized in [5], needs us to work with sequences of independent random variables, so that, for instance, the event that  $\varphi \circ T^i$  is large influences only the value of  $X \circ T^i$  (and not a large number of  $X \circ T^j$ , which would break down our argument).

We work with processes which are in general not sequences of independent random variables. Thus, we shall get back some independence, and then use the former argument to prove that the partial sums of our processes are asymptotically independent. More precisely, we shall introduce processes which are, in some sense, piecewise independent.

Let  $q$  and  $\varepsilon$  be in  $(0, 1)$ . For any positive integers  $n$  and  $k$  with  $k < 2^{(1-q)n}$ , let  $I_{n,k} := \{i \in \mathbb{N} : 2^n + k2^{qn} \leq i < 2^n + (k + 1)2^{qn} - 2^{q\varepsilon n}\}$  and  $J_{n,k} := \{i \in \mathbb{N} : 2^n + (k + 1)2^{qn} - 2^{q\varepsilon n} \leq i < 2^n + (k + 1)2^{qn}\}$ . If  $k \geq 2^{(1-q)n}$ , we put  $I_{n,k} = J_{n,k} = \emptyset$ . We then define  $I := \bigcup_{(n,k) \in \mathbb{N}^2} I_{n,k}$  and  $J := \bigcup_{(n,k) \in \mathbb{N}^2} J_{n,k}$ .

**Definition 3.1** (*Piecewise i.i.d. processes*) Let  $\mathbb{B}$  be a Banach space. We say that a sequence of  $\mathbb{B}$ -valued random variables  $(Y_i)_{i \in \mathbb{N}}$  is a piecewise i.i.d. process with parameters  $q$  and  $\varepsilon$  if:

- the  $(Y_i)_{i \in I}$  are identically distributed;
- if  $i$  belongs to  $J$ , then  $Y_i = 0$ ;
- for all  $n$ , the  $\mathbb{B}^{|I_{n,k}|}$ -valued random variables  $((Y_i)_{i \in I_{n,k}})_{0 \leq k < 2^{(1-q)n}}$  are independent and identically distributed.

In the first three subsections, we shall couple the initial process  $(X \circ T^i, \varphi \circ T^i)$  with different processes (for instance, piecewise i.i.d. processes) and check that such

a coupling can be done in a good way, i.e., so that the trajectories are not too far one from another. Each of those steps will require some conditions on the parameters  $p, q$  and  $\varepsilon$ . The last part of the proof of Theorem 1.7 will consist in finding when all those conditions can be simultaneously satisfied.

### 3.1 Decorrelation for Gibbs–Markov systems

In this subsection, we deal with the random variables  $X$  and  $\varphi$  in one move. Let us put  $Y := (X, \varphi)$ . This variable takes its values in  $\mathbb{B} := \mathbb{R}^2$ , which we endow with the supremum norm. Then  $|Y|_{\text{Lip}(a)} = \max\{|X|_{\text{Lip}(a)}, |\varphi|_{\text{Lip}(a)}\}$  for all  $a$  in the partition  $\pi$ . This implies in particular that if  $X$  and  $\varphi$  satisfy the condition (1.9), then  $Y$  satisfies the same condition.

**Lemma 3.2** *Let  $(\Omega, \mu, T)$  be a mixing Gibbs–Markov map. Let  $Y$  be a function from  $\Omega$  to a Banach space  $(\mathbb{B}, \|\cdot\|_\infty)$ , and let  $q$  and  $\varepsilon$  be in  $(0, 1)$ . We put  $Y_i := Y \circ T^i$ . Let  $(Y_i^*)$  be a piecewise i.i.d. process with parameters  $q$  and  $\varepsilon$  such that, for all integers  $n$  and  $k$ , the sequences  $(Y_i)_{i \in I_{n,k}}$  and  $(Y_i^*)_{i \in I_{n,k}}$  have the same law.*

*Assume furthermore that  $Y$  fulfills the condition (1.9). Then there exists a coupling between  $(Y_i)_{i \in \mathbb{N}}$  and  $(Y_i^*)_{i \in \mathbb{N}}$  such that, almost surely,*

$$\sum_{i \in I} \|Y_i - Y_i^*\|_\infty < +\infty.$$

*Proof* Our proof proceeds in four steps. In the first one, we prove a Lipschitz bound on the iterates of the function  $Y$ . In the second step, we introduce conditional measures, and prove a fast (exponential) decorrelation for characteristic functions of cylinder sets. Then, in the third and fourth steps, we use these two ingredients to construct the desired coupling.

#### First step: Lipschitz control

Let  $\bar{a} := [\bar{a}_0, \dots, \bar{a}_{|I_{n,k}|+2q\varepsilon n-1}]$  be a cylinder of length  $2^{qn} - 2^{q\varepsilon n} + 2^{q\varepsilon n-1} = |I_{n,k}| + 2^{q\varepsilon n-1}$ , and let  $x_{\bar{a}}$  be a point of  $\bar{a}$ . Then, we have for all  $x$  in  $\bar{a}$  and all integers  $i < |I_{n,k}| + 2^{q\varepsilon n-1}$ :

$$\begin{aligned} \|Y \circ T^i x - Y \circ T^i x_{\bar{a}}\|_\infty &\leq |Y|_{\text{Lip}(\bar{a}_i)} d(T^i x, T^i x_{\bar{a}}) \\ &\leq C |Y|_{\text{Lip}(\bar{a}_i)} \lambda^{-(|I_{n,k}|+2^{q\varepsilon n-1}-i)}. \end{aligned}$$

As a corollary, we have:

$$\sum_{i=0}^{|I_{n,k}|-1} \|Y \circ T^i x - Y \circ T^i x_{\bar{a}}\|_\infty \leq C \sum_{i=0}^{|I_{n,k}|-1} |Y|_{\text{Lip}(\bar{a}_i)} \lambda^{-(|I_{n,k}|-i)} \lambda^{-2^{q\varepsilon n-1}}. \tag{3.1}$$



**Second step: Decorrelation for the measure of cylinders**

Let  $y$  be a point of  $\Omega$ . We define:

$$\tilde{\mu}_y^{(2^{qn})} := \sum_{T^{2^{qn}}x=y} g^{(2^{qn})}(x)\delta_x.$$

For any measurable and integrable function  $f$ , and for almost any  $y$ , the integral of  $f$  under the measure  $\tilde{\mu}_y^{(2^{qn})}$  is the conditional expectation of  $f(x)$  knowing that  $T^{2^{qn}}x = y$ . For any cylinder  $\bar{a}$ , one computes  $\tilde{\mu}_y^{(2^{qn})}(\bar{a}) = \mathcal{L}^{2^{qn}}1_{\bar{a}}(y)$ .

Let  $\bar{a}$  be a non-empty cylinder of length  $|I_{n,k}| + 2^{q\epsilon n-1}$ . Let  $x$  be a point of  $T^{|I_{n,k}|+2^{q\epsilon n-1}}\bar{a}$ . We write:

$$\mathcal{L}^{|I_{n,k}|+2^{q\epsilon n-1}}1_{\bar{a}}(x) = g^{(|I_{n,k}|+2^{q\epsilon n-1})}(\bar{a}x) \leq C\mu(\bar{a}),$$

where  $\bar{a}x$  is the unique point of  $\bar{a}$  whose image by  $T^{|I_{n,k}|+2^{q\epsilon n-1}}$  is  $x$ . Thanks to Lemma 1.2, we know that, for all  $x$  and  $y$  in  $T^{|I_{n,k}|+2^{q\epsilon n-1}}\bar{a}$ ,

$$\begin{aligned} \left| \mathcal{L}^{|I_{n,k}|+2^{q\epsilon n-1}}1_{\bar{a}}(x) - \mathcal{L}^{|I_{n,k}|+2^{q\epsilon n-1}}1_{\bar{a}}(y) \right| &= \left| g^{(|I_{n,k}|+2^{q\epsilon n-1})}(\bar{a}x) - g^{(|I_{n,k}|+2^{q\epsilon n-1})}(\bar{a}y) \right| \\ &\leq Cd(x, y)g^{(|I_{n,k}|+2^{q\epsilon n-1})}(\bar{a}x) \\ &\leq Cd(x, y)\mu(\bar{a}). \end{aligned}$$

This implies that the function  $\mathcal{L}^{|I_{n,k}|+2^{q\epsilon n-1}}1_{\bar{a}}$  belongs to  $\text{Lip}^\infty$  with a norm bounded by  $\mu(\bar{a})$ . Since one has  $\mathcal{L}^{2^{qn}}1_{\bar{a}} = \mathcal{L}^{2^{q\epsilon n-1}}\mathcal{L}^{|I_{n,k}|+2^{q\epsilon n-1}}1_{\bar{a}}$ , an application of Eq. (1.4) gives us:

$$\left| \tilde{\mu}_y^{(2^{qn})}(\bar{a}) - \mu(\bar{a}) \right| = \left| \mathcal{L}^{2^{qn}}1_{\bar{a}}(y) - \mu(\bar{a}) \right| \leq C\mu(\bar{a})\rho^{-2^{q\epsilon n-1}}. \tag{3.2}$$

**Third step: Construction of the coupling**

Let  $n$  and  $k$  be in  $\mathbb{N}$ . For all  $i \in I_{n,k}$ , we put  $Y_i := Y \circ T^i$ . Let  $y$  be in  $\Omega$ , and let  $(Y_{i,y})_{i \in I_{n,k}}$  be the process  $(Y \circ T^i x)_{i \in I_{n,k}}$  when the distribution of  $T^{2^n+k2^{qn}}x$  is  $\tilde{\mu}_y^{(2^{qn})}$ . Roughly speaking, we look at the process  $(Y \circ T^i x)_{i \in I_{n,k}}$  knowing that  $T^{2^n+(k+1)2^{qn}}x = y$ . For any  $y$ , let  $(Y_{i,y}^*)$  be distributed as  $(Y_i)$ .

By using the first two steps in this proof, we shall prove in the fourth step that there exists a constant  $C$ , which does not depend on  $n$ , such that for almost any choice of  $y$ , there exists a coupling between  $(Y_{i,y})$  and  $(Y_{i,y}^*)$  such that:

$$\mathbb{P} \left( \sum_{i \in I_{n,k}} \|Y_{i,y} - Y_{i,y}^*\|_\infty > \lambda^{-2^{q\epsilon n-2}} \right) \leq C \min \{ \rho, \lambda \}^{-2^{q\epsilon n-2}}. \tag{3.3}$$

First, let us assume that this bound holds. Let us put  $(Y_i^*) := \int_\Omega (Y_{i,y}^*) d\mu(y)$ . This random variable is naturally coupled with  $(Y_i) = \int_\Omega (Y_{i,y}) d\mu(y)$  on the one hand; on

the other hand, since the distribution of  $(Y_{i,y}^*)$  does not depend on  $y = T^{2^n+(k+1)2^{qn}}x$ , the random variable  $(Y_i^*)_{i \in I_{n,k}}$  is independent from the future starting from time  $2^n + (k + 1)2^{qn}$ , i.e., is independent from the  $\sigma$ -algebra  $T^{-2^n-(k+1)2^{qn}}\mathcal{B}$ .

Using the bound (3.3), and thanks to a measurable selection theorem, we can construct a coupling between  $(Y_i^*)_{i \in I_{n,k}}$  and  $(Y_i)_{i \in I_{n,k}}$  such that:

$$\mathbb{P} \left( \sum_{i \in I_{n,k}} \|Y_i - Y_i^*\|_\infty > \lambda^{-2^{q\epsilon n-2}} \right) \leq C \min \{\rho, \lambda\}^{-2^{q\epsilon n-2}}.$$

We iterate this operation. Let  $((Y_i^*)_{i \in I_{m,k}})_{m \leq n, 0 \leq k < 2^{(1-q)m}}$  be a sequence of independent random variables such that the  $((Y_i^*)_{i \in I_{m,k}})_{0 \leq k < 2^{(1-q)m}}$  are identically distributed and with the same law as  $(Y_i)_{i \in I_{m,0}}$ . We construct a coupling between the random variables  $((Y_i^*)_{i \in I_{m,k}})_{m \leq n, 0 \leq k < 2^{(1-q)m}}$  and  $((Y_i)_{i \in I_{m,k}})_{m \leq n, 0 \leq k < 2^{(1-q)m}}$  such that, for all  $m \leq n$ :

$$\mathbb{P} \left( \sum_{k=0}^{2^{(1-q)m}-1} \sum_{i \in I_{m,k}} \|Y_i - Y_i^*\|_\infty > C2^{-m} \right) \leq C2^{-m}. \tag{3.4}$$

Let  $(Y_i^*)_{i \in I}$  be a process such that the  $((Y_i^*)_{i \in I_{n,k}})_{(n,k) \in \mathbb{N}^2}$  are independent and, for all pair of integers  $(n, k)$ , the  $\mathbb{B}^{|I_{n,k}|}$ -valued random variable  $(Y_i^*)_{i \in I_{n,k}}$  have the same law as  $(Y_i)_{i \in I_{n,k}}$ . Then, by Kolmogorov extension theorem, there exists a coupling between  $(Y \circ T^i)_{i \in I}$  and  $(Y_i^*)_{i \in I}$  such that, for all integer  $m$ , the bound (3.4) holds. By Borel–Cantelli lemma, the sum  $\sum_{i \in I} \|Y_i - Y_i^*\|_\infty$  is almost surely finite. This finishes the proof of Lemma 3.2.

**Fourth step: A sharp coupling**

Now, let us turn to the proof of the bound (3.3). For given  $n$  and  $k < 2^{(1-q)n}$ , we want to couple the random variables  $(Y_{i,y})_{i \in I_{n,k}}$  and  $(Y_{i,y}^*)_{i \in I_{n,k}}$  in a sufficiently sharp way. Let  $\mathcal{X}$  be a random variable with values in  $\Omega$  and whose law is  $\mu$ , and  $\mathcal{X}_y$  be a random variable with values in  $\Omega$  and whose law is  $\tilde{\mu}_y^{(2^{qn})}$ . Then we can identify  $(Y \circ T^i(\mathcal{X}_y))_{i < |I_{n,k}|-1}$  with  $(Y_{i,y})_{i \in I_{n,k}}$  and  $(Y \circ T^i(\mathcal{X}))_{i < |I_{n,k}|-1}$  with  $(Y_{i,y}^*)_{i \in I_{n,k}}$ . This translates the fact that  $(Y_{i,y})_{i \in I_{n,k}}$  and  $(Y_{i,y}^*)_{i \in I_{n,k}}$  are actually the same function on  $\Omega$ , but when the measure on  $T^{2^n+k2^{qn}}x$  is  $\tilde{\mu}_y^{(2^{qn})}$  and  $\mu$  respectively.

From the first step of this proof, we can have a good bound on  $\sum_{i=0}^{|I_{n,k}|-1} Y \circ T^i x - Y \circ T^i z$  whenever  $x$  and  $z$  belong to the same cylinder of length, say,  $|I_{n,k}| + 2^{q\epsilon n-1} = 2^{qn} - 2^{q\epsilon n-1}$ . Now, let us show that  $\mu$  and  $\tilde{\mu}_y^{(2^{qn})}$  give approximately the same weight to each such cylinder.

By Eq. (3.2), for some constants  $C$  and  $\rho < 1$ , for any cylinder  $\bar{a}$  of length  $|I_{n,k}| + 2^{q\epsilon n-1}$ , we have:

$$\left| \tilde{\mu}_y^{(2^{qn})}(\bar{a}) - \mu(\bar{a}) \right| \leq C\mu(\bar{a})\rho^{-2^{q\epsilon n-1}}.$$

Thus, there exists a coupling between  $\mathcal{X}$  and  $\mathcal{X}_y$  such that, with probability at least  $1 - C\rho^{-2q\epsilon n-1}$ , those two random variable take their values in the same cylinder of length  $|I_{n,k}| + 2^{q\epsilon n-1}$ . This induces a coupling between  $(Y_{i,y})_{i \in I_{n,k}}$  and  $(Y_{i,y}^*)_{i \in I_{n,k}}$ . Let us show that such a coupling satisfies the bound (3.3).

By Eq. (3.1), for some constant  $C_{(1)}$ , whenever  $\mathcal{X}$  and  $\mathcal{X}_y$  take their values in the same cylinder  $\bar{a}$  of length  $|I_{n,k}| + 2^{q\epsilon n-1}$ ,

$$\sum_{i \in I_{n,k}} \left\| Y_i - Y_{i,y}^* \right\|_\infty \leq C_{(1)} \sum_{i=0}^{|I_{n,k}|-1} |Y|_{\text{Lip}(\bar{a}_i)} \lambda^{-(|I_{n,k}|-i)} \lambda^{-2q\epsilon n-1}.$$

In order to avoid very large indices, for any event  $A$ , we use the notation  $1(A)$  (instead of  $1_A$ ) for the characteristic function of  $A$ . Let  $\tau$  be in  $(\lambda^{-1}, 1)$ , and let  $\delta > 0$ . Let  $\gamma$  be a random variable taking its values in  $\pi$  such that  $\mathbb{P}(\gamma = a) = \mu(a)$ . Then, for any  $\delta > 0$ :

$$\begin{aligned} & \mathbb{P} \left( \sum_{i \in I_{n,k}} \left\| Y_{i,y} - Y_{i,y}^* \right\|_\infty > \delta \right) \\ & \leq C\rho^{-2q\epsilon n-1} + \sum_{\bar{a} \in \pi^{|I_{n,k}|}} \mu(\bar{a}) 1 \left( C_{(1)} \sum_{i=0}^{|I_{n,k}|-1} |Y|_{\text{Lip}(\bar{a}_i)} \lambda^{-(|I_{n,k}|-i)} \lambda^{-2q\epsilon n-1} > \delta \right) \\ & \leq C\rho^{-2q\epsilon n-1} + \sum_{\bar{a} \in \pi^{|I_{n,k}|}} \mu(\bar{a}) \sum_{i=0}^{|I_{n,k}|-1} 1 \left( C_{(1)} |Y|_{\text{Lip}(\bar{a}_i)} \lambda^{-(|I_{n,k}|-i)} \lambda^{-2q\epsilon n-1} > \frac{\tau^{|I_{n,k}|-i}}{1-\tau} \delta \right) \\ & = C\rho^{-2q\epsilon n-1} + \sum_{i=0}^{|I_{n,k}|-1} \sum_{\bar{a} \in \pi^{|I_{n,k}|}} \mu(\bar{a}) 1 \left( |Y|_{\text{Lip}(\bar{a}_i)} > C_{(2)} (\lambda\tau)^{|I_{n,k}|-i} \lambda^{2q\epsilon n-1} \delta \right) \\ & = C\rho^{-2q\epsilon n-1} + \sum_{i=0}^{|I_{n,k}|-1} \mathbb{P} \left( |Y|_{\text{Lip}(\gamma)} > C_{(2)} (\lambda\tau)^{|I_{n,k}|-i} \lambda^{2q\epsilon n-1} \delta \right). \end{aligned} \tag{3.5}$$

Now, we use Markov’s inequality:

$$\begin{aligned} & \mathbb{P} \left( \sum_{i \in I_{n,k}} \left\| Y_{i,y} - Y_{i,y}^* \right\|_\infty > \delta \right) \\ & \leq C\rho^{-2q\epsilon n-1} + \sum_{i=0}^{|I_{n,k}|-1} \mathbb{P} \left( |Y|_{\text{Lip}(\gamma)} > C_{(2)} \lambda^{2q\epsilon n-1} (\lambda\tau)^{|I_{n,k}|-i} \delta \right) \\ & \leq C\rho^{-2q\epsilon n-1} + \sum_{i=0}^{|I_{n,k}|-1} \frac{\mathbb{E} \left( |Y|_{\text{Lip}(\gamma)} \right)}{C_{(2)} \lambda^{2q\epsilon n-1} (\lambda\tau)^{|I_{n,k}|-i} \delta} \end{aligned}$$

$$\begin{aligned}
 &\leq C\rho^{-2q\epsilon n-1} + \frac{\sum_{a \in \pi} \mu(a) |Y|_{\text{Lip}(a)}}{C_{(2)}\lambda^{2q\epsilon n-1}\delta} \sum_{i=0}^{|I_{n,k}|-1} (\lambda\tau)^{-(|I_{n,k}|-i)} \\
 &\leq C\rho^{-2q\epsilon n-1} + \frac{\lambda\tau}{1-\lambda\tau} \frac{\sum_{a \in \pi} \mu(a) |Y|_{\text{Lip}(a)}}{C_{(2)}\lambda^{2q\epsilon n-1}\delta}.
 \end{aligned}
 \tag{3.6}$$

Let us take, for instance,  $\delta = \lambda^{-2q\epsilon n-2}$ . Then, we have:

$$\mathbb{P} \left( \sum_{i \in I_{n,k}} \|Y_{i,y} - Y_{i,y}^*\|_{\infty} > \lambda^{-2q\epsilon n-2} \right) \leq C \min \{ \rho, \lambda \}^{-2q\epsilon n-2}.$$

□

### 3.2 Controlling the gaps

In this section, we check that we do not lose too much when we introduce the gaps  $J$ . The control of the gaps from the sequence  $(\varphi_i) := (\varphi \circ T^i)$  comes from the fact that a large step has little chance to happen on a small enough set. The part concerning the sequence  $(X_i) := (X \circ T^i)$  is slightly more tricky, as we need to introduce some independence. We start with the sequence  $(\varphi_i)$ .

**Lemma 3.3** *Assume that  $\varphi$  satisfies the condition (1.7). Let  $r \in (1 - (1 - \epsilon)q, 1)$ . Almost surely, for all large enough integer  $N$ ,*

$$\sum_{\substack{i \leq N \\ i \in J}} \varphi_i \leq \psi^*(N^r).
 \tag{3.7}$$

*Proof* Let  $r' \in (1 - (1 - \epsilon)q, r)$ , let  $C > 1$  and let  $n$  be a positive integer. In the following computation, we cut the random variables  $\varphi_i$  at threshold  $\psi^*(2^{r'n})$  and then use Markov inequality:

$$\begin{aligned}
 &\mathbb{P} \left( \sum_{\substack{2^n \leq i < 2^{n+1} \\ i \in J}} \varphi_i > \psi^*(C2^{r'n}) \right) \\
 &\leq \mathbb{P} \left( \exists 2^n \leq i < 2^{n+1}, i \in J, \varphi_i > \psi^*(C2^{r'n}) \right) \\
 &\quad + \mathbb{P} \left( \sum_{\substack{2^n \leq i < 2^{n+1} \\ i \in J}} \min\{\varphi_i, \psi^*(C2^{r'n})\} > \psi^*(C2^{r'n}) \right)
 \end{aligned}$$

$$\begin{aligned} &\leq 2^{(1-(1-\varepsilon)q)n} \mathbb{P}\left(\varphi > \psi^*(C2^{r'n})\right) + 2^{(1-(1-\varepsilon)q)n} \frac{\mathbb{E}\left(\min\{\varphi, \psi^*(C2^{r'n})\}\right)}{\psi^*(C2^{r'n})} \\ &\leq \frac{2^{(1-(1-\varepsilon)q)n}}{\psi(\psi^*(C2^{r'n}))} + 2^{(1-(1-\varepsilon)q)n} \frac{\mathbb{E}\left(\min\{\varphi, \psi^*(C2^{r'n})\}\right)}{\psi^*(C2^{r'n})}. \end{aligned}$$

Using Karamata’s theorem (Proposition 1.5.8 in [2]), we get the following bound:

$$\begin{aligned} \mathbb{E}\left(\min\{\varphi, \psi^*(C2^{r'n})\}\right) &\leq \int_0^{\psi^*(C2^{r'n})} \frac{1}{\psi(t)} dt + \frac{\psi^*(C2^{r'n})}{\psi(\psi^*(C2^{r'n}))} \\ &\leq C' \frac{\psi^*(C2^{r'n})}{\psi(\psi^*(C2^{r'n}))}. \end{aligned}$$

Gluing the last two inequalities together, we obtain for all large enough  $n$ :

$$\mathbb{P}\left(\sum_{\substack{2^n \leq i < 2^{n+1} \\ i \in J}} \varphi_i > \psi^*(C2^{r'n})\right) \leq C' \frac{2^{(1-(1-\varepsilon)q)n}}{\psi(\psi^*(C2^{r'n}))} \leq C' 2^{-[r'-(1-(1-\varepsilon)q)n]},$$

which is summable. Thanks to Borel–Cantelli Lemma, we know that almost surely, for all large enough integer  $n$ ,

$$\sum_{\substack{2^n \leq i < 2^{n+1} \\ i \in J}} \varphi_i \leq \psi^*(2^{r'n}).$$

Hence, almost surely, there exists a constant  $C$  such that, for all large enough integer  $N$ , if we put  $n := \lfloor \log_2 N \rfloor$ , we have:

$$\sum_{\substack{i \leq N \\ i \in J}} \varphi_i \leq \sum_{\substack{i < 2^{n+1} \\ i \in J}} \varphi_i \leq C \sum_{k=0}^n \psi^*(2^{r'k}) \leq C(n+1)\psi^*(2^{r'n}).$$

As a consequence of Lemma 1.6, for all large enough  $N$  one has:

$$\sum_{\substack{i \leq N \\ i \in J}} \varphi_i \leq C(n+1)\psi^*(2^{r'n}) \leq \psi^*(2^{r'n}) \leq \psi^*(N^r).$$

□

We have managed to control the partial sum of the elements of the sequence  $(\varphi \circ T^i)$  which belong to  $J$ . Now, we shall obtain the same kind of bounds for the sequence  $(X \circ T^i)$ .

**Lemma 3.4** *Assume that  $X \in \mathbb{L}^p(\Omega, \mu)$  for some  $p > 2$ . Let  $r \in (1 - (1 - \varepsilon)q, 1)$ . Almost surely, for all large enough integer  $N$ ,*

$$\left| \sum_{\substack{i \leq N \\ i \in J}} X_i \right| \leq N^{\frac{r}{2}}. \tag{3.8}$$

*Proof* Let  $n$  be an integer, let  $k < 2^{(1-q)n}$ , and let  $(X_i^*)_{i \in J_{n,k}}$  be a sequence of random variables distributed as  $(X_i)_{i \in J_{n,k}}$ , so that  $((X_i^*)_{i \in J_{n,k}})_{0 \leq k < 2^{(1-q)n}}$  is a sequence of independent and identically distributed random variables. As in Sect. 3.1, there exists a coupling between  $(X_i)$  and  $(X_i^*)$  such that:

$$\mathbb{P} \left( \sum_{k=0}^{2^{(1-q)n}-1} \sum_{i \in J_{n,k}} |X_i - X_i^*| > C2^{-n} \right) \leq C2^{-n}.$$

The difference  $\sum_{i \in J} |X_i - X_i^*|$  is almost surely bounded, so that it is sufficient to prove the lemma for the sequence  $(X_i^*)_{i \in J}$ . For given  $i$  and  $j$  such that  $i < j$ , we want for now an upper bound on the expectation of  $\left| \sum_{\substack{\ell \in J \\ i \leq \ell < j}} X_\ell^* \right|^p$ . Notice that  $\{\ell \in J : i \leq \ell < j\}$  can be seen as a sequence of blocks  $B$  such that the  $(X_\ell^*)_{\ell \in B}$  are all independent: these blocks  $B$  are exactly the sets  $\{\ell \in J_{n,k} : i \leq \ell < j\}$ , where  $(n, k)$  span  $\mathbb{N}^2$ . Let  $\mathbb{B}$  be the set of all blocks in  $\{\ell \in J : i \leq \ell < j\}$ . By Rosenthal inequality for sequences of independent random variables (Theorem 3 in [17]), we get for some constant  $C$  which depends only on  $p$ :

$$\mathbb{E} \left( \left| \sum_{\substack{\ell \in J \\ i \leq \ell < j}} X_\ell^* \right|^p \right) \leq C \left[ \left( \sum_{B \in \mathbb{B}} \mathbb{E} \left( \left| \sum_{\ell \in B} X_\ell^* \right|^2 \right) \right)^{\frac{p}{2}} + \sum_{B \in \mathbb{B}} \mathbb{E} \left( \left| \sum_{\ell \in B} X_\ell^* \right|^p \right) \right]. \tag{3.9}$$

Now, we use again a version of Rosenthal inequality 1.5. There exists a constant  $C$  such that, for all blocks  $B$ , no matter what their length is,

$$\mathbb{E} \left( \left| \sum_{\ell \in B} X_\ell^* \right|^2 \right) \leq C|B|,$$

and:

$$\mathbb{E} \left( \left| \sum_{\ell \in B} X_\ell^* \right|^p \right) \leq C|B|^{\frac{p}{2}}.$$

Hence, Eq. (3.9) becomes:

$$\begin{aligned} \mathbb{E} \left( \left| \sum_{\substack{\ell \in J \\ i \leq \ell < j}} X_\ell^* \right|^p \right) &\leq C \left[ \left( \sum_{B \in \mathbb{B}} |B| \right)^{\frac{p}{2}} + \sum_{B \in \mathbb{B}} |B|^{\frac{p}{2}} \right] \\ &\leq C \left( \sum_{B \in \mathbb{B}} |B| \right)^{\frac{p}{2}} = C \text{Card}\{\ell \in J : i \leq \ell < j\}^{\frac{p}{2}}. \end{aligned}$$

A result by Serfling (Corollary B1 in [18]) allows us to control the moments of the supremum of the norm of the partial sum if we only have a control on the moments of the norm of the partial sum (this is, in nature, similar to Kolmogorov’s inequality):

$$\mathbb{E} \left( \sup_{\substack{i \in J \\ i \leq N}} \left| \sum_{\substack{\ell \in J \\ 0 \leq \ell < i}} X_\ell^* \right|^p \right) \leq C \text{Card}\{\ell \in J : \ell \leq N\}^{\frac{p}{2}} \leq CN^{\frac{(1-(1-\varepsilon)q)p}{2}}.$$

Let  $r' \in (1 - (1 - \varepsilon)q, r)$ . By Markov inequality, we get for all integer  $n$ :

$$\mathbb{P} \left( \sup_{\substack{i \in J \\ i \leq 2^n}} \left| \sum_{\substack{\ell \in J \\ 0 \leq \ell < i}} X_\ell^* \right| > 2^{\frac{r'n}{2}} \right) \leq C \frac{2^{\frac{(1-(1-\varepsilon)q)pn}{2}}}{2^{\frac{pr'n}{2}}}.$$

This quantity is summable in  $n$ . By Borel–Cantelli Lemma, almost surely, for all large enough integer  $N$ ,

$$\left| \sum_{\substack{i \leq N \\ i \in J}} X_i^* \right| \leq \sup_{\substack{i \in J \\ i \leq 2^{\lfloor \log_2 N \rfloor + 1}}} \left| \sum_{\substack{\ell \in J \\ 0 \leq \ell < i}} X_\ell^* \right| \leq C 2^{\frac{r' \lfloor \log_2 N \rfloor}{2}} \leq CN^{\frac{r'}{2}} \leq N^{\frac{r}{2}}.$$

□

### 3.3 Csáki–Földes argument

We shall prove in this section the following result:

**Proposition 3.5** *Let  $(X_i^*, \varphi_i^*)_{i \in \mathbb{N}}$  be a piecewise i.i.d. process with parameters  $q$  and  $\varepsilon$ . Let  $(\bar{X}_i)_{i \in \mathbb{N}}$  and  $(\bar{\varphi}_i)_{i \in \mathbb{N}}$  be two independent processes, such that  $(\bar{X}_i)_{i \in \mathbb{N}}$  and  $(X_i^*)_{i \in \mathbb{N}}$  have the same law, and so do  $(\bar{\varphi}_i)_{i \in \mathbb{N}}$  and  $(\varphi_i^*)_{i \in \mathbb{N}}$ .*

*Assume that there exist  $p > 2$  and a constant  $C$  such that, for all  $n$ ,*

$$\mathbb{E} \left( \sup_{i < 2^{qn}} \left| \sum_{\ell=2^n}^{2^n+i} X_\ell^* \right|^p \right) \leq C 2^{\frac{pqn}{2}}. \tag{3.10}$$

If  $\varphi$  satisfies the condition (1.7), then there exist  $r \in (0, 1)$  and a coupling between  $(X_i^*, \varphi_i^*)_{i \in \mathbb{N}}$  and  $(\bar{X}_i, \bar{\varphi}_i)_{i \in \mathbb{N}}$  such that, almost surely, for all large enough integer  $N$ ,

$$\begin{aligned} \left| \sum_{i=0}^{N-1} X_i^* - \sum_{i=0}^{N-1} \bar{X}_i \right| &\leq N^{\frac{r}{2}}, \\ \left| \sum_{i=0}^{N-1} \varphi_i^* - \sum_{i=0}^{N-1} \bar{\varphi}_i \right| &\leq \psi^*(N^r). \end{aligned} \tag{3.11}$$

Let  $r$  be in  $(0, 1)$ . Let  $(X_i^{(1)}, \varphi_i^{(1)})$  and  $(X_i^{(2)}, \varphi_i^{(2)})$  be two mutually independent processes, which have the same law as  $(X_i^*, \varphi_i^*)_{i \in \mathbb{N}}$ . We construct a process  $(X'_i, \varphi'_i)$  from the processes  $(X_i^{(1)}, \varphi_i^{(1)})$  and  $(X_i^{(2)}, \varphi_i^{(2)})$  in (almost) the same fashion as Csáki and Földes did in their article [5]. We say that a block  $(X_i^{(j)}, \varphi_i^{(j)})_{i \in I_{n,k}}$  is large if there exists some integer  $i \in I_{n,k}$  such that  $\varphi_i^{(j)} \geq \psi^*(2^{rn})$ , and otherwise we say that the block is small. Let  $n$  and  $k$  be two integers.

- if either the block  $(X_i^{(1)}, \varphi_i^{(1)})_{i \in I_{n,k}}$  or the block  $(X_i^{(2)}, \varphi_i^{(2)})_{i \in I_{n,k}}$  is large, put  $(X'_i, \varphi'_i)_{i \in I_{n,k}} := (X_i^{(1)}, \varphi_i^{(1)})_{i \in I_{n,k}}$ ;
- if both the block  $(X_i^{(1)}, \varphi_i^{(1)})_{i \in I_{n,k}}$  and the block  $(X_i^{(2)}, \varphi_i^{(2)})_{i \in I_{n,k}}$  are small, put  $(X'_i, \varphi'_i)_{i \in I_{n,k}} := (X_i^{(2)}, \varphi_i^{(2)})_{i \in I_{n,k}}$ .

The process  $(X'_i, \varphi'_i)$  is piecewise i.i.d. with parameters  $q$  and  $\varepsilon$ . Since a block  $(X'_i, \varphi'_i)_{i \in I_{n,k}}$  is large if and only if  $(X_i^{(1)}, \varphi_i^{(1)})_{i \in I_{n,k}}$  is large, one checks easily that the process  $(X'_i, \varphi'_i)$  has the same distribution as the initial process  $(X_i^*, \varphi_i^*)$ . Moreover, the process  $(X_i^{(2)}, \varphi_i^{(1)})$  is distributed as the process  $(\bar{X}_i, \bar{\varphi}_i)$ .

We shall show that, by construction, the processes  $(X'_i, \varphi'_i)$  and  $(X_i^{(2)}, \varphi_i^{(1)})$  satisfy the equations (3.11). Therefore, the natural coupling between  $(X'_i, \varphi'_i)$  and  $(X_i^{(2)}, \varphi_i^{(1)})$  yields the desired coupling between  $(X_i^*, \varphi_i^*)$  and  $(\bar{X}_i, \bar{\varphi}_i)$ .

**Lemma 3.6** Assume that  $\varphi$  fulfills the condition (1.7). Let  $r^* \in (\beta + (1 - \beta)r, 1)$ . Almost surely, for all large enough integer  $N$ ,

$$\left| \sum_{i=0}^{N-1} \varphi'_i - \sum_{i=0}^{N-1} \varphi_i^{(1)} \right| \leq \psi^*(N^{r^*}). \tag{3.12}$$

*Proof* The random variables  $\varphi'_i$  and  $\varphi_i^{(1)}$  are different only in the blocks where they are both small. Thus:

$$\begin{aligned} &\left| \sum_{i=0}^{N-1} \varphi'_i - \sum_{i=0}^{N-1} \varphi_i^{(1)} \right| \\ &\leq \sum_{n=0}^{\lfloor \log_2 N \rfloor + 1} \sum_{k=0}^{2^{(1-q)n} - 1} \left( \mathbb{1}_{\{(X'_i, \varphi'_i)_{i \in I_{n,k}} \text{ small}\}} \sum_{i \in I_{n,k}} \varphi'_i + \mathbb{1}_{\{(X_i^{(1)}, \varphi_i^{(1)})_{i \in I_{n,k}} \text{ small}\}} \sum_{i \in I_{n,k}} \varphi_i^{(1)} \right). \end{aligned}$$



Now, we will bound the part with the sums  $\sum_{i \in I_{n,k}} \varphi'_i$ ; the part with the  $\varphi_i^{(1)}$  can be dealt with exactly in the same way. Let  $r' \in (\beta + (1 - \beta)r, r^*)$ . Let  $n$  be an integer. Because of Markov inequality and Karamata’s theorem (Proposition 1.5.8 in [2]), we have:

$$\begin{aligned} & \mathbb{P} \left( \sum_{k=0}^{2^{(1-q)n}-1} 1_{\{(X'_i, \varphi'_i)_{i \in I_{n,k}} \text{ small}\}} \sum_{i \in I_{n,k}} \varphi'_i > \psi^*(2^{r'n}) \right) \\ & \leq \frac{2^{(1-q)n} \mathbb{E} \left( 1_{\{(X'_i, \varphi'_i)_{i \in I_{n,0}} \text{ small}\}} \sum_{i \in I_{n,0}} \varphi'_i \right)}{\psi^*(2^{r'n})} \\ & \leq \frac{2^n \left( \sum_{\ell=0}^{\lfloor \psi^*(2^{r'n}) \rfloor} \frac{1}{\psi(\ell)} \right)}{\psi^*(2^{r'n})} \\ & \leq C 2^{(1-r)n} \frac{\psi^*(2^{r'n})}{\psi^*(2^{r'n})}. \end{aligned}$$

Since  $r' > \beta + (1 - \beta)r$ , we can choose  $\kappa > 1 - r$  such that  $r' > r + \kappa\beta$ . Then, by Lemma 1.6, we get:

$$\begin{aligned} \mathbb{P} \left( \sum_{k=0}^{2^{(1-q)n}-1} 1_{\{(X'_i, \varphi'_i)_{i \in I_{n,k}} \text{ small}\}} \sum_{i \in I_{n,k}} \varphi'_i > \psi^*(2^{r'n}) \right) & \leq C 2^{(1-r)n} \frac{\psi^*(2^{r'n})}{\psi^*(2^{r'n})} \\ & \leq C 2^{-[\kappa - (1-r)]n}. \end{aligned}$$

The right hand side is summable, so that we can use Borel–Cantelli Lemma. Almost surely, for all large enough integer  $n$ ,

$$\sum_{k=0}^{2^{(1-q)n}-1} \left( 1_{\{(X'_i, \varphi'_i)_{i \in I_{n,k}} \text{ small}\}} \sum_{i \in I_{n,k}} \varphi'_i \right) \leq \psi^*(2^{r'n}).$$

We sum this inequality for  $n$  going from 0 to  $\lfloor \log_2 N \rfloor + 1$ , and finish as in the proof of Lemma 3.3. Almost surely, for all large enough integer  $N$ ,

$$\left| \sum_{i=0}^{N-1} \varphi'_i - \sum_{i=0}^{N-1} \varphi_i^{(1)} \right| \leq C \lfloor \log_2 N \rfloor \psi^*(N^{r'}) \leq \psi^*(N^{r^*}).$$

□

**Lemma 3.7** Assume that there exist  $p > 2$  and a constant  $C$  such that, for all  $n$ ,

$$\mathbb{E} \left( \sup_{i < 2^{qn}} \left| \sum_{\ell=2^n}^{2^n+i} X_\ell^* \right|^p \right) \leq C 2^{\frac{pqn}{2}}.$$

Let  $z > 0$  such that  $1 + (p - 2)q/2 < p(z + r - 1)$ . Almost surely, for all large enough integer  $N$ ,

$$\left| \sum_{i=0}^{N-1} X'_i - \sum_{i=0}^{N-1} X_i^{(2)} \right| \leq N^z. \tag{3.13}$$

*Proof* Heuristically, this lemma means that, for large values of  $n$ , most blocks are small; hence, the large blocks have little influence upon the partial sums of the  $X_i$ . Naturally, we need a bound on the number of large blocks. Let  $n$  be an integer. We denote by  $v_n^{(j)}$  the number of large blocks of the process  $(X_i^{(j)}, \varphi_i^{(j)})_{2^n \leq i < 2^{n+1}}$ ; this variable has a binomial distribution  $B(M, P)$  with parameters  $M = 2^{(1-q)n}$  and a probability  $P$  which can be bounded by  $2^{qn} \mathbb{P}(\varphi \geq \psi^*(2^{rn})) \leq 2^{-(r-q)n}$ . Thanks to Markov inequality, we show an exponential bound on the number of large blocks:

$$\begin{aligned} \mathbb{P}(v_n^{(j)} > 2^{(1-r)n+1}) &\leq e^{-2^{(1-r)n+1}} \mathbb{E}(e^{v_n^{(j)}}) \\ &\leq e^{-2^{(1-r)n+1}} (1 + (e - 1)2^{(q-r)n}) 2^{(1-q)n} \\ &\leq e^{-2^{(1-r)n+1}} e^{(e-1)2^{(q-r)n+(1-q)n}} \\ &= e^{-(3-e)2^{(1-r)n}}. \end{aligned}$$

Let  $M_n^{(j)} := \sup_{k < 2^{(1-q)n}} \sup_{i \in I_{n,k}} \left| \sum_{\ell=2^n+k2^{qn}}^i X_\ell^{(j)} \right|$ . Let us take an integer  $N$ , and let  $n \leq \lfloor \log_2 N \rfloor + 1$ . On each block  $I_{n,k}$ , the processes  $X'$  and  $X^{(2)}$  differ only if both  $X^{(1)}$  and  $X^{(2)}$  are large, and in this case the sums  $\sum_{\ell=2^n+k2^{qn}}^i X'_\ell$  and  $\sum_{\ell=2^n+k2^{qn}}^i X_\ell^{(2)}$  differ at most by  $M_n^{(1)} + M_n^{(2)}$ . Hence:

$$\left| \sum_{i=0}^{N-1} X'_i - \sum_{i=0}^{N-1} X_i^{(2)} \right| \leq \sum_{n=0}^{\lfloor \log_2 N \rfloor + 1} (v_n^{(1)} + v_n^{(2)})(M_n^{(1)} + M_n^{(2)}).$$

Let  $0 < z' < z$  be such that the inequality  $1 + (p - 2)q/2 < p(z' + r - 1)$  remains true. We compute:

$$\begin{aligned} &\mathbb{P}((v_n^{(1)} + v_n^{(2)})(M_n^{(1)} + M_n^{(2)}) > 2^{z'n}) \\ &\leq 2\mathbb{P}(v_n^{(j)} > 2^{(1-r)n+1}) + 2\mathbb{P}\left(M_n^{(j)} > \frac{2^{z'n} 2^{-(1-r)n}}{8}\right) \\ &\leq 2e^{-(3-e)2^{(1-r)n}} + 2 \cdot 2^{(1-q)n} \mathbb{P}\left(\sup_{i < 2^{qn}} \left| \sum_{\ell=2^n}^{2^n+i} X_\ell^{(j)} \right| > \frac{2^{(z'+r-1)n}}{8}\right) \\ &\leq 2e^{-(3-e)2^{(1-r)n}} + C2^{(1-q)n} \frac{\mathbb{E}\left(\sup_{i < 2^{qn}} \left| \sum_{\ell=2^n}^{2^n+i} X_\ell^{(j)} \right|^p\right)}{2^{p(z'+r-1)n}} \\ &\leq 2e^{-(3-e)2^{(1-r)n}} + C2^{(1-q)n} 2^{\frac{pqn}{2}} 2^{-p(z'+r-1)n}. \end{aligned}$$

This quantity is summable in  $n$ . Thanks to Borel–Cantelli Lemma, we deduce that almost surely, for all large enough integer  $N$ ,

$$\left| \sum_{i=0}^{N-1} X'_i - \sum_{i=0}^{N-1} X_i^{(2)} \right| \leq CN^{z'}.$$

Up to taking even larger values of  $N$ , the inequality (3.13) holds.

In order to finish the proof of Proposition 3.5, we need to choose the parameters  $r$  (the threshold parameters used in the coupling) and  $z$  in a suitable way. We recall that  $r$  belongs to  $(0, 1)$  and  $z$  to  $(0, 1/2)$  (so that Lemma 3.7 be non-trivial). We only need to check that those two parameters can be chosen such that:

$$1 + \frac{p-2}{2}q < p(z+r-1).$$

If we take  $r = 1$  and  $z = 1/2$ , and since  $p$  is strictly larger than 2, this is true for all  $q \in (0, 1)$ . This inequality holds if we choose  $r$  close enough to 1 and  $z$  close enough to 1/2.

### 3.4 End of the proof

Now, we want to end the proof of Theorem 1.7. Let  $X$  and  $\varphi$  be functions satisfying the hypotheses of Theorem 1.7. For all  $i$ , we write  $(X_i, \varphi_i) = (X \circ T^i, \varphi \circ T^i)$ . We want to couple  $(X_i, \varphi_i)$  with a process  $(\tilde{X}_i, \tilde{\varphi}_i)$  such that  $(\tilde{X}_i)$  and  $(\tilde{\varphi}_i)$  have the same distribution as  $(X_i)$  and  $(\varphi_i)$  respectively, and such that  $(\tilde{X}_i)$  and  $(\tilde{\varphi}_i)$  are independent. Let  $q$  and  $\varepsilon$  be in  $(0, 1)$ . Let  $(\tilde{X}_i)$  and  $(\tilde{\varphi}_i)$  be processes with the same distribution as  $(X_i)$  and  $(\varphi_i)$  respectively. Here are the different steps of the proof, with the lemmas we use at each step:

- we couple the process  $(X_i, \varphi_i)$  with a process  $(X_i^*, \varphi_i^*)$  which is piecewise i.i.d. with parameters  $q$  and  $\varepsilon$ , thanks to Lemma 3.2. We use Lemmas 3.3 and 3.4 to check that  $(X_i^*, \varphi_i^*)$  is asymptotically close to the original process.
- we use Lemma 3.2 to couple  $(\tilde{X}_i)$  with a process  $(\bar{X}_i)$  which is piecewise i.i.d. with parameters  $q$  and  $\varepsilon$ , and check with Lemma 3.4 that these two processes are asymptotically close.
- we use Lemma 3.2 to couple  $(\tilde{\varphi}_i)$  with a process  $(\bar{\varphi}_i)$  which is piecewise i.i.d. with parameters  $q$  and  $\varepsilon$ , and check with Lemma 3.3 that these two processes are asymptotically close.
- we couple the two random variables  $(\tilde{X}_i, \bar{X}_i)$  and  $(\tilde{\varphi}_i, \bar{\varphi}_i)$  in the trivial (independent) way. This yields a coupling between  $(\tilde{X}_i, \tilde{\varphi}_i)$  and  $(\bar{X}_i, \bar{\varphi}_i)$ . In addition,  $(\tilde{X}_i)$  and  $(\tilde{\varphi}_i)$  are independent, so are  $(\bar{X}_i)$  and  $(\bar{\varphi}_i)$ , the processes  $(\tilde{X}_i)$  and  $(X_i^*)$  have the same distribution, and so do the processes  $(\tilde{\varphi}_i)$  and  $(\varphi_i^*)$ .
- we couple the process  $(X_i^*, \varphi_i^*)$  and the process  $(\bar{X}_i, \bar{\varphi}_i)$  according to Proposition 3.5.

### 4 Limit theorems

We now prove the convergence in distribution we claimed in the first section, Theorems 1.11 and 1.12, as well as an almost everywhere bound which holds in the setting of Theorem 1.11.

In this section, we consider a dynamical system  $(\Omega, \mu, T)$  which induces a Gibbs–Markov map on a Borel set  $A$ , with  $\mu(A) = 1$ . Moreover, we assume that the return times in  $A$  satisfy  $\mu|_A(\varphi_A \geq N) = 1/\psi(N)$  for some function  $\psi$  on  $\mathbb{R}_+$  with regular variation of index  $\beta \in [0, 1]$  at infinity. It will be clear that Theorem 1.12 does not need the full strength of these assumptions; we will discuss it in Sect. 4.3.

#### 4.1 Proof of the generalized CLT

Before proving our second main theorem, we must show a few technical lemmas. The first one deals with the limit of some processes indexed by a random time.

**Lemma 4.1** *Let  $(A_N, B_N)_{N \in \mathbb{N}}$  be a sequence of random variables such that the  $A_N$  are nonnegative integers and the  $B_N$  belong to some Polish space  $\Omega$ . We assume that:*

- *for some sequence  $(a_N)$  such that  $\lim_{N \rightarrow +\infty} a_N = +\infty$ , the process  $(A_N/a_N)$  converges in distribution to some random variable  $\mathcal{A}$  on  $\mathbb{R}_+$ ;*
- *$\mathbb{P}(\mathcal{A} > 0) = 1$ ;*
- *the process  $(B_N)$  converges in distribution to some random variable  $\mathcal{B}$  on  $\Omega$ ;*
- *the processes  $(A_N)$  and  $(B_N)$  are independent.*

*Then  $(A_N/a_N, B_{A_N})$  converges in distribution to a random variable  $(\mathcal{A}^*, \mathcal{B}^*)$  whose distribution is the same as  $(\mathcal{A}, \mathcal{B})$ ; in particular, the two components of the limit distribution are independent.*

*Proof* Obviously, the sequence of random variables  $(A_N/a_N, B_{A_N})$  is tight, and thus has weak limit points.

Let  $0 < m < M$ ; put  $I := (m, M)$  and let  $J$  be a subset of  $\Omega$ . Assume that  $\mathbb{P}(\mathcal{A} \in \partial I) = \mathbb{P}(\mathcal{B} \in \partial J) = 0$ . Let  $\varepsilon > 0$ . Since  $(B_N)$  converges to  $\mathcal{B}$  and  $a_N$  goes to  $+\infty$ , for large enough  $N$ , one has for all  $k \geq ma_N$ :

$$|\mathbb{P}(B_k \in J) - \mathbb{P}(\mathcal{B} \in J)| \leq \varepsilon.$$

Up to taking larger  $N$ , we can also assume that:

$$|\mathbb{P}(A_N/a_N \in I) - \mathbb{P}(\mathcal{A} \in I)| \leq \varepsilon.$$

Then we deduce that, for large enough  $N$ :

$$\begin{aligned} & |\mathbb{P}(A_N \in a_N I, B_{A_N} \in J) - \mathbb{P}(\mathcal{A} \in I)\mathbb{P}(\mathcal{B} \in J)| \\ & \leq \left| \sum_{k \in a_N I} \mathbb{P}(A_N = k, B_k \in J) - \mathbb{P}(A_N \in a_N I)\mathbb{P}(\mathcal{B} \in J) \right| \end{aligned}$$

$$\begin{aligned}
 &+ |\mathbb{P}(A_N \in a_N I) \mathbb{P}(\mathcal{B} \in J) - \mathbb{P}(\mathcal{A} \in I) \mathbb{P}(\mathcal{B} \in J)| \\
 &\quad \sum_{k \in a_N I} \mathbb{P}(A_N = k) |\mathbb{P}(B_k \in J) - \mathbb{P}(\mathcal{B} \in J)| \\
 &+ |\mathbb{P}(A_N \in a_N I) - \mathbb{P}(\mathcal{A} \in I)| \mathbb{P}(\mathcal{B} \in J) \\
 \leq &\varepsilon \sum_{k \in a_N I} \mathbb{P}(A_N = k) + \varepsilon \mathbb{P}(\mathcal{B} \in J) \leq 2\varepsilon.
 \end{aligned}$$

Hence, any limit distribution  $\mu$  of  $(A_N/a_N, B_{A_N})$  satisfies  $\mu((m, M) \times J) = \mathbb{P}(\mathcal{A} \in (m, M))\mathbb{P}(\mathcal{B} \in J)$  for any continuity sets  $(m, M)$  and  $J$ . Since most balls are continuity sets, such boxes  $(m, M) \times J$  generate the Borel sigma-algebra on  $\mathbb{R}_+ \times \Omega$ , so that  $\mu$  is a product measure.

The next result describes the limit behavior of local times for the maps which induce a Gibbs–Markov map. It is folklore if the parameter  $\beta$  belongs to  $[0, 1)$ , and all the ingredients can then be found in Jon Aaronson’s book [1]. The version we present also deals with the case  $\beta = 1$ .

**Proposition 4.2** *Let  $(\Omega, \mu, T)$  be an ergodic dynamical system which induces a Gibbs–Markov map on a Borel set  $A$ . Let  $f$  be in  $\mathbb{L}^1(\Omega, \mu)$  and  $\nu \ll \mu$  be a probability measure.*

*If the random variable  $\varphi_A$  on  $(A, \mu|_A)$  fulfills the condition (1.8) for some  $\beta \in [0, 1)$ , then:*

$$\frac{1}{\text{sinc}(\beta\pi)\psi(N)} \sum_{i=0}^{N-1} f \circ T^i \rightarrow \int f \, d\mu \cdot Y_\beta, \tag{4.1}$$

where the convergence is in distribution when the left-hand side is seen as a random variable from  $(\Omega, \nu)$  to  $\mathbb{R}$ .

*If the random variable  $\varphi_A$  on  $(A, \mu|_A)$  fulfills the condition (1.8) with  $\beta = 1$ , then:*

$$\frac{1}{N \sum_{k=0}^{N-1} \frac{1}{\psi(k)}} \sum_{i=0}^{N-1} f \circ T^i \rightarrow \int f \, d\mu, \tag{4.2}$$

where the convergence is in distribution when the left-hand side is seen as a random variable from  $(\Omega, \nu)$  to  $\mathbb{R}$ .

*Proof* By Lemma 6.5 in [8], the analytic family of renewal operators  $T(z)(\nu) = \mathcal{L}(z^{\varphi_A} \nu)$  acting on  $\mathbb{L}^1(A, \mu|_A)$  behave nicely when  $|z| < 1$ . Hence, the conditions  $(H1')$  and  $(H2)$  of [16] are satisfied. Then, the conclusions of Theorem 3.6 in [16] hold. This gives us the pointwise dual ergodicity of the system with the normalizing sequences we used. The Corollary 3.7.3 in [1] yields the conclusion. In the case  $\beta = 1$ , the random variable  $Y_1$  does not appear because its value is 1 almost surely.

Then, we define a pair of useful objects: the local time in a Borel subset, and the return times.

**Definition 4.3** (*Local time and return time*) For any  $N$  in  $\mathbb{N}$  and  $x$  in  $\Omega$ , we define the local time in  $A$  at time  $N$  by:

$$\xi_N(x) := \sum_{i=1}^N 1_A \circ T^i x = \text{Card}\{1 \leq i \leq N : T^i x \in A\}. \tag{4.3}$$

In addition, if  $x$  belongs to  $A$ , then we define for any  $N$  in  $\mathbb{N}$  the return times by:

$$t_N(x) := \sum_{i=0}^{N-1} \varphi_A \circ T_A^i x. \tag{4.4}$$

The following lemma allows us to control in a locally uniform way the growth of the local time. It states, roughly, that under the usual assumption on the return times, the local time  $\xi_N$  at time  $N$  is of order  $\psi(N)$ .

**Lemma 4.4** *Let  $(\Omega, \mu, T)$  be a dynamical system which induces an mixing Gibbs–Markov map on a Borel set  $A$ . Assume that the random variable  $\varphi_A$  on  $(A, \mu|_A)$  fulfills the condition (1.8). Let  $r > 0$  and  $r^* > r$ . Then,  $\mu|_A$ -almost surely, for all large enough integer  $N$ ,*

$$\sup_{k \leq N} (\xi_{k+\psi^*(\psi(N)^r)} - \xi_k) \leq \psi(N)^{r^*}. \tag{4.5}$$

*Proof* Most of this proof comes from the proof of Lemma 3.5 in [5]. We use a probabilistic notation, all the probabilities being taken with respect to the measure  $\mu|_A$ . For any  $n$  in  $\mathbb{N}$ , we put:

$$\mathbb{P}(n) := \mathbb{P}\left(\sup_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k) > (n-1)^{r^*}\right).$$

Since the functions  $\psi(k)$  and  $\xi_k$  are nondecreasing in  $k$ , we see that if  $\sup_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k) \leq (n-1)^{r^*}$  then, for all  $\psi^*(n-1) < N \leq \psi^*(n)$ , we have:

$$\begin{aligned} \sup_{k \leq N} (\xi_{k+\psi^*(\psi(N)^r)} - \xi_k) &\leq \sup_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k) \\ &\leq (n-1)^{r^*} \leq \psi(N)^{r^*}. \end{aligned}$$

Hence, it is enough to prove that  $\sup_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k) \leq (n-1)^{r^*}$  holds for all large enough  $n$ , and by Borel–Cantelli lemma we only have to ensure that  $\mathbb{P}(n)$  is summable. In addition, the maximum  $\sup_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k)$  is reached just before a jump of  $\xi_k$ , when  $k = t_\ell - 1$  for some  $\ell$ . This is due to the fact that as long as  $\xi_k$  is constant (in other words, for  $t_{\ell-1} \leq k \leq t_\ell - 1$  for some  $\ell$ ), the function  $\xi_{k+\psi^*(n^r)} - \xi_k$  is non-decreasing. By going one step further (which increases  $\xi_k$  by 1), we get:

$$\sup_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k) \leq \sup_{k: t_k \leq \psi^*(n)} (\xi_{t_k+\psi^*(n^r)} - \xi_{t_k}) + 1.$$

Since we can decrease ever so slightly the value of  $r^*$  and take arbitrarily large values of  $n$ , we shall ignore the  $+1$  term. Now, we compute:

$$\begin{aligned} \mathbb{P}(n) &\leq \mathbb{P}\left(\sup_{k: t_k \leq \psi^*(n)} (\xi_{t_k+\psi^*(n^r)} - \xi_{t_k}) > (n-1)^{r^*}\right) \\ &\leq \mathbb{P}\left(\sup_{k: t_k \leq \psi^*(n)} (\xi_{t_k+\psi^*(n^r)} - \xi_{t_k}) > (n-1)^{r^*}, \xi_{\psi^*(n)} \leq n^2\right) \\ &\quad + \mathbb{P}\left(\xi_{\psi^*(n)} > n^2\right), \end{aligned}$$

and, since the measure  $\mu|_A$  is  $T_A$ -invariant, we get:

$$\mathbb{P}(n) \leq n^2 \mathbb{P}\left(\xi_{\psi^*(n^r)} > (n-1)^{r^*}\right) + \mathbb{P}\left(\xi_{\psi^*(n)} > n^2\right).$$

Now, notice that:

$$\mathbb{P}\left(\xi_{\psi^*(n^r)} > (n-1)^{r^*}\right) = \mathbb{P}\left(\sum_{k=0}^{(n-1)^{r^*}-1} \varphi_A \circ T_A^k < \psi^*(n^r)\right).$$

Let  $q$  in  $(0, 1)$  be such that  $(1-q)r^* > r$ . We shall consider only the values of  $\varphi_A \circ T_A^{k(n-1)^{qr^*}}$  for  $k$  going from 0 to  $(n-1)^{(1-q)r^*} - 1$ , so as to gain some independence:

$$\begin{aligned} &\mathbb{P}\left(\sum_{k=0}^{(n-1)^{r^*}-1} \varphi_A \circ T_A^k < \psi^*(n^r)\right) \\ &\leq \mathbb{P}\left(\sup_{0 \leq k < (n-1)^{(1-q)r^*}-1} \varphi_A \circ T_A^{k(n-1)^{qr^*}} < \psi^*(n^r)\right). \end{aligned}$$

The function  $\varphi_A$  is constant on each set  $a$  of the Gibbs–Markov partition of  $A$ . Let  $(\varphi'_k)_{k \geq 0}$  be a family of independent, identically distributed random variables with the same law as  $\varphi_A$ . As we did in the proof of Lemma 3.2, we can couple the family of random variables  $(\varphi_A \circ T_A^{k(n-1)^{qr^*}})_{0 \leq k < (n-1)^{(1-q)r^*}-1}$  with  $(\varphi'_k)_{k \geq 0}$ , with an exponential bound on their differences. For some constants  $\rho > 1$  and  $C, C'$  and for all large enough  $n$ ,

$$\mathbb{P}\left(\sup_{0 \leq k < (n-1)^{(1-q)r^*}-1} |\varphi_A \circ T_A^{k(n-1)^{qr^*}} - \varphi'_k| > C'\right) \leq C(n-1)^{(1-q)r^*} \rho^{-(n-1)^{qr^*}},$$

which yields:

$$\begin{aligned} & \mathbb{P} \left( \sup_{0 \leq k < (n-1)^{(1-q)r^*} - 1} \varphi_A \circ T_A^{k(n-1)^{qr^*}} < \psi^*(n^r) \right) \\ & \leq \mathbb{P} \left( \sup_{0 \leq k < (n-1)^{(1-q)r^*} - 1} \varphi'_k < \psi^*(n^r) + C' \right) + C(n-1)^{(1-q)r^*} \rho^{-(n-1)^{qr^*}}. \end{aligned}$$

We can always re-write the argument with a slightly larger value of  $r$  to dominate the constant  $C'$ , so we shall forget about the later. We get:

$$\begin{aligned} & \mathbb{P} \left( \sum_{k=0}^{(n-1)^{r^*} - 1} \varphi_A \circ T_A^k < \psi^*(n^r) \right) \\ & \leq \mathbb{P} (\varphi_A < \psi^*(n^r))^{(n-1)^{(1-q)r^*}} + C(n-1)^{(1-q)r^*} \rho^{-(n-1)^{qr^*}} \\ & = \left( 1 - \frac{1}{\psi(\psi^*(n^r))} \right)^{(n-1)^{(1-q)r^*}} + C(n-1)^{(1-q)r^*} \rho^{-(n-1)^{qr^*}} \\ & \leq e^{-\frac{(n-1)^{(1-q)r^*}}{n^r}} + C(n-1)^{(1-q)r^*} \rho^{-(n-1)^{qr^*}}. \end{aligned}$$

The right hand side is summable; by taking  $r = 1$  and  $r^* = 2$  we see that  $n^2 \mathbb{P} (\xi_{\psi^*(n)} \geq n^2)$  is also summable. Hence,  $\mathbb{P}(n)$  is summable.

We shall also need the same kind of control on the partial sum of  $X_f$ , where  $f$  is an observable on  $\Omega$  such that  $\int_{\Omega} f \, d\mu = 0$ .

**Lemma 4.5** *Let  $(A, d, \mu|_A, T_A)$  be a mixing Gibbs–Markov map. Let  $X \in \mathbb{L}^p(A, \mu|_A)$ , with  $p > 2$ , be such that  $\int_A X \, d\mu|_A = 0$  and the regularity condition (1.9) is satisfied.*

*For all  $r \in [0, 1]$  and all  $r^* > 2/p + (1 - 2/p)r$ , almost surely, for all large enough integer  $N$ ,*

$$\sup_{k \leq N} \sup_{i \leq N^r} \left| \sum_{j=k}^{k+i-1} X \circ T_A^j \right| \leq N^{\frac{r^*}{2}}. \tag{4.6}$$

*Proof* We begin our proof by cutting the set of integers into pieces of length  $2^n$ , where  $n$  belongs to  $\mathbb{N}$ . Each of these pieces shall be cut in  $2^{(1-r)n}$  segments of length  $2^{rn}$ , on which we can control the variation of the partial sums of  $X \circ T^i$ . In other words, we have for all  $r' > r$ :

$$\begin{aligned} & \mathbb{P} \left( \sup_{2^n \leq k < 2^{n+1}} \sup_{i \leq 2^{rn}} \left| \sum_{j=k}^{k+i-1} X \circ T^j \right| \geq 2^{\frac{r'(n-1)}{2}} \right) \\ & \leq \mathbb{P} \left( \sup_{k < 2^{(1-r)n}} \sup_{i \leq 2^{rn}} \left| \sum_{j=2^n+k2^{rn}}^{2^n+k2^{rn}+i-1} X \circ T^j \right| \geq \frac{2^{\frac{r'(n-1)}{2}}}{3} \right) \end{aligned}$$



$$\begin{aligned} &\leq 2^{(1-r)n} \mathbb{P} \left( \sup_{i \leq 2^{rn}} \left| \sum_{j=0}^{i-1} X \circ T^j \right| \geq \frac{2^{\frac{r'(n-1)}{2}}}{3} \right) \\ &\leq C 2^{(1-r)n} 2^{-\frac{r'pn}{2}} \mathbb{E} \left( \sup_{i \leq 2^{rn}} \left| \sum_{j=0}^{i-1} X \circ T^j \right|^p \right). \end{aligned}$$

By Rosenthal inequality 1.5, there exists a constant  $C$  such that, for all  $N$ ,

$$\mathbb{E} \left( \left| \sum_{i=0}^{N-1} X \circ T^i \right|^p \right) \leq CN^{\frac{p}{2}}.$$

We use again Corollary B1 in [18]:

$$\mathbb{E} \left( \sup_{i \leq 2^{rn}} \left| \sum_{j=0}^{i-1} X \circ T^j \right|^p \right) \leq C 2^{\frac{rpn}{2}}.$$

Hence, for some constant  $C$ , we have for all integer  $n$ :

$$\mathbb{P} \left( \sup_{2^n \leq k < 2^{n+1}} \sup_{i \leq 2^{rn}} \left| \sum_{j=k}^{k+i-1} X \circ T^j \right| \geq 2^{\frac{r'n}{2}} \right) \leq C 2^{(1-r)n} 2^{-\frac{(r'-r)pn}{2}}.$$

Now, assume that  $r'$  belongs to  $(2/p + (1 - 2/p)r, r^*)$ . The right-hand side of this inequality is summable, and by Borel–Cantelli lemma, almost surely there exists a constant  $C$  such that:

$$\sup_{k \leq N} \sup_{i \leq N^r} \left| \sum_{j=k}^{k+i-1} X \circ T^j \right| \leq CN^{\frac{r'}{2}}.$$

Since  $r' < r^*$ , the lemma is proved.

At last, we have all the tools we need to prove Theorem 1.11. For the convenience of our readers, we recall that this theorem states that, for dynamical systems which induce a mixing Gibbs–Markov map and for which the inverse of the tail of the return time has regular variation of order  $\beta \in [0, 1)$ , the partial sum  $\sum_{i=0}^{N-1} f \circ T^i$  of a nice function (i.e., smooth enough, and integrable enough) with zero average is of order  $\sqrt{\psi(n)}$ . Moreover, as  $N$  goes to infinity, the following convergence in law (with respect to any absolutely continuous probability distribution) occurs:

$$\frac{1}{\sqrt{\text{sinc}(\beta\pi)\psi(N)}} \sum_{i=0}^{N-1} f \circ T^i \rightarrow \sigma(f)\sqrt{Y_\beta}\mathcal{N},$$

where  $Y_\beta$  is a Mittag–Leffler random variable of order  $\beta$ , where  $\mathcal{N}$  is a standard Gaussian random variable, and where  $Y_\beta$  and  $\mathcal{N}$  are independent.

*Proof of Theorem 1.11* So as to manipulate lighter expressions, we put:

$$K := \frac{1}{\sqrt{\text{sinc}(\beta\pi)}}.$$

We shall proceed in three steps. In the first, we assume that  $x$  is chosen in  $A$  according to the distribution  $\mu|_A$ , and get a limit theorem for the induced transformation. In the second, we will stick with the assumption that  $x$  is chosen in  $A$  according to the distribution  $\mu|_A$  and get our limit theorem. At last, we will get the theorem for any starting distribution which is absolutely continuous with respect to  $\mu$ .

**First step: Sums for the induced transformation on A**

We choose  $x$  in  $A$  with probability  $\mu|_A$ . Let  $(\tilde{X}_i, \tilde{t}_i)$  be a process such that  $(\tilde{X}_i)$  and  $(X_f \circ T_A^i)$  have the same law (as processes),  $(\tilde{t}_i)$  and  $(\varphi_A \circ T_A^i)$  have the same law, and  $(\tilde{X}_i)$  and  $(\tilde{t}_i)$  are independent. If we denote by  $\tilde{\xi}_i = (\tilde{\Phi})^*(i)$  the generalized inverse of  $\tilde{\Phi}(N) := \sum_{i=0}^{N-1} \tilde{t}_i$ , then  $(\tilde{\xi}_i)$  has the same law as  $(\xi_i)$  and is independent of  $(\tilde{X}_i)$ . We couple the processes  $(\tilde{X}_i, \tilde{t}_i)$  and  $(X_f \circ T_A^i, \varphi_A \circ T_A^i)$  in a way satisfying the conclusions of Theorem 1.7. By Proposition 1.4, Proposition 4.2 and Lemma 4.1:

$$K \left( \frac{\tilde{\xi}_N}{\psi(N)} \right)^{\frac{1}{2}} \frac{1}{\sqrt{\tilde{\xi}_N}} \sum_{\ell=0}^{\tilde{\xi}_N-1} \tilde{X}_\ell \rightarrow \sigma(f)\sqrt{Y_\beta}\mathcal{N},$$

where the convergence is in distribution,  $Y_\beta$  and  $\mathcal{N}$  are independent,  $Y_\beta$  is a standard Mittag–Leffler distribution of order  $\beta$  and  $\mathcal{N}$  is a standard Gaussian random variable, and where the variance  $\sigma(f)^2$  is given by Eq. (1.13). In addition,  $\sigma(f) = 0$  if and only if  $X_f$  is a coboundary, hence if and only if  $f$  is a coboundary.

Moreover, as stated by Theorem 1.7, for some  $r < 1$ , almost surely, for all large enough  $n$ :

$$\frac{K}{\sqrt{\psi(N)}} \left| \sum_{\ell=0}^{\tilde{\xi}_N-1} \tilde{X}_\ell - X_f \circ T_A^\ell \right| \leq \frac{K}{\psi(N)^{\frac{1-r}{2}}} \left( \frac{\tilde{\xi}_N}{\psi(N)} \right)^{\frac{r}{2}},$$

which tends in probability to 0 by Proposition 4.2. Therefore,

$$\frac{K}{\sqrt{\psi(N)}} \sum_{\ell=0}^{\tilde{\xi}_N-1} X_f \circ T_A^\ell \rightarrow \sigma(f)\sqrt{Y_\beta}\mathcal{N}.$$

Now, notice that for all  $N$ , we have  $\xi_{t_N} = N$  and  $t_{\xi_N} \leq N$ , so that:

$$\tilde{\xi}_N - \xi_N = \xi_{t_{\tilde{\xi}_N}} - \xi_N \leq \xi_{t_{\xi_N}} - \xi_{t_{\xi_N}}.$$

Let  $\kappa > 1$ . By Lemma 4.4 (take  $r = 1$ ), we know that almost surely, for all large enough integer  $N$ , the function  $\xi_N$  (respectively  $\tilde{\xi}_N$ ) is bounded by  $\psi(N)^\kappa$ . Since the coupling between  $t_N$  and  $\tilde{t}_N$  satisfies the conclusions of Theorem 1.7, almost surely, for all large enough integer  $N$ :

$$\left| t_{\xi_N} - \tilde{t}_{\tilde{\xi}_N} \right| \leq \psi^*(\psi(N)^{\kappa r}).$$

Since this is true for all  $\kappa > 1$ , for all  $r' \in (r, 1)$ , the difference  $\left| t_{\xi_N} - \tilde{t}_{\tilde{\xi}_N} \right|$  is bounded by  $\psi^*(\psi(N)^{r'})$  for all large enough  $N$ . By Lemma 4.4, for all  $r^* \in (r, 1)$  and for all large enough  $N$ , we have:

$$\tilde{\xi}_N - \xi_N \leq \psi(N)^{r^*}.$$

We deal with  $\xi_N - \tilde{\xi}_N$  in the same way, and we obtain  $\left| \xi_N - \tilde{\xi}_N \right| \leq \psi(N)^{r^*}$  for all  $r^* \in (r, 1)$  and all large enough  $N$ . Let  $\kappa > 1$ , and  $r' \in (2/p + (1 - 2/p)r^*, 1)$ . By Lemma 4.5, we get for all large enough  $N$ :

$$\begin{aligned} & \frac{K}{\sqrt{\psi(N)}} \left| \sum_{\ell=0}^{\tilde{\xi}_N-1} X_f \circ T_A^\ell - \sum_{\ell=0}^{\xi_N-1} X_f \circ T_A^\ell \right| \\ & \leq \frac{K}{\sqrt{\psi(N)}} \sup_{k \leq \psi(N)^\kappa} \sup_{i \leq \psi(N)^{r^*}} \left| \sum_{\ell=k}^{k+i-1} X_f \circ T_A^\ell \right| \leq K \psi(N)^{\frac{r'\kappa-1}{2}}. \end{aligned}$$

If we take  $\kappa$  close enough to 1, this converges to 0 as  $N$  goes to infinity. Hence, we have:

$$\frac{K}{\sqrt{\psi(N)}} \sum_{\ell=0}^{\xi_N-1} X_f \circ T_A^\ell \rightarrow \sigma(f) \sqrt{Y_\beta} \mathcal{N}. \tag{4.7}$$

**Second step: Limit sums with starting point in A**

We still choose  $x$  in  $A$  with probability  $\mu|_A$ . We want to get rid of the return times in the expression (4.7).

For any  $\ell \geq 0$  and any  $\varepsilon > 0$ , we have:

$$\mathbb{P}(X_{|f|} \circ T_A^\ell > \ell^{\frac{1+\varepsilon}{p}}) \leq C \ell^{-(1+\varepsilon)}.$$

Hence, by Borel–Cantelli lemma, almost surely, there exists a constant  $C$  such that  $X_{|f|} \circ T_A^\ell \leq C \ell^{\frac{1+\varepsilon}{p}}$ . If we replace  $\ell$  by  $\xi_N - 1$ , we obtain that almost surely, there exists a constant  $C$  such that  $X_{|f|} \circ T_A^{\xi_N-1} \leq C \xi_N^{\frac{1+\varepsilon}{p}}$ . By Lemma 4.4, for any  $\kappa > 1$ , almost surely, for all large enough  $n$ , we have  $X_{|f|} \circ T_A^{\xi_N-1} \leq C \psi(N)^{\frac{(1+\varepsilon)\kappa}{p}}$ . This gives us the following bounds for large enough values of  $N$ :

$$\begin{aligned} \frac{1}{\sqrt{\psi(N)}} \left| \sum_{i=\ell_{\varepsilon_N}}^{N-1} f \circ T^i \right| &\leq \frac{1}{\sqrt{\psi(N)}} X_{|f|} \circ T_A^{\xi_{N-1}} \\ &\leq C\psi(N)^{\frac{(1+\varepsilon)\kappa}{p}-\frac{1}{2}}. \end{aligned}$$

If we take  $\varepsilon$  close enough to 0 and  $\kappa$  close enough to 1, the expression above converges almost surely to 0 as  $N$  goes to  $+\infty$ . Hence, we get:

$$\begin{aligned} &\frac{K}{\sqrt{\psi(N)}} \sum_{i=0}^{N-1} f \circ T^i \\ &= \frac{K}{\sqrt{\psi(N)}} \left( \sum_{\ell=0}^{\xi_{N-1}} X_f \circ T_A^\ell + \sum_{i=\ell_{\varepsilon_N}}^{N-1} f \circ T^i \right) \rightarrow \sigma(f)\sqrt{Y_\beta}\mathcal{N}, \end{aligned} \tag{4.8}$$

where the convergence is in distribution when the left hand side is seen as a real-valued function on the probability space  $(\Omega, \mu_{|A})$ .

**Third step: Strong distributional convergence**

This last step is a straightforward application of Theorem 1 in [21], with one hypothesis left to prove. Let  $N$  be an integer. We compute for any  $\varepsilon > 0$ :

$$\mu \left( \frac{K}{\sqrt{\psi(N)}} \left| \sum_{i=0}^{N-1} f \circ T^i - f \circ T \circ T^i \right| > \varepsilon \right) \leq 2\mu \left( |f| > C\sqrt{\psi(N)}\varepsilon \right),$$

and since  $f$  is in  $\mathbb{L}^1(\Omega, \mu)$  the right hand side converges to 0 as  $N$  goes to  $+\infty$ .

Hence, we can apply the conclusions of Theorem 1 in [21]: the convergence in equation (4.8) holds not only on  $(\Omega, \mu_{|A})$ , but on  $(\Omega, \nu)$  for any probability measure  $\nu \ll \mu$ .

*Remark 4.6* A closer look at this proof shows that the assumption  $X_{|f|} \in \mathbb{L}^p(A, \mu_{|A})$  and  $\int_\Omega f \, d\mu = 0$  in Theorem 1.11 can be replaced by the following weaker set of assumptions:

- $\sup_{i \leq \varphi_A} \left| \sum_{\ell=0}^{i-1} f \circ T^\ell \right| \in \mathbb{L}^p(A, \mu_{|A})$ ;
- there exists a  $M > 0$  such that  $\mu(|f| > M) < +\infty$ ;
- $\int_A X_f \, d\mu_{|A} = 0$ .

This improvement is enough to extend our limit theorems to “non-integrable functions with zero average”, which do not belong to  $\mathbb{L}^1(\Omega, \mu)$ . We give some examples at the end of Sect. 2.

### 4.2 Almost everywhere bound

We wish to present to the reader a last result which, while somewhat beyond the scope of this article, can be easily proved with the tools used in the previous sections and may improve our understanding of the behavior of the Birkhoff sums in infinite ergodic theory.

**Theorem 4.7** *Let  $(\Omega, \mu, T)$  be a dynamical system which induces a mixing Gibbs–Markov map on a Borel set  $A$ . Assume that the inverse of the tail of the random variable  $\varphi_A$  on  $(A, \mu|_A)$  is regularly varying with index  $\beta \in [0, 1)$ . Let  $f$  in  $\mathbb{L}^1(\Omega, \mu)$ , with  $\int_{\Omega} f \, d\mu = 0$ , be such that the random variable  $X_{|f|}$  belongs to  $\mathbb{L}^p(A, \mu|_A)$  for some  $p > 2$  and that condition (1.9) is satisfied.*

*Then, for any  $\varepsilon > 0$  and for  $\mu$ -almost every point  $x$  in  $\Omega$ :*

$$\lim_{N \rightarrow +\infty} \frac{1}{N^{\frac{\beta}{2} + \varepsilon}} \sum_{i=0}^{N-1} f(T^i x) = 0. \tag{4.9}$$

The proof of this theorem uses the same tools as the proof of Theorem 1.11, albeit in a much cruder fashion.

*Proof* For now, we choose  $x$  in  $A$  with probability  $\mu|_A$ . Almost surely,  $\xi_N$  increases to  $+\infty$ . By Lemma 4.5 (take  $r = 1$ ), this implies that for all  $\varepsilon > 0$ , almost surely for all large enough  $N$ ,

$$\left| \sum_{\ell=0}^{\xi_N-1} \left( \sum_{i=0}^{\varphi_A-1} f \circ T^i \right) \circ T_A^\ell \right| \leq \xi_N^{\frac{1}{2} + \varepsilon}. \tag{4.10}$$

By Lemma 4.4 (take  $r = 1$ ), for all  $\varepsilon > 0$ , almost surely for all large enough  $N$ ,

$$\xi_N \leq \psi(N)^{1+\varepsilon}. \tag{4.11}$$

Putting Eqs. (4.10) and (4.11) together, we get that for all  $\varepsilon > 0$ , almost surely for all large enough  $N$ ,

$$\left| \sum_{\ell=0}^{\xi_N-1} \left( \sum_{i=0}^{\varphi_A-1} f \circ T^i \right) \circ T_A^\ell \right| \leq \psi(N)^{\frac{1}{2} + \varepsilon}.$$

Then, we can copy the second step of the proof of Theorem 1.11; for all  $\varepsilon > 0$ , this gives us:

$$\left| \sum_{k=0}^{N-1} f \circ T^k \right| \leq \psi(N)^{\frac{1}{2} + \varepsilon},$$

almost surely for all large enough  $N$ .

Now, we want to get a result for  $\mu$ -almost every  $x$  in  $\Omega$ . Notice that some iterate  $T^k x$  of  $\mu$ -almost every  $x$  in  $\Omega$  lies in the subset of  $A$  (of full  $\mu|_A$  measure) of the points which satisfy the desired bound. Moreover, the sum of the  $f \circ T^i x$  before  $T^k x$  enters this set is also finite for  $\mu$ -almost every  $x$ . Hence, the theorem holds.

*Remark 4.8* We can again prove this theorem with a weaker set of assumption, which does not require  $f$  to be integrable. The hypotheses  $X_{|f|} \in \mathbb{L}^p(A, \mu|_A)$  and  $\int_{\Omega} f \, d\mu = 0$  in Theorem 4.7 can be replaced by:

- $\sup_{i \leq \varphi_A} \left| \sum_{\ell=0}^{i-1} f \circ T^\ell \right| \in \mathbb{L}^p(A, \mu|_A)$ ;
- $\int_A X_f \, d\mu|_A = 0$ .

The rate of convergence of Theorem 4.7 may be improved until one gets a law of iterated logarithm, as was done by Marcus and Rosen in [14] (Theorem 1.1) for random walks on  $\mathbb{Z}$  and  $\mathbb{Z}^2$ .

### 4.3 A generalized CLT with $\beta = 1$

We finish this article by proving Theorem 1.12. We cannot rely on asymptotic independence, so we shall use two tricks to prove the convergence in distribution. The first trick consists in noticing that a Mittag–Leffler random variable of parameter 1 is almost surely constant, and thus in a way independent from any other random variable. The second trick, used for instance in [3, Section 2.3], is used to manage the random fluctuations of the local time around its limit.

*Proof of Theorem 1.12* In the following, we use the notation  $\tilde{\psi}(N) := N \sum_{k=0}^{N-1} \frac{1}{\psi(k)}$ . Sums with negative range are defined by  $\sum_{n=a}^{b-1} u_n = -\sum_{n=b}^{a-1} u_n$  if  $a > b$ .

As in the proof of Theorem 1.11, we will proceed in three steps: we first get a limit theorem for the induced transformation, then a limit theorem when  $x$  is chosen in  $A$  according to the distribution  $\mu|_A$ , and at last a limit theorem with arbitrary absolutely continuous initial distribution. Since the last two steps can be copied from the proof of Theorem 1.11, we will only deal with the first argument. We want to prove that, under the law  $\mu|_A$ , the following convergence in distribution happens:

$$\frac{1}{\sqrt{\tilde{\psi}(N)}} \sum_{\ell=0}^{\xi_N-1} X_f \circ T_A^\ell \rightarrow \sigma(f)\mathcal{N}, \tag{4.12}$$

where  $\sigma(f)^2 = \int_A X_f^2 \, d\mu + 2 \sum_{i=1}^{+\infty} \int_A X_f \cdot X_f \circ T_A^i \, d\mu$  and  $\mathcal{N}$  is a standard Gaussian random variable. We divide the sum in the following way:

$$\frac{1}{\sqrt{\tilde{\psi}(N)}} \sum_{\ell=0}^{\xi_N-1} X_f \circ T_A^\ell = \frac{1}{\sqrt{\tilde{\psi}(N)}} \left( \sum_{\ell=0}^{\tilde{\psi}(N)-1} X_f \circ T_A^\ell + \sum_{\ell=\tilde{\psi}(N)}^{\xi_N-1} X_f \circ T_A^\ell \right).$$

Then, the Central Limit Theorem 1.4 applied to  $X_f$  yields the convergence in distribution:

$$\frac{1}{\sqrt{\tilde{\psi}(N)}} \sum_{\ell=0}^{\tilde{\psi}(N)-1} X_f \circ T_A^\ell \rightarrow \sigma(f)\mathcal{N},$$

and all we have to do is to bound the error term. By Proposition 4.2, the function  $\xi_N/\tilde{\psi}(N)$  converges in distribution to a random variable which is almost surely equal to 1. Hence, it converges in probability. Let  $\varepsilon > 0$ . For large enough  $N$ ,

$$\mathbb{P}(|\xi_n - \tilde{\psi}(N)| > \varepsilon \tilde{\psi}(N)) \leq \varepsilon.$$

By Rosenthal inequality 1.5, there is a constant  $C$  depending only on  $f$  and  $p$  such that, for all  $n$ ,

$$\mathbb{E} \left( \sup_{0 < k \leq n} \left| \sum_{\ell=0}^{k-1} X_f \circ T_A^\ell \right|^p \right) \leq Cn^{\frac{p}{2}},$$

which yields, via Markov’s inequality,

$$\begin{aligned} & \mathbb{P} \left( \frac{1}{\sqrt{\tilde{\psi}(N)}} \left| \sum_{\ell=\tilde{\psi}(N)}^{\xi_N-1} X_f \circ T_A^\ell \right| > \varepsilon^{\frac{1}{3}} \right) \\ & \leq \varepsilon + \mathbb{P} \left( \sup_{0 < k \leq \varepsilon \tilde{\psi}(N)} \left| \sum_{\ell=0}^{k-1} X_f \circ T_A^\ell \right| > \varepsilon^{\frac{1}{3}} \sqrt{\tilde{\psi}(N)} \right) \leq \varepsilon + \frac{C(\varepsilon \tilde{\psi}(N))^{\frac{p}{2}}}{\varepsilon^{\frac{p}{3}} \tilde{\psi}(N)^{\frac{p}{2}}} \\ & = \varepsilon + C\varepsilon^{\frac{p}{6}}. \end{aligned}$$

Thus, the error term converges in probability to 0. This proves the convergence in distribution of Eq. (4.12).

### References

1. Aaronson, J.: An Introduction to Infinite Ergodic Theory. American Mathematical Society, Providence (1997)
2. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular variation. Encyclopedia of Mathematics and its Applications (1987)
3. Chazottes, J.-R., Gouëzel, S.: On almost-sure versions of classical limit theorems for dynamical systems. Probab. Theory Relat. Fields **138**, 195–234 (2007)
4. Csáki, E., Földes, A.: On asymptotic independence and partial sums. Asymptotic methods in probability and statistics, A volume in honour of Miklós Csörgő, pp. 373–381. Elsevier, Amsterdam (2000)
5. Csáki, E., Földes, A.: Asymptotic independence and additive functionals. J. Theor. Probab. **13**, 1123–1144 (2000)
6. Dobrushin, R.L.: Two limit theorems for the simplest random walk on a line (Russian). Uspekhi Matematicheskikh Nauk **10**, 139–146 (1955)

7. Gaspard, P., Wang, X.J.: Sporadicity: between periodic and chaotic dynamical behaviors. *Proc. Natl. Acad. Sci.* **85**, 4591–4595 (1988)
8. Gouëzel, S.: Sharp polynomial estimates for the decay of correlations. *Isr. J. Math.* **139**, 29–65 (2004)
9. Gouëzel, S.: Central limit theorem and stable laws for intermittent maps. *Probab. Theory Relat. Fields* **128**, 82–122 (2004)
10. Gouëzel, S.: Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.* **38**, 1639–1671 (2010)
11. Gouëzel, S.: Correlation asymptotics from large deviations in dynamical systems with infinite measure. *Colloq. Math.* **125**, 193–212 (2011)
12. Jain, N.C., Pruitt, W.E.: Maximal increments of local time of a random walk. *Ann. Probab.* **15**, 1461–1490 (1987)
13. Liverani, C., Saussol, B., Vaienti, S.: A probabilistic approach to intermittency. *Ergod. Theory Dyn. Syst.* **19**, 671–685 (1999)
14. Marcus, M.B., Rosen, J.: Laws of the iterated logarithm for the local times of recurrent random walks on  $\mathbb{Z}^2$  and of Lévy processes and random walks in the domain of attraction of Cauchy random variables. *Annales de l'Institut Henri Poincaré (B). Probabilités et Statistiques* **30**, 467–499 (1994)
15. Melbourne, I., Terhesiu, D.: Operator renewal theory and mixing rates for dynamical systems with infinite measure. *Invent. Math.* **189**, 61–110 (2012)
16. Melbourne, I., Terhesiu, D.: First and higher order uniform dual ergodic theorems for dynamical systems with infinite measure. *Isr. J. Math.* (to appear)
17. Rosenthal, H.P.: On the subspaces of  $\mathbb{L}^p$  ( $p \geq 2$ ) spanned by sequences of independent random variables. *Isr. J. Math.* **8**, 273–303 (1970)
18. Serfling, R.J.: Moment inequalities for the maximum cumulative sum. *Ann. Math. Stat.* **41**, 1227–1234 (1970)
19. Thaler, M.: Estimates of the invariant densities of endomorphisms with indifferent fixed points. *Isr. J. Math.* **37**, 303–314 (1980)
20. Woess, W.: Random walks on infinite graphs and groups. *Cambridge Tracts in Mathematics*, Cambridge University Press, Cambridge (2000)
21. Zweimüller, R.: Mixing limit theorems for ergodic transformations. *J. Theor. Probab.* **20**, 1059–1071 (2007)