Bounds on Regeneration Times and Limit Theorems for Subgeometric Markov Chains

Randal Douc, Arnaud Guillin, Éric Moulines

BOUNDS ON REGENERATION TIMES AND LIMIT THEOREMS FOR SUBGEOMETRIC MARKOV CHAINS

BORNES DES TEMPS DE RÉGÉNÉRATION ET THÉORÈMES LIMITES POUR DES CHAÎNES DE MARKOV SOUS-GÉOMÉTRIQUES

RANDAL DOUC *, ARNAUD GUILLIN, AND ERIC MOULINES

Abstract. This paper studies limit theorems for Markov Chains with general state space under conditions which imply subgeometric ergodicity. We obtain a central limit theorem and moderate deviation principles for additive not necessarily bounded functional of the Markov chains under drift and minorization conditions which are weaker than the Foster-Lyapunov conditions. The regeneration-split chain method and a precise control of the modulated moment of the hitting time to small sets are employed in the proof.

AMS 2000 MSC 60J10

Stochastic monotonicity; rates of convergence; Markov chains

Résumé. Nous établissons dans ce papier des théorèmes limites pour des chaînes de Markov à espace d'état général sous des conditions impliquant l'ergodicité sous géométrique. Sous des conditions de dérive et de minorisation plus faibles que celles de Foster-Lyapounov, nous obtenons un théorème de limite centrale et un principe de déviation modérée pour des fonctionnelles additives non nécessairement bornées de la chaîne de Markov. La preuve repose sur la méthode de régénération et un contrôle précis du moment modulé de temps d'atteinte d'ensembles petits.

1. Introduction

This paper studies limit theorems and deviation inequalities for a positive Harris recurrent Markov chain $\{X_k\}_{k\geq 0}$ on a general state space X equipped with a countably generated σ-field $\mathcal{X}$. Results of this type for geometrically ergodic Markov chains are now well established: see for instance (Meyn and Tweedie, 1993, Chapter 17) for the central limit theorem and the law of iterated logarithm, de Acosta and Chen (1998), Chen (1999) for moderate deviations results. However, the more subtle subgeometrical case is not nearly as well understood (see for instance Djellout and Guillin (2001)).

These results can be obtained by using the regeneration method constructed via the splitting technique on returns to small sets. These methods typically require bounds for modulated moments of the excursions between two regenerations. In practice, one most often control the corresponding modulated moment of the excursion between two small set return times rather than regeneration times. Our first result in section 2 relate these two bounds, extending to subgeometrical case results reported earlier in the geometric case by Roberts and Tweedie (1999). We then apply these bounds in sections 3, 4 and 5. In section 3, we establish a CLT and Berry-Esseen bounds, sharpening estimates given in Bolthausen (1982). In section 4, we establish a Moderate Deviation Principle for possibly unbounded additive functionals of the Markov chains, extending results obtained earlier for bounded functionals and *atomic* chains by Djellout and Guillin (2001). Finally, in section 5, we give deviation inequality for unbounded additive functionals of the Markov Chain.

Following Nummelin and Tuominen (1983), we denote by Λ_0 the set of sequences such that $r(n)$ is non decreasing and $\log r(n)/n \downarrow 0$ as $n \to \infty$ and by Λ the set of sequences for which $r(n) > 0$ for all $n \in \mathbb{N}$ and for which there exists an $r_0 \in \Lambda_0$ which is equivalent to r in the sense that

$$0 < \liminf_{n\to\infty} \frac{r(n)}{r_0(n)} \quad \text{and} \quad \limsup_{n\to\infty} \frac{r(n)}{r_0(n)} < \infty \, .$$

Without loss of generality, we assume that $r(0) = 1$ whenever $r \in \Lambda_0$. Examples of subgeometric sequences include: polynomial sequences $r(n) = (n+1)^\delta$ ($\delta > 0$), or subexponential sequences, $r(n) = (n+1)^\delta e^{cn^\gamma}$ ($\delta > 0$, $c > 0$ and $\gamma \in (0,1)$).

Denote by P the transition kernel of the chain and for $n \geq 1$, P^n the n-th iterate of the kernel. For any signed measure μ on $(\mathsf{X}, \mathcal{X})$, we denote by $\|\mu\|_f \stackrel{\text{def}}{=} \sup_{|g|\leq f} |\mu(g)|$ the f-total variation norm. Let $f : \mathsf{X} \to [1, \infty)$ be

a measurable function and $\{r(k)\} \in \Lambda$. We shall call $\{X_k\}$ (f,r)-ergodic (or f-ergodic at rate $\{r(k)\}$) if P is aperiodic, ϕ-irreducible and positive Harris recurrent Markov chain and

$$\lim_{n\to\infty} r(n)\|P^n(x,\cdot) - \pi\|_f < \infty\,, \quad \text{for all } x \in \mathsf{X}\,. \tag{1.1}$$

where π is the unique stationary distribution of the chain. If (1.1) holds for $f \equiv 1$, then we call $\{X_k\}$ r-ergodic (or ergodic at rate r). For positive Harris recurrent Markov chain (Meyn and Tweedie, 1993, Chapter V) there exists some (and indeed infinitely many) small sets satisfying for some constant m and some probability measure ν, the minorisation condition: $P^m(x,\cdot) \geq \epsilon\nu(\cdot)$, $x \in C$. In what follows, for simplicity of exposition, we shall consider the "strongly aperiodic case" $m = 1$, that is

A 1. There exist $\epsilon \in (0,1]$, a probability measure ν on $(\mathsf{X}, \mathcal{X})$ such that $\nu(C) = 1$ and for all $x \in C$, $A \in \mathcal{X}$, $P(x,A) \geq \epsilon\nu(A)$.

The general m case can be straightforwardly, but to the price of heavy notations and calculus (considering for example easy extensions of i.i.d. theorem to the 1-dependent case), recovered from the proofs presented here. Fundamental to our methodology will be the regeneration technique (see (Nummelin, 1984, chapter IV). The existence of small sets enables the use of the splitting construction to create atoms and to use regeneration methods, similar to those on countable spaces. In particular, each time the chain reaches C, there is a possibility for the chain to regenerate. Each time the chain is at $x \in C$, a coin is tossed with probability of success ϵ. if the toss is successful, then the chain is moved according to the probability distribution ν, otherwise, according to $(1-\epsilon)^{-1}\{P(x,\cdot) - \epsilon\nu(\cdot)\}$. Overall, the dynamic of the chain is not affected by this coin toss, but at each time the toss is successful, the chains regenerates with regeneration distribution ν independent from x. We denote by $\tau = \inf\{k \geq 1, X_k \in C\}$ and $\sigma = \inf\{k \geq 0, X_k \in C\}$ the first return and hitting time to C and by $\check{\tau} = \inf\{k \geq 1, (X_k, d_k) \in C \times \{1\}\}$ and $\check{\sigma} = \inf\{k \geq 0, (X_k, d_k) \in \mathsf{X} \times \{1\}\}$. Let f be a non-negative function and $r \in \Lambda$ a subgeometric sequence and μ a probability measure on $(\mathsf{X}, \mathcal{X})$. Our main result gives a bound to the (f,r)-modulated expectation of moments $\check{\mathbb{E}}_{\check{\mu}}\left[\sum_{k=1}^{\check{\sigma}} r(k) f(X_k)\right]$ of the regeneration time (where $\check{\mathbb{E}}_{\check{\mu}}$ is the expectation associated to the split chain; see below) in terms of the corresponding moment of $\tilde{\mathbb{E}}_{\mu}\left[\sum_{k=0}^{\tau} r(k) f(X_k)\right]$ and constants depending only and explicitly on ϵ and

ν and on the sequence r. Here, $\tilde{\mathbb{E}}_\mu$ denotes the expectation associated to a Markov chain with initial distribution μ and moving according to P outside C and the residual kernel $(1-\epsilon)^{-1}\{P(x,\cdot)-\epsilon\nu(\cdot)\}$ inside C.

Because finding bounds for $\tilde{\mathbb{E}}_\mu\left[\sum_{k=0}^{\tau} r(k)f(X_k)\right]$ is not always easy, we will consider bounds for this quantity derived from a "subgeometric" condition recently introduced in Douc et al. (2004), which might be seen, in the subgeometrical case, as an analog to the Foster-Lyapunov drift condition for geometrically ergodic Markov Chains. We obtain, using these drift conditions, explicit bounds for the (f,r)-modulated expectation of moments of the regeneration times in terms of the constants in **A1**, the sequence r and the constants appearing in the drift conditions. With these results, we obtain limit theorems for additive functionals and deviations inequalities, under conditions which are easy to check.

2. Bounds for regeneration time

We proceed by recalling the construction of the split chain (Nummelin, 1984, Chapter 4). For $x \in C$ and $A \in \mathcal{X}$ define the kernel Q as follows,

$$Q(x,A) = \begin{cases} (1-\epsilon\mathbb{1}_C(x))^{-1}\left\{P(x,A)-\epsilon\mathbb{1}_C(x)\nu(A)\right\} & 0 \le \epsilon\mathbb{1}_C(x) < 1, \\ \delta_x(A) & \epsilon\mathbb{1}_C(x) = 1 \end{cases} \tag{2.1}$$

Define now, on the product space $\check{\mathsf{X}} = \mathsf{X}\times\{0,1\}$ equipped with the product σ-algebra $\mathcal{X}\otimes\mathcal{P}(0,1)$ where $\mathcal{P}(0,1) \stackrel{\text{def}}{=} \{\emptyset,\{0\},\{1\},\{0,1\}\}$ the split kernel as follows:

$$\check{P}(x,0;A\times\{0\}) = \int_A Q(x,dy)\{1-\epsilon\mathbb{1}_C(y)\} \quad \check{P}(x,0;A\times\{1\}) = \epsilon Q(x,A\cap C)$$
$$\check{P}(x,1;A\times\{0\}) = \int_A \nu(dy)\{1-\epsilon\mathbb{1}_C(y)\} \qquad \check{P}(x,1;A\times\{1\}) = \epsilon\nu(A\cap C)\,.$$

For μ be a probability measure on $(\mathsf{X},\mathcal{X})$, define the split probability $\check{\mu}$ on $(\mathsf{X}\times\{0,1\},\mathcal{X}\otimes\mathcal{P}(\{0,1\})$ by

$$\check{\mu}(A\times\{0\}) = \int_A \{1-\epsilon\mathbb{1}_C(y)\}\mu(dy)\,, \qquad A\in\mathcal{X}\,, \tag{2.2}$$
$$\check{\mu}(A\times\{1\}) = \epsilon\mu(A\cap C)\,. \tag{2.3}$$

We denote by $\check{\mathbb{P}}_{\check{\mu}}$ and $\check{\mathbb{E}}_{\check{\mu}}$ the probability and the expectation on $(\mathsf{X}^{\mathbb{N}}\times\{0,1\}^{\mathbb{N}}$, $\mathcal{X}^{\mathbb{N}}\otimes\mathcal{P}^{\otimes\mathbb{N}}(\{0,1\}))$ associated to the Markov chain $\{X_n,d_n\}_{n\ge0}$ with initial distribution $\check{\mu}$ and transition kernel $\check{P}$. The definition of the split kernel implies

that

$$\check{\mathbb{P}}\left(X_{n+1}\in A\,|\,\mathcal{F}_n^X\vee\mathcal{F}_{n-1}^d\right)=P(X_n,A) \tag{2.4}$$

$$\check{\mathbb{P}}\left(d_n=1\,|\,\mathcal{F}_n^X\vee\mathcal{F}_{n-1}^d\right)=\epsilon\mathbb{1}_C(X_n) \tag{2.5}$$

$$\check{\mathbb{P}}\left(X_{n+1}\in A\,|\,\mathcal{F}_n^X\vee\mathcal{F}_{n-1}^d;d_n=1\right)=\nu(A)\;, \tag{2.6}$$

where for $n\geq 0$, $\mathcal{F}_n^d=\sigma(d_k,k\leq n)$ and by convention $\mathcal{F}_{-1}^d=\{\emptyset,\Omega)$. Condition (2.4) simply states that $\{X_n\}_{n\geq 0}$ is a Markov chain w.r.t. the filtration $(\mathcal{F}_n^X\vee\mathcal{F}_{n-1}^d,n\geq 0)$. Condition (2.5) means that the probability of getting a head ($d_n=1$) as the n-th toss is equal to $\epsilon\mathbb{1}_C(X_n)$, independently of the previous history $\mathcal{F}_{n-1}^X$ and of the $n-1$ previous toss. Condition (2.6) says that, if head is obtained at the n-th toss ($d_n=1$), then the next transition obeys the transition law ν independently of the past history of the chain and of the tosses. This means in particular that $\mathsf{X}\times\{1\}$ is a proper atom. From conditions (2.4), (2.5) and (2.6), we have

$$\check{\mathbb{P}}\left(X_{n+1}\in A\,|\,\mathcal{F}_n^X\vee\mathcal{F}_{n-1}^d;d_n=0\right)=Q(X_n,A)\;.$$

We denote respectively by $\tilde{\mathbb{P}}_\mu$ and $\tilde{\mathbb{E}}_\mu$ the probability and the expectation on $(\mathsf{X}^{\mathbb{N}},\mathcal{X}^{\otimes\mathbb{N}})$ of a Markov chain with initial distribution μ and transition kernel Q.

Denote by $\{\sigma_j\}_{j\geq 0}$ are the successive hitting times of $\{X_n\}$ to the set C

$$\sigma_0\stackrel{\text{def}}{=}\inf\{n\geq 0,X_n\in C\}\quad\text{and}\quad\sigma_j=\inf\{n>\sigma_{j-1},X_n\in C\},\quad j\geq 1\;, \tag{2.7}$$

and by N_n the number of visits of $\{X_n\}$ to the set C before time n,

$$N_n=\sum_{i=0}^{n}\mathbb{1}_C(X_n)=\sum_{j=0}^{\infty}\mathbb{1}_{\{\sigma_j\leq n\}} \tag{2.8}$$

Define by $\check{\sigma}$ the hitting time of the atom of the split chain $\mathsf{X}\times\{1\}$,

$$\check{\sigma}\stackrel{\text{def}}{=}\{k\geq 0,d_k=1\}\;. \tag{2.9}$$

The stopping time $\check{\sigma}$ is a regeneration time and ν is a regeneration measure, *i.e.* the distribution of X_n conditional to $\check{\sigma}=n$ is ν independently of the past history of the chain. The following proposition relates the functionals of the regeneration time under the probability associated to the split chain $\check{\mathbb{P}}_{\check{\mu}}$ to the corresponding functionals of the chain $\{X_n\}$ under the probability $\tilde{\mathbb{P}}_\mu$.

Proposition 1. *Assume **A1**. Let μ a probability measure on $(\mathsf{X}, \mathcal{X})$. Let $\{\xi_n\}$ be a non-negative $\mathcal{F}^X$-adapted process and let S be a $\mathcal{F}^X$-stopping time. Then,*

$$\check{\mathbb{E}}_{\check{\mu}}\left[\xi_S \mathbb{1}_{\{S<\check{\sigma}\}}\right] = \tilde{\mathbb{E}}_{\mu}\left[\xi_S (1-\epsilon)^{N_S} \mathbb{1}_{\{S<\infty\}}\right] , \tag{2.10}$$

$$\check{\mathbb{E}}_{\check{\mu}}\left[\xi_{\check{\sigma}} \mathbb{1}_{\{\check{\sigma}<\infty\}}\right] = \epsilon \sum_{j=0}^{\infty} (1-\epsilon)^j \tilde{\mathbb{E}}_{\mu}\left[\xi_{\sigma_j} \mathbb{1}_{\{\sigma_j<\infty\}}\right] . \tag{2.11}$$

The proof is given in the Appendix A.

We will now apply the proposition above to functionals of the form $\xi_n := \sum_{k=0}^{n} r(k) g(X_k)$ where g is a non-negative function and $r \in \Lambda$ is a sequence, to relate the bounds of the (g, r)-modulated expectation of moments of regeneration time to the (f, r)-modulated expectation of moments of the hitting time.

Proposition 2. *Assume **A1**. Let $\{r(n)\}_{n\geq 0}$ be a sequence such that, for some K, $r(n+m) \leq K r(n) r(m)$, for all $(n, m) \in \mathbb{N} \times \mathbb{N}$. Let $g : \mathsf{X} \to [1, \infty)$ be a measurable function. For $x \in \mathsf{X}$, define*

$$W_{r,g}(x) \stackrel{\text{def}}{=} \tilde{\mathbb{E}}_x \left[\sum_{k=1}^{\tau} r(k) g(X_k)\right], \tag{2.12}$$

Then, for any $x \in \mathsf{X}$,

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} r(k) g(X_k)\right] \leq$$
$$r(0) g(x) + W_{r,g}(x) \mathbb{1}_{C^c}(x) + \epsilon^{-1}(1-\epsilon)\, K \left(\sup_C W_{r,g}\right) \check{\mathbb{E}}_{\check{\delta}_x}[r(\check{\sigma})] . \tag{2.13}$$

If $g \equiv 1$ and $r(n) = \beta^n$, this proposition may be seen as an extension of (Roberts and Tweedie, 1999, Theorem 2.1), which relates the generating function of the regeneration time to that of the hitting time to C. Subgeometric sequences $r \in \Lambda_0$ also satisfies the inequality $r(n+m) \leq r(n) r(m)$. There is however a striking difference with geometric sequence. Whereas for a geometric sequence $\liminf_{n\to\infty} r(n) / \sum_{k=0}^{n} r(k) > 0$, for subgeometric sequence we have on the contrary $\limsup_{n\to\infty} r(n) / \sum_{k=0}^{n} r(k) = 0$. This implies that, whereas $\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} r(k) g(X_k)\right]$ and $\check{\mathbb{E}}_{\check{\delta}_x}[r(\check{\sigma})]$ are of the same order of magnitude in the geometric case, the second is negligible compared to the first one in the subgeometric case. In particular,

Corollary 3. *Assume **A1**. For any function $g : \mathsf{X} \to [0, \infty)$, there exists a constant b_g (depending only and explicitly on ϵ and $\sup_C W_{1,g}$) such that*

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} g(X_k)\right] \leq g(x) + W_{1,g}(x)\mathbb{1}_{C^c}(x) + b_g \ . \tag{2.14}$$

For any $r \in \Lambda_0$ and $\delta > 0$, there exists a constant b_r (depending only and explicitly on ϵ, δ, r and $\sup_C W_{r,1}$) such that

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} r(k)\right] \leq (1+\delta)W_{r,1}(x)\mathbb{1}_{C^c}(x) + b_r \ . \tag{2.15}$$

In general, of course, $\sup_C W_{1,g}$ and $\sup_C W_{r,1}$ is not easy to find analytically and, as in other approaches to this problem, we will consider bounds on these quantities using "subgeometric drift" conditions as introduced in Douc et al. (2004), generalising a condition implying rieamnnian convergence stated in Jarner and Roberts (2001) (see also Fort and Moulines (2000)). This condition may be seen as an analogue for subgeometrically ergodic Markov chain of the Foster-Lyapunov condition for geometrically ergodic Markov chain.

A 2. There exist a concave, non decreasing, differentiable function $\varphi : [1, +\infty) \to \mathbb{R}^+$, a measurable function $V : \mathsf{X} \to [1, \infty)$ and positive constants b satisfying $\varphi(1) > 0$, $\lim_{v\to\infty} \varphi(v) = \infty$, $\lim_{v\to\infty} \varphi'(v) = 0$, $\sup_{x\in C} V(x) < \infty$ and

$$PV \leq V - \varphi \circ V + b\mathbb{1}_C \ ,$$

where the set C is given in **A1**.

This drift condition has been checked in a large number of examples arising for example in queueing theory, Markov Chain Monte Carlo, time-series analysis (see for example Jarner and Roberts (2001),Douc et al. (2004)). Examples of functions φ satisfying **A2** include of course polynomial functions $\varphi(v) = (v+1)^\alpha$ for $\alpha \in (0, 1)$ but also more general functions like $\varphi(v) = \log^\alpha(v + 1)$ for some $\alpha > 0$, or $\varphi(v) = (v + d)/\log(v + d)^\alpha$, for some $\alpha > 0$ and sufficiently large constant d. We refer to Douc et al. (2004) for precise statements giving both drift functions and rate φ for these examples. Define

$$\Phi(v) \stackrel{\text{def}}{=} \int_1^v \frac{dx}{\varphi(x)}. \tag{2.16}$$

The function $\Phi : [1, \infty) \to [0, \infty)$ is increasing and $\lim_{v\to\infty} \Phi(v) = \infty$ (see (Douc et al., 2004, Section 2)). Define, for $u \in [1, \infty)$,

$$r_\varphi(u) \stackrel{\text{def}}{=} \varphi \circ \Phi^{-1}(u)/\varphi \circ \Phi^{-1}(0), \tag{2.17}$$

where Φ^{-1} is the inverse of Φ. The function $u \mapsto r_\varphi(u)$ is log-concave and thus the sequence $\{r_\varphi(k)\}$ is subgeometric. Polynomial functions $\varphi(v) = v^\alpha$, $\alpha \in (0,1)$ are associated to polynomial sequences $r_\varphi(k) = (1+(1-\alpha)k)^{\alpha/(1-\alpha)}$. Functions like $\varphi(v) = c(v+d)/\log^\alpha(v+d)$ ($\alpha \in (0,1)$ and sufficiently large d) are associated to subexponential sequences,

$$r_\varphi(n) \asymp n^{-\alpha/(1+\alpha)} \exp\left(\{c(1+\alpha)n\}^{1/(1+\alpha)}\right) .$$

where for two sequences $\{u_k\}$ and $\{v_k\}$ of positive numbers, $u_k \asymp v_k$ if

$$0 < \liminf_{k\to\infty} \frac{u_k}{v_k} \le \limsup_{k\to\infty} \frac{u_k}{v_k} < \infty .$$

(Douc et al., 2004, Proposition 2.2) shows that, under **A1-2**, for all $x \in \mathsf{X}$,

$$\mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1} \varphi\circ V(X_k)\right] \le V(x) + b\mathbb{1}_C(x) , \tag{2.18}$$

$$\mathbb{E}_x\left[\sum_{k=0}^{\tau_C-1} r_\varphi(k)\right] \le \{V(x) - 1 + br_\varphi(1)\mathbb{1}_C(x)\}/\varphi(1) . \tag{2.19}$$

This implies, using Tuominen and Tweedie (1994) that a Markov Chain satisfying **A1-2** is both $(1, r_\varphi)$- and $(f,1)$-ergodic. Denote by $\mathcal{G}(\varphi)$ the set of measurable functions satisfying:

$$\mathcal{G}(\varphi) \stackrel{\text{def}}{=} \{\psi : [1,\infty) \to \mathbb{R}, \psi \text{ is non decreasing, } \psi/\varphi \text{ is non increasing}\} . \tag{2.20}$$

Similarly to (2.16), for all $\psi \in \mathcal{G}(\varphi)$, define the function

$$\Phi_\psi : v \mapsto \int_1^v \frac{\psi}{\varphi}(u)du . \tag{2.21}$$

The function Φ_ψ is concave, non decreasing and, because $[\psi/\varphi](u) \le [\psi/\varphi](1)$, $\Phi_\psi(u) \le [\psi/\varphi](1)\,(u-1)$ for all $u \ge 1$. The results of the previous section are used to derive explicit bounds for

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} \psi\circ V(X_k)\right] \quad \text{and} \quad \check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} r_\varphi(k)\right]$$

where ψ is any function in $\mathcal{G}(\varphi)$. The following theorem, proved in section B, establishes bounds for the modulated moment of the excursion of the split chain to the atom $\mathsf{X}\times\{1\}$ as a function of the drift condition.

Theorem 4. *Assume **A1-2**. Then, there exists finite constant B_ψ (depending only and explicitly on the constants appearing in the assumptions) such that for*

all $x \in \mathsf{X}$, $\psi \in \mathcal{G}(\varphi)$,

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k)\right] \leq \Phi_\psi \circ V(x) \mathbb{1}_{C^c}(x) + B_\psi \ , \tag{2.22}$$

For any $\delta > 0$, *there exists a finite constant* B_φ *(depending only and explicitly on the constants appearing in the assumptions and* $\delta > 0$*) such that*

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} r_\varphi(k)\right] \leq (1+\delta) V(x) \mathbb{1}_{C^c}(x) + B_\varphi \ , \tag{2.23}$$

For any $c \in (0,1)$ *and* $K \geq 1$, *there exists a finite constant* κ *(depending only and explicitly on the constants appearing in the assumptions) such that for any* $\psi \in \mathcal{G}(\varphi)$, *and* $x \in \mathsf{X}$,

$$\check{\mathbb{P}}_{\check{\delta}_x}\left(\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k) \geq M\right) \leq \kappa \left[\frac{1}{\Phi^{-1}\left\{cM/\psi(K)\right\}} + \frac{\Phi_\psi(K)+1}{(1-c)MK}\right] V(x) \ , \tag{2.24}$$

The rates of convergence for the tail of the excursions may be obtained by optimizing the choice of the constant K with respect to M. As an illustration, consider first the case where $\psi \equiv 1$. Since $\lim_{s\to\infty} \varphi(s) = \infty$, then

$$\lim_{K\to\infty} \frac{\Phi(K)}{K} = \lim_{K\to\infty} \frac{1}{K} \int_1^K \frac{ds}{\varphi(s)} = 0 \ .$$

Therefore, by letting $K \to \infty$ in the right hand side of (2.24) and then, taking $c = 1$,

$$\check{\mathbb{P}}_{\check{\delta}_x}\left(\check{\sigma} \geq M\right) \leq \kappa V(x)/\Phi^{-1}(M) \ .$$

Note that this bound could have been obtained directly by using the Markov inequality with the bound (2.23) of the f-modulated moment of the excursion. Consider now the case: $\psi \equiv \varphi$. By construction, for any $K \geq 1$, $(\Phi_\psi(K) + 1)/K \leq 1$ and for any positive u, $\Phi^{-1}(u) \geq \varphi(1)u + 1$. Taking $K = 1$ in (2.24), Theorem 4 shows that, for some constant κ,

$$\check{\mathbb{P}}_{\check{\delta}_x}\left(\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k) \geq M\right) \leq \kappa V(x)/M \ ,$$

which could have been again deduced from the Markov inequality applied to the bound for the excursion (2.22). The expression (2.24) thus allows to retrieve these two extreme situations. Eq. (2.24) also allows to interpolate the rates for functions growing more slowly than $\varphi \circ V$.

We give now two examples of convergence rates derived from the previous theorem by balancing the two terms of the right hand side appearing in (4).

Polynomial ergodicity. By Eqs (2.16) and (2.17), if $\varphi(v) = v^\alpha$ (with $\alpha \in (0,1)$), then $r_\varphi(k) = (1+(1-\alpha)k)^{\alpha/(1-\alpha)}$ and $\Phi^{-1}(u) \asymp (1-\alpha)^{1/(1-\alpha)} u^{1/(1-\alpha)}$ as $u \to \infty$. Choose $\beta \in (0,\alpha)$ and set $\psi(u) = u^\beta$. Then, $\Phi_\psi(v) = (1+\beta-\alpha)^{-1}(v^{1+\beta-\alpha}-1)$ and the optimal rate in the right hand side of (2.24) is obtained by setting $K = M^{\frac{\alpha}{\beta+(\alpha-\beta)(1-\alpha)}}$. With this choice of K, (2.24) implies that

$$\check{\mathbb{P}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} V^\beta(X_k) \geq M\right] \leq \kappa V(x) M^{-\frac{\alpha}{\beta+(\alpha-\beta)(1-\alpha)}} .$$

This bound shows how the rate of convergence of the tail depends on the tail behavior of the function g and of the mixing rate of the Markov Chain.

Subexponential ergodicity. Assume that $\varphi(v) = c(v+d)(\log(v+d))^{-\alpha}$ for some positive constants c and α and sufficiently large d. Then $\Phi^{-1}(k) \asymp \mathrm{e}^{(c(1+\alpha)k)^{1/(1+\alpha)}}$. Choose for example $\psi(v) = |\log|^\beta(1+v)$, $v \in \mathbb{R}^+$. By optimising the bound w.r.t. K, (2.24) yields:

$$\check{\mathbb{P}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} |\log|^\beta[V(X_k)] \geq M\right] \leq \kappa \mathrm{e}^{-cM^{\frac{1}{1+\alpha+\beta}}} V(x) ,$$

for some constants c and C which does not depend of β or M. Similarly, for $\psi(v) = (1+v)^\beta$ with $\beta \in (0,1)$, there exists a constant $\kappa < \infty$,

$$\check{\mathbb{P}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} V^\beta(X_k) \geq M\right] \leq \kappa M^{\frac{-1}{\beta}} \log^{\frac{2\alpha\beta-1+\beta-\alpha}{\beta}}(M) V(x) .$$

3. Central Limit Theorem and Berry-Esseén Bounds

As a first elementary application of the results obtained in the previous section, we will derive conditions upon which a Central Limit Theorem holds for the normalized sum $S_n(f) \stackrel{\text{def}}{=} n^{-1/2}\sum_{i=1}^n (f(X_k) - \pi(f))$ where π is the stationary distribution for the chain. For u, v two vectors of $\mathbb{R}^d$, denote by $\langle u, v\rangle$ the standard scalar product and $\|u\| = (\langle u, u\rangle)^{1/2}$ the associated norm.

Theorem 5. *Assume **A1-2**. Let ψ be a function such that ψ^2 and $\psi\Phi_\psi$ belong to $\mathcal{G}(\varphi)$. Then, for any function $f : \mathsf{X} \to \mathbb{R}$ such that $\sup_{\mathsf{X}} \frac{|f|}{\psi\circ V} < \infty$,*

$$\int f^2 d\pi + \int |f| \sum_{k=1}^\infty |P^k f - \pi(f)| d\pi < \infty .$$

If in addition $\boldsymbol{\sigma}^2(f) > 0$, where

$$\boldsymbol{\sigma}^2(f) \stackrel{\text{def}}{=} \int \{f - \pi(f)\}^2 d\pi + 2\int f \sum_{k=1}^\infty P^k\{f - \pi(f)\} d\pi , \tag{3.1}$$

then, for any initial probability measure μ on $(\mathsf{X}, \mathcal{X})$ satisfying $\mu(\Phi_\psi) < \infty$, $\sqrt{n}S_n(f)$ converges in distribution to a zero-mean Gaussian variable with variance $\boldsymbol{\sigma}^2(f)$.

Polynomial ergodicity: Assume that $\varphi(v) = v^\alpha$ for some $\alpha \in (1/2, 1)$ and choose $\psi(v) = v^\beta$ for some $\beta \in [0, \alpha]$. Then, $\Phi_\psi(v) = (1+\beta-\alpha)^{-1}(v^{1+\beta-\alpha} - 1)$ and the conditions of theorem 5 are satisfied if $\alpha > 1/2$ and $\beta \in [0, \alpha - 1/2]$. This is equivalent to the condition used in the CLT (Jarner and Roberts, 2001, Theorem 4.4) for polynomially ergodic Markov chains. Note that, if $\alpha < 1/2$, then the moment of order two of the hitting time $\check{\sigma}$ is not necessarily finite, and the CLT does not necessarily holds in this case.

Subexponential ergodicity: Theorem 5 allows to derive a CLT under conditions which imply subexponential convergence. Assume that $\varphi(v) = (d + v)\log^{-\alpha}(d + v)$, for some $\alpha > 0$ and sufficiently large d. The condition of Theorem 5 are satisfied for $\psi(v) \propto v^{1/2}\{\log(v)\}^{-(\alpha+\delta)}$ for $\delta > 0$.

By strengthening the assumptions, it is possible to establish a Berry-Esseén Theorem with an explicit control of the constants.

Theorem 6. *In addition to the assumptions of Theorem 5, suppose that the functions ψ^3, $\psi^2\Phi_\psi$ and $\psi\Phi_{\psi\Phi_\psi}$ belong to $\mathcal{G}(\varphi)$. Let μ be a probability measure on $(\mathsf{X}, \mathcal{X})$ such that $\mu(\Phi_\psi) < \infty$. Then, there exist a constant κ depending only and explicitly on the constants appearing in the assumptions (**A**1-2) and on the probability measure μ such that, for any function $f : \mathsf{X} \to \mathbb{R}$ such that $\sup_{\mathsf{X}} \frac{|f|}{\psi\circ V} < \infty$ and $\boldsymbol{\sigma}^2(f) > 0$,*

$$\sup_t \left| \mathbb{P}_\mu \left(n^{-1/2} S_n(f)/\boldsymbol{\sigma}(f) \le t \right) - G(t) \right| \le \kappa n^{-1/2} , \tag{3.2}$$

where G is the standard normal distribution function.

Berry-Esseen theorems have been obtained for Harris-recurrent Markov chains under moment and strongly mixing conditions by Bolthausen (1982). The use of the results obtained above allow to check these conditions directly from the drift condition. A side result, which is not fully exploited here because of the lack of space, is the availability of an explicit computable expression for the constant κ, which allows to investigate to assess deviation of the normalized sum for finite sample. This provides an other mean to get "honest" evaluation of the convergence of the Markov chain, under conditions which are less stringent than the ones outlined in Jones and Hobert (2001), based on total

variation distance. It is interesting to compare our conditions with those derived in (Bolthausen, 1982, Theorem 1), in the polynomial case, *i.e.* $\varphi(v) = v^\alpha$ $\alpha \in (0,1)$. It is straightforward to verify that the conditions of the Theorem 6 are satisfied by $\psi(v) = v^\beta$ if $\alpha > 2/3$ and $\beta \in [0, \alpha - 2/3]$. On the other hand, the strong mixing rate of this chain is $r(n) = n^{-\alpha/(1-\alpha)}$ (see Douc et al. (2004) and the maximum value of p such that $\pi(V^{p\beta}) < \infty$ is $p = \alpha/\beta$. The Bolthausen condition $\sum_{k=1}^{\infty} k^{(p+3)/(p-3)} r(n) < \infty$, is therefore satisfied again if $\alpha > 2/3$ and $\beta \in [0, \alpha - 2/3)$, the value $\alpha - 2/3$ being this time excluded.

4. Moderate deviations

The main goal of this section is to generalize the MDP result of Djellout-Guillin Djellout and Guillin (2001) from the atomic case to the 1-small set case. We will indicate in the proof the easy modifications needed to cover the general case.

4.1. **Moderate deviations for bounded functions.**

We first consider MDP for bounded mapping, including non separable case (the functional empirical process and the trajectorial case).

Theorem 7. *Assume conditions **A1-2**. Then, for all sequence $\{b_n\}$ satisfying, for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \left(\frac{\sqrt{n}}{b_n} + \frac{b_n}{n} \right) = 0, \tag{4.1}$$

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \left(\frac{n}{\Phi^{-1}(\varepsilon b_n)} \right) = -\infty \,, \tag{4.2}$$

for all initial measure μ satisfying $\mu(V) < \infty$, for all bounded measurable function $f : \mathsf{X} \to \mathbb{R}^d$ such that $\pi(f) = 0$ and for all closed set $F \subset \mathbb{R}^d$, we have

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \mathbb{P}_\mu \left(\frac{1}{b_n} \sum_{k=0}^{n-1} f(X_k) \in F \right) \le - \inf_{x\in F} J_f(x) \,,$$

where J_f is a good rate function, defined by

$$J_f(x) \stackrel{\text{def}}{=} \sup_{\lambda\in\mathbb{R}^d} \left(\langle \lambda, x \rangle - (1/2)\boldsymbol{\sigma}^2(\lambda, f) \right), \tag{4.3}$$

and $\boldsymbol{\sigma}^2$ is defined by (3.1).

The proof is given in section D. de Acosta (1997) proved that the moderate deviation lower bound holds for all bounded function and all initial measure provided that the chain is ergodic of degree 2, *i.e.* for all set $B \in \mathcal{X}$ such

that $\pi(B) > 0$, $\int_B \mathbb{E}_x[\tau_B^2]\pi(dx) < \infty$, where $\tau_B \stackrel{\text{def}}{=} \inf\{k \geq 1, X_k \in B\}$ is the return-time to the set B. It turns out that, under the assumptions **A1-2**, the condition (4.1)-(4.2) implies that the Markov chain is ergodic of degree 2. Note indeed that the conditions (4.1)-(4.2) implies that $\lim_{k\to\infty} k/r_\varphi(k) = 0$. The definition (2.17) of $\{r_\varphi(k)\}$ implies that for some positive c, $\varphi(v) \geq c\sqrt{v}$, for any $v \in [1,\infty)$ and Lemma 12 (stated and proved in section D) shows that this condition implies that the chain is ergodic of degree two. Thus, Theorem 7 together with (de Acosta, 1997, Theorem 3.1) establish the full MDP for bounded additive functionals.

Condition (4.1)-(4.2) linking ergodicity and speed of the MDP may be seen as the counterpart for Markov chains of the condition of Ledoux (1992) for the MDP of i.i.d. random variable linking the tail of this random variable with the speed of the MDP. Let us give examples of the range of speed of the MDP allowed as the function of the ergodicity rate.

Polynomial ergodicity: By Eqs (2.16) and (2.17), if $\varphi(v) = v^\alpha$ (with $\alpha \in (0,1)$), then $r_\varphi(k) \asymp k^{\alpha/(1-\alpha)}$ and $\Phi^{-1}(k) \asymp k^{1/(1-\alpha)}$. Therefore, condition (4.1)-(4.2) is fulfilled as soon as for any $\alpha \in (1/2,1)$ by any sequence $\{b_n\}$ satisfying

$$\lim_{n\to\infty} \left\{ \frac{\sqrt{n}}{b_n} + \frac{\sqrt{n\log n}}{b_n} \right\} = 0 \, .$$

Subexponential ergodicity: Assume that $\varphi(v) = (v+d)(\log(v+d))^{-\alpha}$ for some $\alpha > 0$ and sufficiently large d. Then, $\Phi^{-1}(k) \asymp e^{ck^{1/(1+\alpha)}}$ for some constant c. The condition (4.1)-(4.2) is fulfilled by any speed sequence $\{b_n\}$ satisfying

$$\lim_{n\to\infty} \left\{ \frac{\sqrt{n}}{b_n} + \frac{b_n}{n^{\frac{1+\alpha}{1+2\alpha}}} \right\} = 0 \, .$$

The result can be extended to the empiral measure of a Markov chain. Assume that X is a Polish space and denote by $\mathsf{M}(\mathsf{X})$ the set of finite Borel signed measures on X. Denote by $B(\mathsf{X})$ the collection of bounded measurable functions on X. We equip $\mathsf{M}(\mathsf{X})$ with the smallest topology such that the maps $\nu \mapsto \int_{\mathsf{X}} f d\nu$ are continuous for each $f \in B(\mathsf{X})$, commonly referred to as the τ-topology. The σ-algebra $\mathcal{M}(\mathsf{X})$ on $\mathsf{M}(\mathsf{X})$ is defined to be the smallest σ-algebra such that for each $f \in B(\mathsf{X})$, the map $\nu \mapsto f d\nu$ is measurable. Define the empirical measure L_n as

$$L_n = \frac{1}{b_n} \sum_{k=0}^{n-1} (\delta_{X_k} - \pi) \, .$$

For any $B \in \mathcal{M}(\mathsf{X})$, we denote by $\mathrm{int}_\tau(B)$ and $\mathrm{clos}_\tau(B)$ the interior and the closure of the set B w.r.t. the τ-topology.

Theorem 8. *Under the assumptions of Theorem 7, for every probability measure $\mu \in \mathsf{M}(\mathsf{X})$ satisfying $\mu(V) < \infty$, and any $B \in \mathcal{M}(\mathsf{X})$*

$$\limsup_n \frac{n}{b_n^2} \log \mathbb{P}_\mu(L_n \in B) \leq - \inf_{\gamma \in \mathrm{clos}_\tau(B)} I_0(\gamma) \ ,$$
$$\liminf_n \frac{n}{b_n^2} \log \mathbb{P}_\mu(L_n \in B) \geq - \inf_{\gamma \in \mathrm{int}_\tau(B)} I_0(\gamma)$$

where for $\gamma \in \mathsf{X}$, setting $\bar{f} = f - \pi(f)$,

$$I_0(\gamma) = \sup_{f \in B(\mathsf{X})} \left[\int f d\gamma - \frac{1}{2} \left(\int \bar{f}^2 d\gamma + 2 \int \bar{f} \sum_{k=1}^{\infty} P^k \bar{f} d\pi \right) \right] . \tag{4.4}$$

The proof can be directly adapted from the proof of (de Acosta, 1997, Theorem 3.2) and is omitted for brevity. An explicit expression of the good rate function can be found in (de Acosta, 1997, Theorem 4.1). Other MDP principles (for instance, for the supremum of the empirical process) can be obtained, using the results obtained previously by Djellout and Guillin (2001). To save space, we do not pursue in this direction.

4.2. **Moderate deviations for unbounded functionals of Markov chains.** We give here conditions allowing to consider unbounded functions. These conditions make a trade-off between the ergodicity of the Markov Chain, the range of speed for which a moderate deviation principle may be established and the control of the tails of the functions.

Theorem 9. *Assume **A1-2** and that there exist a function $\psi \in \mathcal{G}(\varphi)$ and a sequence $\{K_n\}$ such that $\lim_{n\to\infty} K_n = \infty$ and, for any positive ε,*

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \left(\frac{n}{\Phi^{-1}(\varepsilon b_n / \psi(K_n))} \right) = -\infty \ , \tag{4.5}$$

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \left(\frac{n \Phi_\psi(K_n)}{\varepsilon b_n K_n} \right) = -\infty \ . \tag{4.6}$$

Then, for any initial distribution μ satisfying $\mu(V) < \infty$ and any measurable function $f : \mathsf{X} \to \mathbb{R}^d$ such that $\sup_{\mathsf{X}} \|f\|/\psi \circ V$, the sequence $\{\sigma_n^2(\lambda, f, \mu)\}$ where

$$\sigma_n^2(\lambda, f, \mu) \stackrel{\mathrm{def}}{=} \mathbb{E}_\mu \left[\left(\frac{1}{n} \sum_{k=0}^{n-1} \{ f(X_k) - \pi(f) \} \right)^2 \right] \ ,$$

has a limit $\sigma^2(\lambda, f)$ (which does not depend on μ) and $\mathbb{P}_\mu\left[L_n(f) \in \cdot\right]$ satisfies a moderate deviation principle with speed b_n^2/n and good rate function $\mathcal{J}_f$,

$$\mathcal{J}_f(x) = \sup_{\lambda \in \mathbb{R}^d} \left[\langle \lambda, x\rangle - (1/2)\sigma^2(\lambda, f)\right] .$$

Moreover, if $\psi^2 + \psi\Phi_\psi \in \mathcal{G}(\varphi)$, then $\sigma^2(\lambda, f) = \boldsymbol{\sigma}^2(\langle \lambda, f\rangle)$ and $\mathcal{J}_f = J_f$.

Polynomial ergodicity: By Eqs (2.16) and (2.17), if $\varphi(v) = v^\alpha$ (with $\alpha \in (1/2, 1)$), then $r_\varphi(k) \asymp k^{\alpha/(1-\alpha)}$ and $\Phi^{-1}(k) \asymp k^{1/(1-\alpha)}$. Choose $\psi(v) = v^\beta$ with $\beta < \alpha - 1/2$. Then the MDP holds for for any sequence $\{b_n\}$ such that $\lim_{n\to\infty}\{\frac{\sqrt{n}}{b_n} + \frac{b_n}{\sqrt{n\log n}}\} = 0$. It is worthwhile to note that the speed which can be achieved are the same than in the bounded case.

Subexponential ergodicity: Assume now that $\varphi(v) = (v + d)(\log(v + d))^{-\alpha}$ for some $\alpha > 0$ and sufficiently marge d. Then Letting $\psi(v) = (\log(1 + v))^\beta$ for some $\beta > 0$, then Theorem 9 shows the MDP with speed $b_n = n^a$ for a such that

$$\frac{1}{2} < a < \frac{\beta + 1 + \alpha}{2\beta + 1 + 2\alpha}.$$

Letting $\psi(v) = (1 + v)^\beta$ with $\beta < 1/2$, then Theorem 9 shows that the MDP principle holds for any sequence $\{b_n\}$ such that $\lim_{n\to\infty}\{\frac{\sqrt{n}}{b_n} + \frac{b_n}{\sqrt{n\log n}}\} = 0$.

5. Deviation inequalities

We now investigate some exponential deviation inequalities for $\mathbb{P}_\mu(\sum_{k=0}^n f(X_i) > \varepsilon n)$ valid for each n where f is a bounded and centered function w.r.t. π. This is to be compared to Bernstein's inequality for i.i.d. variables or more precisely to the Fuk and Nagaev (1971) inequality adapted to Markov chains, (as done in a previous work of Clémençon (2001)) except that in this paper, the Markov chain is not geometrically but sub-geometrically ergodic. Extensions to the case of unbounded functions can be tackled using result of Theorem 4.

Theorem 10. *Assume that f is bounded and centered with respect to π and the assumptions of Theorem 1. Then, for any initial measure μ satisfying $\mu(V) < \infty$, for any positive $\varepsilon > 0$ and $n > n_0(\varepsilon)$, there exists L, K (independent of n and ϵ) such that, for all positive y*

$$\mathbb{P}_\mu\left(\left\|\sum_{k=0}^{n-1} f(X_k)\right\| > \varepsilon n\right) \leq \frac{Ln}{\Phi^{-1}\left(\frac{\varepsilon n}{\|f\|_\infty}\right)} + \frac{Ln}{\Phi^{-1}\left(\frac{y}{\|f\|_\infty}\right)} + e^{-\frac{n\varepsilon^2}{K\|f\|_\infty^2 + \varepsilon y}}.$$

The proof is given in section E. Let us give a few comments on the obtained rate in some examples: with $\|f\|_\infty \leq 1$, for $n \geq n_1$

(1) $\varphi(v) = (1+v)^\alpha$ for $\alpha \in (1/2, 1)$, then there exists K

$$\mathbb{P}_\mu \left(\left\| \sum_{k=0}^{n-1} f(X_k) \right\| > \varepsilon n \right) \leq K \frac{\log(n)^{\frac{1}{1-\alpha}}}{\varepsilon^{\frac{2}{1-\alpha}} n^{\frac{\alpha}{1-\alpha}}}$$

(2) $\varphi(v) = (1+v)\log(c+v)^{-\alpha}$ for positive α, then there exists K, L

$$\mathbb{P}_\mu \left(\left\| \sum_{k=0}^{n-1} f(X_k) \right\| > \varepsilon n \right) \leq K \; e^{-L(n\varepsilon)^{\frac{1}{2+\alpha}}}.$$

The polynomial rate shown in the first case is better than the one derived by Rosenthal's inequality, and considering that we in fact only consider integrability assumptions, are not so far from optimal when considering stronger assumptions as weak Poincare inequalities. The subgeometric case is less satisfactory in the sense that when α is near 0, we hope to achieve a n in the exponential (obtained for example via Cramer argument) whereas we obtained instead $\sqrt{n}$. The gap here, due to Fuk-Nagaev's inequality, is fullfilled only asymptotically via the moderate deviations result, and is left for deviation inequalities for further study.

Appendix A. Proof of Propositions 1 and 2

Proof of the Proposition 1. We first prove by induction that for all $n \geq 0$ and all functions $f_0, \ldots, f_n \in \mathcal{F}^+(\mathsf{X})$,

$$\check{\mathbb{E}}_{\check{\mu}} \left[\prod_{i=0}^{n} f_i(X_i) \mathbb{1}_{\{\check{\sigma} > n\}} \right] = \tilde{\mathbb{E}}_\mu \left[\prod_{i=0}^{n} f_i(X_i)(1-\epsilon)^{N_n} \right]. \tag{A.1}$$

We first establish the result for $n = 0$. For $f \in \mathcal{F}^+(\mathsf{X})$ we have

$$\check{\mathbb{E}}_{\check{\mu}}[f(X_0)\mathbb{1}_{\{\check{\sigma}>0\}}] = \check{\mathbb{E}}_{\check{\mu}}[f(X_0)\mathbb{1}_{\{d_0=0\}}] =$$
$$(1-\epsilon)\int_C f(x)\mu(dx) + \int_{C^c} f(x)\mu(dx) = \int_{\mathsf{X}} \{1 - \epsilon \mathbb{1}_C(x)\} f(x)\mu(dx),$$

Assume now that the result holds up to order n, for some $n \geq 0$. Similarly, for any $f \in \mathcal{F}^+(\mathsf{X})$,

$$\begin{aligned}
&\check{\mathbb{E}}[f(X_{n+1})\mathbb{1}_{\{d_{n+1}=0\}} \mid \mathcal{F}_n^X \vee \mathcal{F}_n^d]\mathbb{1}_{\{d_n=0\}} \\
&\quad = \check{\mathbb{E}}[f(X_{n+1})\{1-\epsilon\mathbb{1}_C(X_{n+1})\} \mid \mathcal{F}_n^X \vee \mathcal{F}_n^d]\mathbb{1}_{\{d_n=0\}} \\
&\quad = \tilde{\mathbb{E}}[f(X_{n+1})\{1-\epsilon\mathbb{1}_C(X_{n+1})\} \mid X_n]\mathbb{1}_{\{d_n=0\}}
\end{aligned}$$

Therefore, by the recurrence assumption,

$$\begin{aligned}
&\check{\mathbb{E}}_{\check{\mu}}\left[f_{n+1}(X_{n+1})\prod_{i=0}^{n} f_i(X_i)\mathbb{1}_{\{\check{\sigma}>n+1\}}\right] \\
&\quad=\check{\mathbb{E}}_{\check{\mu}}\left[\tilde{\mathbb{E}}[f_{n+1}(X_{n+1})\{1-\epsilon\mathbb{1}_C(X_{n+1})\}\mid X_n]\prod_{i=0}^{n} f_i(X_i)\mathbb{1}_{\{\check{\sigma}>n\}}\right] \\
&\quad=\tilde{\mathbb{E}}_{\mu}\left[\tilde{\mathbb{E}}[f_{n+1}(X_{n+1})\{1-\epsilon\mathbb{1}_C(X_{n+1})\}\mid X_n]\prod_{i=0}^{n} f_i(X_i)(1-\epsilon)^{N_n}\right] \\
&\quad=\tilde{\mathbb{E}}_{\mu}\left[f_{n+1}(X_{n+1})\prod_{i=0}^{n} f_i(X_i)(1-\epsilon)^{N_{n+1}}\right],
\end{aligned}$$

showing (A.1). Therefore, the two measures on $(\mathsf{X}^{n+1},\mathcal{X}^{\otimes(n+1)})$ defined respectively by

$$\begin{aligned}
&A\mapsto\check{\mathbb{E}}_{\check{\mu}}\left[\mathbb{1}_A(X_0,\ldots,X_n)\mathbb{1}_{\{\check{\sigma}\geq n\}}\right] \quad\text{and}\\
&A\mapsto\tilde{\mathbb{E}}_{\mu}\left[\mathbb{1}_A(X_0,\ldots,X_1)(1-\epsilon)^{N_n}\right]
\end{aligned}$$

are equal on the monotone class $\mathcal{C}\stackrel{\text{def}}{=}\{A, A=A_0\times\cdots\times A_n, A_i\in\mathcal{X}\}$ for any n, and thus these two measures coincide on the product σ-algebra. The proof of (2.10) follows upon conditioning upon the events $\{S=n\}$. We now prove (2.11). By definition of the hitting time $\check{\sigma}$ to the atom $\mathsf{X}\times\{1\}$, $\xi_{\check{\sigma}}\mathbb{1}_{\{\check{\sigma}<\infty\}}$ may be expressed as

$$\xi_{\check{\sigma}}\mathbb{1}_{\{\check{\sigma}<\infty\}}=\xi_{\sigma_0}\mathbb{1}_{\{d_{\sigma_0}=1\}}\mathbb{1}_{\{d_{\sigma_0}<\infty\}}+\sum_{j=1}^{\infty}\xi_{\sigma_j}\mathbb{1}_{\{d_{\sigma_j}=1\}}\mathbb{1}_{\{\sigma_{j-1}<\check{\sigma}\}}\mathbb{1}_{\{\sigma_j<\infty\}}.$$

Note that

$$\check{\mathbb{E}}\left[\mathbb{1}_{\{d_{\sigma_j}=1\}}\xi_{\sigma_j}\,\middle|\,\mathcal{F}^X_{\sigma_j}\right]\mathbb{1}_{\{\sigma_j<\infty\}}=\epsilon\,\check{\mathbb{E}}\left[\xi_{\sigma_j}\,\middle|\,\mathcal{F}^X_{\sigma_j}\right]\mathbb{1}_{\{\sigma_j<\infty\}}$$

and $(1-\epsilon)^{N_{\sigma_j}}\mathbb{1}_{\{\sigma_j<\infty\}}=(1-\epsilon)^{j+1}\mathbb{1}_{\{\sigma_j<\infty\}}$. The proof follows from the identity

$$\begin{aligned}
&\check{\mathbb{E}}_{\check{\mu}}[\xi_{\sigma_j}\mathbb{1}_{\{\sigma_j<\infty\}}\mathbb{1}_{\{\sigma_{j-1}<\check{\sigma}\}}]=\check{\mathbb{E}}_{\check{\mu}}[\check{\mathbb{E}}[\xi_{\sigma_j}\mathbb{1}_{\{\sigma_j<\infty\}}\mid\mathcal{F}^X_{\sigma_{j-1}}\vee\mathcal{F}^d_{\sigma_{j-1}}]\mathbb{1}_{\{d_{\sigma_{j-1}}=0\}}\mathbb{1}_{\{\sigma_{j-1}<\check{\sigma}\}}]=\\
&\tilde{\mathbb{E}}_{\mu}[\tilde{\mathbb{E}}[\xi_{\sigma_j}\mathbb{1}_{\{\sigma_j<\infty\}}\mid\mathcal{F}^X_{\sigma_{j-1}}](1-\epsilon)^{N_{\sigma_{j-1}}}\mathbb{1}_{\{\sigma_{j-1}<\infty\}}]=(1-\epsilon)^j\tilde{\mathbb{E}}_{\mu}[\xi_{\sigma_j}\mathbb{1}_{\{\sigma_j<\infty\}}].
\end{aligned}$$

□

Proof of the Proposition 2. Without loss of generality we assume that $\sup_C W_{r,g} < \infty$ (otherwise the inequality is trivial). In the case $\epsilon = 1$, Proposition 2 is elementary since by Proposition 1, it then holds that

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=1}^{\check{\sigma}} r(k)g(X_k)\right] = W_{r,g}(x)\mathbb{1}_{C^c}(x).$$

Consider now the case $\epsilon \in (0,1)$. By applying Proposition 1, we obtain:

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=1}^{\check{\sigma}} r(k)g(X_k)\right] = \epsilon W_{r,g}(x)\mathbb{1}_{C^c}(x) + \epsilon \sum_{j=1}^{\infty}(1-\epsilon)^j \tilde{\mathbb{E}}_x\left[\sum_{k=1}^{\sigma_j} r(k)g(X_k)\right]. \quad \text{(A.2)}$$

For $j \geq 1$, write

$$\tilde{\mathbb{E}}_x\left[\sum_{k=1}^{\sigma_j} r(k)g(X_k)\right] = W_{r,g}(x)\mathbb{1}_{C^c}(x) + \sum_{\ell=0}^{j-1}\tilde{\mathbb{E}}_x\left[\sum_{k=\sigma_\ell+1}^{\sigma_{\ell+1}} r(k)g(X_k)\right].$$

Under the stated assumptions, for all $n, m \geq 0$, $r(n+m) \leq Kr(n)r(m)$. This and the strong Markov property imply, for $x \in \{W_{r,g} < \infty\}$:

$$\tilde{\mathbb{E}}_x\left[\sum_{k=\sigma_\ell+1}^{\sigma_{\ell+1}} r(k)g(X_k)\right] = \tilde{\mathbb{E}}_x\left[\sum_{k=1}^{\tau\circ\theta^{\sigma_\ell}} r(k+\sigma_\ell)g(X_{k+\sigma_\ell})\right]$$
$$\leq K\tilde{\mathbb{E}}_x\left[r(\sigma_\ell)W_{r,g}(X_{\sigma_\ell})\right] \leq K\left(\sup_C W_{r,g}\right)\tilde{\mathbb{E}}_x[r(\sigma_\ell)],$$

where θ is the shift operator. Plugging this bound into (A.2) and using again Proposition 1, we obtain,

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=1}^{\check{\sigma}} r(k)g(X_k)\right] \leq W_{r,g}(x)\mathbb{1}_{C^c}(x) + K\left(\sup_C W_{r,g}\right)\sum_{j=1}^{\infty}\epsilon(1-\epsilon)^j\sum_{\ell=0}^{j-1}\tilde{\mathbb{E}}_x[r(\sigma_\ell)]$$
$$= W_{r,g}(x)\mathbb{1}_{C^c}(x) + K\left(\sup_C W_{r,g}\right)\sum_{\ell=0}^{\infty}(1-\epsilon)^{\ell+1}\tilde{\mathbb{E}}_x[r(\sigma_\ell)]$$
$$= W_{r,g}(x)\mathbb{1}_{C^c}(x) + \epsilon^{-1}(1-\epsilon)K\left(\sup_C W_{r,g}\right)\check{\mathbb{E}}_{\check{\delta}_x}[r(\check{\sigma})].$$

□

Proof of corollary 3. For any $r \in \Lambda$, $\lim_{n\to\infty} r(n)/\sum_{k=1}^{n} r(k) = 0$. As a consequence, for any $r(n) \in \Lambda$ and any $\delta > 0$, $N_{r,\delta}$ defined by

$$N_{r,\delta} \stackrel{\text{def}}{=} \sup \left\{ n \geq 1, r(n)/\sum_{k=1}^{n} r(k) \geq \delta \right\} , \tag{A.3}$$

is finite. For any $n \geq 0$, the definition of $N_{r,\delta}$ implies $r(n) \leq \delta \sum_{k=1}^{n} r(k) + r(N_{r,\delta})$. Hence, for any $x \in \mathsf{X}$,

$$\check{\mathbb{E}}_{\check{\delta}_x}[r(\check{\sigma})] \leq \delta \check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=1}^{\check{\sigma}} r(k)\right] + r(N_{r,\delta})$$

The proof of (2.15) then follows by choosing δ sufficiently small so that $(1 - \epsilon^{-1}(1-\epsilon)\sup_C W_{r,1}\delta)^{-1} \leq 1 + \delta$. □

Appendix B. Proof of Theorem 4

We preface the proof by the following elementary lemma.

Lemma 11. *Assume **A**2. Then, for any $\psi \in \mathcal{G}(\varphi)$ there exists b_ψ (depending only and explicitly on b, ψ and φ) such that, for all $x \in \mathsf{X}$,*

$$Q(x, \Phi_\psi \circ V) \leq \Phi_\psi \circ V(x) - \psi \circ V(x) + b_\psi \mathbb{1}_C(x) \tag{B.1}$$

Proof. Since Φ_ψ is concave, differentiable, non decreasing, the Jensen inequality implies, for $x \notin C$

$$\begin{aligned} P[\Phi_\psi \circ V] \leq \Phi_\psi(PV) &\leq \Phi_\psi(V - \varphi \circ V) \\ &\leq \Phi_\psi \circ V + \Phi'_\psi(V)(-\varphi \circ V) \leq \Phi_\psi \circ V - \psi \circ V \end{aligned}$$

and

$$\sup_C Q(x, \Phi_\psi \circ V) \leq \Phi_\psi \left[(1-\epsilon)^{-1} \left\{ \sup_C PV - \epsilon \nu(V) \right\} \right] .$$

The proof follows. □

Proof. By Corollary 3, we may write

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k)\right] \leq \tilde{\mathbb{E}}_x\left[\sum_{k=0}^{\tau-1} \psi \circ V(X_k)\right] \mathbb{1}_{C^c}(x) + \sup_C \psi \circ V + b_g \, . \tag{B.2}$$

On the other hand, the comparison Theorem (Meyn and Tweedie, 1993, Theorem 11.3.1) and the drift condition (B.1) implies that

$$\tilde{\mathbb{E}}_x\left[\sum_{k=0}^{\tau-1} \psi \circ V(X_k)\right] \mathbb{1}_{C^c}(x) \leq \Phi_\psi \circ V(x) \mathbb{1}_{C^c}(x) \, .$$

The proof of (2.22) follows. The proof of (2.23) is along the same lines using (2.15) instead of (2.14).

We now consider (2.24). Define $\eta \stackrel{\text{def}}{=} \inf\{k \geq 0, V(X_k) \geq K\}$. We consider first the event $\left\{\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k) \geq M, \eta \geq \check{\sigma}\right\}$, on which $\psi \circ V(X_k)$ remains bounded by $\psi(K)$. Therefore, on $\{\eta \geq \check{\sigma}\}$, $\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k) \leq (\check{\sigma}+1)\psi(K)$, which implies that

$$\left\{\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k) \geq M, \eta \geq \check{\sigma}\right\} \subset \{\check{\sigma} \geq M/\psi(K)\} \; .$$

We now consider the complementary event: $\left\{\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k) \geq M, \eta < \check{\sigma}\right\}$. We take $c \in (0,1)$, Note that, if $\check{\sigma} < cM/\psi(K)$, then, $\sum_{k=0}^{\eta-1} \psi \circ V(X_k) \leq \eta\psi(K) \leq cM$ which implies that $\sum_{k=\eta}^{\check{\sigma}} \psi \circ V(X_k) \geq (1-c)M$. Therefore,

$$\begin{aligned}\left\{\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k) \geq M, \eta < \check{\sigma}\right\} &\subset \{\check{\sigma} \geq cM/\psi(K)\} \\ &\cup \left\{\eta \leq \check{\sigma} \leq cM/\psi(K) \; , \sum_{k=\eta}^{\check{\sigma}} \psi \circ V(X_k) \geq (1-c)M\right\} \; .\end{aligned}$$

Therefore,

$$\begin{aligned}\check{\mathbb{P}}_{\check{\delta}_x}\left(\sum_{k=0}^{\check{\sigma}} \psi \circ V(X_k) \geq M\right) &\leq 2\check{\mathbb{P}}_{\check{\delta}_x}\left(\check{\sigma} \geq cM/\psi(K)\right) + \\ &\check{\mathbb{P}}_{\check{\delta}_x}\left(\eta \leq \check{\sigma} \leq cM/\psi(K) \; , \sum_{k=\eta}^{\check{\sigma}} \psi \circ V(X_k) \geq (1-c)M\right) \; . \quad \text{(B.3)}\end{aligned}$$

The first term of the right hand side of (B.3) is bounded using the Markov inequality with (2.23),

$$\check{\mathbb{P}}_{\check{\delta}_x}\left\{\check{\sigma} \geq cM/\psi(K)\right\} \leq \frac{\check{\mathbb{E}}_{\check{\delta}_x}\left\{\sum_{k=0}^{\check{\sigma}} r_\varphi(k)\right\}}{\Phi^{-1}\left\{cM/\psi(K)\right\}} \leq \kappa_0 \frac{V(x)\mathbb{1}_{C^c}(x)+1}{\Phi^{-1}\left\{cM/\psi(K)\right\}} \; ,$$

for some finite constant κ_0. Similarly, the Markov inequality and the strong Markov property imply, using Eq. (2.22),

$$\begin{aligned}&\check{\mathbb{P}}_{\check{\delta}_x}\left\{\eta \leq \check{\sigma} \leq cM/\psi(K), \sum_{k=\eta}^{\check{\sigma}} \psi \circ V(X_k) \geq (1-c)M\right\} \\ &\quad \leq \frac{1}{(1-c)M}\check{\mathbb{E}}_{\check{\delta}_x}\left(\mathbb{1}_{\{\eta \leq \check{\sigma}\}}\check{\mathbb{E}}\left\{\sum_{k=\eta}^{\check{\sigma}} \psi \circ V(X_k) \middle| \mathcal{F}_\eta^X\right\}\right) \\ &\quad \leq \frac{\kappa_1}{(1-c)M}\check{\mathbb{E}}_{\check{\delta}_x}\left[(\Phi_\psi \circ V(X_\eta)+1)\mathbb{1}_{\{\check{\sigma} \geq \eta\}}\right] \; ,\end{aligned}$$

for some constant κ_1. The function $u \mapsto \Phi_\psi(u)/u$ is non-increasing. Therefore, $(\Phi_\psi \circ V(X_\eta) + 1)\mathbb{1}_{\{\eta<\infty\}} \leq K^{-1}(\Phi_\psi(K) + 1)V(X_\eta)\mathbb{1}_{\{\eta<\infty\}}$, which implies that

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[(\Phi_\psi \circ V(X_\eta) + 1)\mathbb{1}_{\{\check{\sigma}\geq\eta\}}\right] \leq \frac{(\Phi_\psi(K) + 1)}{K}\check{\mathbb{E}}_{\check{\delta}_x}\left[V(X_\eta)\mathbb{1}_{\{\check{\sigma}\geq\eta\}}\right] .$$

We now prove that there exists a constant κ_2 such that, for any $x \in \mathsf{X}$,

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[V(X_\eta)\mathbb{1}_{\{\eta\leq\check{\sigma}\}}\right] \leq \kappa_2 V(x) . \tag{B.4}$$

Since η is $\mathcal{F}^X$-stopping time, using proposition 1, (2.10), we may write

$$\check{\mathbb{E}}_{\check{\delta}_x}\left[V(X_\eta)\mathbb{1}_{\{\eta<\check{\sigma}\}}\right] = \tilde{\mathbb{E}}_x\left[V(X_\eta)(1-\epsilon)^{N_\eta}\mathbb{1}_{\{\eta<\infty\}}\right] .$$

By conditioning upon the successive visit to the set C, the RHS of the previous equation may be expressed as

$$\tilde{\mathbb{E}}_x\left[V(X_\eta)(1-\epsilon)^{N_\eta}\mathbb{1}_{\{\eta<\infty\}}\right] = \tilde{\mathbb{E}}_x\left[V(X_\eta)\mathbb{1}_{\{\eta<\sigma_0\}}\right] + \sum_{j=1}^{\infty}(1-\epsilon)^j\tilde{\mathbb{E}}\left[V(X_\eta)\mathbb{1}_{\{\sigma_{j-1}\leq\eta<\sigma_j\}}\right] . \tag{B.5}$$

Because $V(X_\eta)\mathbb{1}_{\{\eta<\sigma_0\}} \leq V(X_{\eta\wedge\sigma_0})$ and $\eta\wedge\sigma_0$ is a $\mathcal{F}^X$-stopping time, the comparison Theorem ((Meyn and Tweedie, 1993, Theorem 11.3.1)) implies that, under **A2**,

$$\tilde{\mathbb{E}}_x\left[V(X_\eta)\mathbb{1}_{\{\eta<\sigma_0\}}\right] \leq V(x) + b\mathbb{1}_C(x) . \tag{B.6}$$

Similarly, for any $j \geq 1$, we may write

$$V(X_\eta)\mathbb{1}_{\{\sigma_{j-1}\leq\eta<\sigma_j\}} \leq V(X_{\sigma_j\wedge\eta})\mathbb{1}_{\{\sigma_{j-1}\leq\eta\}} \leq V(X_{\tau\wedge\eta})\circ\theta^{\sigma_{j-1}}\mathbb{1}_{\{\sigma_{j-1}\leq\eta\}} ,$$

and the comparison Theorem and the strong Markov property imply that

$$\tilde{\mathbb{E}}_x\left[V(X_\eta)\mathbb{1}_{\{\sigma_{j-1}\leq\eta<\sigma_j\}}\right] \leq \left(\sup_C V + b\right) . \tag{B.7}$$

By combining the relations (B.5), (B.6) and (B.7), we therefore obtain the bound

$$\tilde{\mathbb{E}}_x\left[V(X_\eta)(1-\epsilon)^{N_\eta}\mathbb{1}_{\{\eta<\infty\}}\right] \leq V(x) + b\mathbb{1}_C(x) + \frac{(1-\epsilon)}{\epsilon}\left\{\sup_C V + b\right\} ,$$

showing (B.4) and concluding the proof. □

Appendix C. Proof of Theorem 5

Proof of Theorem 5. By (Meyn and Tweedie, 1993, Theorem 17.3.6), we only need to check that $I = \check{\mathbb{E}}_{\check{\nu}} \left\{ \left(\sum_{k=0}^{\check{\sigma}} |f|(X_k) \right)^2 \right\} < \infty$. We may write $I = I_1 + 2I_2$ where the two terms I_1 and I_2 are respectively defined by

$$
\begin{aligned}
I_1 &\stackrel{\text{def}}{=} \check{\mathbb{E}}_{\check{\nu}} \left[\sum_{k=0}^{\check{\sigma}} f^2(X_k) \right] \\
I_2 &\stackrel{\text{def}}{=} \check{\mathbb{E}}_{\check{\nu}} \left[\sum_{k=0}^{\check{\sigma}} |f|(X_k) \sum_{\ell=k+1}^{\check{\sigma}} |f|(X_\ell) \right] \\
&= \check{\mathbb{E}}_{\check{\nu}} \left[\sum_{k=0}^{\check{\sigma}} |f|(X_k) \check{\mathbb{E}}_{X_k, d_k} \left\{ \sum_{\ell=0}^{\check{\sigma}} |f|(X_\ell) \right\} \right]
\end{aligned}
$$

The proof follows using Theorem 4. □

Appendix D. Proof of Theorem 7

Lemma 12. *Assume that **A1-2** hold for some function φ such that* $\inf_{v \in [1,\infty)} \frac{\varphi(v)}{\sqrt{v}} > 0$. *Then, the chain is ergodic of degree two.*

Proof. Recall that for a phi-irreducible Markov Chain, the stationary distribution π is a maximal irreducibility measure (see for instance (Meyn and Tweedie, 1993, Proposition 10.4.9)), Therefore any set $C \in \mathcal{X}$ such that $\pi(B) > 0$ is accessible. In addition, for any non-negative measurable function f, $\pi(f) = \int_B \pi(dx) \mathbb{E}_x \left(\sum_{k=0}^{\tau_B - 1} f(X_k) \right)$. A direct calculation shows that

$$
\mathbb{E}_x[\tau_B^2] = 2\mathbb{E}_x \left[\sum_{k=0}^{\tau_B - 1} \mathbb{E}_{X_k}[\tau_B] \right] - \mathbb{E}_x[\tau_B] \, .
$$

Therefore, the Markov chain is ergodic of degree 2 if and only if for any $B \in \mathcal{X}$, $\int_{\mathsf{X}} \pi(dx) \mathbb{E}_x[\tau_B] < \infty$. The Jensen inequality (see for instance (Jarner and Roberts, 2001, Lemma 3.5)) shows that there exists two positive constants c_0 and b_0 such that $P\sqrt{V} \leq \sqrt{V} - c_0 + b_0 \mathbb{1}_C$, and by (Meyn and Tweedie, 1993, Theorem 14.2.3), for any $x \in \mathsf{X}$, and any B such that $\pi(B) > 0$, there exists a constant $c(B)$ such that, for ny $x \in \mathsf{X}$,

$$
\mathbb{E}_x[\tau_B] \leq \sqrt{V(x)} + c(B) \, .
$$

Applying to the inequality $PV + c\sqrt{V} \leq V + b\mathbb{1}_C$ shows that $\pi(\sqrt{V}) < \infty$, which concludes the proof. □

We will only give here the scheme of the proof generalizing the approach of Djellout and Guillin (2001), based on the renewal method introduced (for discrete Markov chains) by Doeblin. Let us first recall the following crucial result due to Arcones-Ledoux: suppose that $\{U_i\}$ are i.i.d. random variables, then $b_n^{-1} \sum_{k=1}^n U_k$ satisfies a MDP if and only if

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \left(n \, \mathbb{P} \left(\|U_k\| \geq b_n \right) \right) = -\infty \, ,$$

and the rate function is the natural quadratic one. Note that by an easy approximation argument (at least in the finite dimensional case) and thus generalizing result by Chen (1997), the previous condition gives also the MDP for a 1-dependent sequence $\{U_i\}$.

The renewal approach consists in splitting the sum $S_n \stackrel{\text{def}}{=} \sum_{i=0}^{n-1} f(X_i)$ into four different terms:

$$S_n = \sum_{k=1}^{e(n)} \xi_k + S_{\check{\sigma}_0 \wedge n} + \left(\sum_{k=1}^{i(n)-1} \xi_k - \sum_{k=1}^{e(n)} \xi_k \right) + \sum_{j=(l(n)+1)}^{n-1} f(X_j) \qquad \text{(D.1)}$$

where $\check{\sigma}_0 \stackrel{\text{def}}{=} \check{\sigma}$ and $\check{\sigma}_k = \inf\{n > \check{\sigma}_{k-1}; d_n = 1\}$ are the successive return times to the atom of the split chain, $i(n) \stackrel{\text{def}}{=} \sum_{k=0}^{n-1} \mathbb{1}(d_k = 1)$ is the number of visits the atom before n, $e(n) = \lfloor \epsilon \pi(C) n \rfloor$ is the expected number of visits to the atom before n, $l(n) \stackrel{\text{def}}{=} \check{\sigma}_{(i(n)-1)\wedge 0}$ is the index of the last visit to the chain to the atom and $\xi_k \stackrel{\text{def}}{=} \sum_{j=\check{\sigma}_{k-1}+1}^{\check{\sigma}_k} f(X_j)$ is the f-modulated moment of the excursion between two successive visits to the atom.

The general idea is to show that only the first term contributes to the moderate deviation principle. To this end we make the following remark: it can be easily checked that $\{\xi_k\}$ is a sequence of i.i.d. random variables with common distribution

$$\check{\mathbb{P}}(\xi_1 \in \cdot) = \check{\mathbb{P}}_{\check{\nu}} \left(\sum_{j=0}^{\check{\sigma}_0} f(X_j) \in \cdot \right) \, .$$

Note that, when $m > 1$ in (A1) then the sequence becomes 1-dependent but essentially the same argument can be carried out. Under (4.5)-(4.6), it is easily seen that

$$\lim_{n\to\infty} \frac{n}{b_n^2} \log \left\{ n \, \check{\mathbb{P}}_{\check{\nu}} \left(\left\| \sum_{k=0}^{\check{\sigma}} f(X_k) \right\| \geq b_n \right) \right\} = -\infty \, ,$$

so that the first term satifies a MDP, the identification of the rate function being easily handled.

Consider now the three remaining terms of the right hand side of (D.1). We have to show that, for any positive ε

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \check{\mathbb{P}}_{\check{\mu}} \left(\|S_{\check{\sigma}_0 \wedge n}\| \geq \varepsilon b_n \right) = -\infty, \tag{D.2}$$

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \check{\mathbb{P}}_{\check{\mu}} \left(\left\| \sum_{j=l(n)+1}^{n-1} f(X_j) \right\| \geq \varepsilon b_n \right) = -\infty, \tag{D.3}$$

$$\limsup_{n\to\infty} \frac{n}{b_n^2} \log \check{\mathbb{P}}_{\check{\mu}} \left(\left\| \sum_{j=1}^{i(n)-1} \xi_j - \sum_{k=1}^{e(n)} \xi_k \right\| \geq \varepsilon b_n \right) = -\infty. \tag{D.4}$$

Remark that the condition ensuring the MDP gives directly the first two needed limits. The last one is more delicate, but as seen from the proof done in the atomic case it merely resumes to the MDP of $\left(\check{\sigma}_k - \check{\sigma}_{k-1} - (\epsilon\pi(C))^{-1}\right)$ (given by Arcones-Ledoux result and (4.5)) which enables us to prove that in the sense of moderate deviations the difference $|i(n) - e(n)|$ can be arbitrarily considered of size $\lfloor \delta n \rfloor$ (δ beeing arbitrary), and the MDP of the sum of $\lfloor \delta n \rfloor$ blocks (ξ_k). This last term being clearly negligible as δ is arbitrary.

Proof of the Theorem 8. The proof of Theorem 8 follows from the projective limit theorem and from the moderate deviation principle for bounded functions (as stated in Theorem 7). The key point consists in checking that the rate function as expressed in Eq. (4.3), Theorem 7 coincides with the one obtained by the projective limit theorem (see for instance de Acosta (1997) and de Acosta and Chen (1998)). □

Appendix E. Proof of Theorem 10

We will the same decomposition than in the moderate deviations proof, i.e. decomposition (D.1)

$$S_n = S_{(\check{\sigma}_0)\wedge n} + \sum_{k=1}^{i(n)-1} \xi_k + \sum_{j=(l(n)+1)}^{n-1} f(X_j). \tag{E.1}$$

We bound $\mathbb{P}_\mu\left(\left\|\sum_{k=0}^{n-1} f(X_k)\right\| > \varepsilon n\right)$ by $\sum_{i=1}^4 I_i$, where

$$I_1 \stackrel{\text{def}}{=} \check{\mathbb{P}}_{\check{\mu}}\left(\left\|\sum_{k=0}^{n-1} f(X_k)\right\| > \varepsilon n, \check{\sigma}_0 > n\right)$$
$$I_2 \stackrel{\text{def}}{=} \check{\mathbb{P}}_{\check{\mu}}\left(\|S_{\check{\sigma}_0}\| > \frac{\varepsilon n}{3}\right)$$
$$I_3 \stackrel{\text{def}}{=} \check{\mathbb{P}}_{\check{\mu}}\left(\left\|\sum_{k=1}^{i(n)-1} \xi_k\right\| > \frac{\varepsilon n}{3}\right)$$
$$I_4 \stackrel{\text{def}}{=} \check{\mathbb{P}}_{\check{\mu}}\left(\left\|\sum_{j=(l(n)+1)}^{n-1} f(X_j)\right\| > \frac{\varepsilon n}{3}\right)$$

where

$$I_1 \leq \check{\mathbb{P}}_{\check{\mu}}\left(\check{\sigma}_0 > n\right) \leq \frac{L}{\Phi^{-1}(n)}\ ,$$

by Theorem 4 if $\mu(V) < \infty$. Remark also

$$I_2 \leq \check{\mathbb{P}}_{\check{\mu}}\left(\check{\sigma}_0 > \frac{\varepsilon n}{\|f\|_\infty}\right) \leq \frac{L}{\Phi^{-1}\left(\frac{\varepsilon n}{\|f\|_\infty}\right)}\ ,$$

and if $\nu(V)$ is bounded,

$$I_4 \leq \check{\mathbb{P}}_{\check{\mu}}\left(\max_{k\leq/+1}(\check{\sigma}_k - \check{\sigma}_{k-1}) > \frac{\varepsilon n}{\|f\|_\infty}\right)$$
$$\leq (n+1)\check{\mathbb{P}}_{\check{\nu}}\left(\check{\sigma}_0 > \frac{\varepsilon n}{\|f\|_\infty} - 1\right) \leq \frac{L(n+1)}{\Phi^{-1}\left(\frac{\varepsilon n}{\|f\|_\infty}\right)}\ .$$

For the last term, note

$$\begin{aligned} I_3 &\leq \check{\mathbb{P}}_{\check{\mu}}\left(\max_{i\leq n}\left\|\sum_{k=1}^{i}\xi_k\right\| > \frac{\varepsilon n}{3}\right) \\ &\leq 2\check{\mathbb{P}}_{\check{\mu}}\left(\left\|\sum_{k=1}^{n}\xi_k\right\| > \frac{\varepsilon n}{6}\right) \end{aligned}$$

where the last step follows by Ottaviani's inequality for i.i.d.r.v. if for n large enough

$$\max_{i\leq n}\check{\mathbb{P}}_{\check{\mu}}\left(\left\|\sum_{k=i}^{n}\xi_k\right\| > \frac{\varepsilon n}{6}\right) \leq 1/2.$$

By Chebyschev's inequality, independence and zero mean of the (ξ_k), it is sufficient to choose n such that

$$n \geq \frac{72\|f\|_\infty^2 \check{\mathbb{E}}_{\check{\nu}}((\check{\sigma}+1)^2)}{\varepsilon^2},$$

where $\check{\mathbb{E}}_{\check{\nu}}((\check{\sigma}+1)^2)$ is finite (and can be easily evaluated) under our drift condition.

By using the Fuk-Nagaev inequality for the remaining term, we get that for all $y > 0$

$$\begin{aligned}\check{\mathbb{P}}_{\check{\mu}}\left(\left\|\sum_{k=1}^{[n/2]+1}\xi_{2k}\right\| > \frac{\varepsilon n}{12}\right) &\leq ([n/2]+1)\check{\mathbb{P}}_{\check{\mu}}\left(\|\xi_1\| > y\right) + \exp\left(-\frac{([n/2]+1)\varepsilon^2}{(9\check{\mathbb{E}}\xi_1^2+\varepsilon y)}\right) \\ &\leq \frac{L([n/2]+1)}{\Phi^{-1}\left(\frac{y}{\|f\|_\infty}\right)} + \exp\left(-\frac{([n/2]+1)\varepsilon^2}{(9\check{\mathbb{E}}\xi_1^2+\epsilon y)}\right),\end{aligned}$$

where $\check{\mathbb{E}}\xi_1^2$ is easily controlled under the drift condition. This concludes the proof.

References

Bolthausen, E. (1982). The Berry-Esseén theorem for strongly mixing Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **60** 283–289.

Chen, X. (1997). Moderate deviations for m-dependent random variables with Banach space values. *Statist. Probab. Lett.* **35** 123–134.

Chen, X. (1999). Limit theorems for functionals of ergodic Markov chains with general state space. *Mem. Amer. Math. Soc.* **139** xiv+203.

Clémençon, S. J. M. (2001). Moment and probability inequalities for sums of bounded additive functionals of regular Markov chains via the Nummelin splitting technique. *Statist. Probab. Lett.* **55** 227–238.

de Acosta, A. (1997). Moderate deviations for empirical measures of Markov chains: lower bounds. *Ann. Probab.* **25** 259–284.

de Acosta, A. and Chen, X. (1998). Moderate deviations for empirical measures of Markov chains: upper bounds. *J. Theoret. Probab.* **11** 1075–1110.

Djellout, H. and Guillin, A. (2001). Moderate deviations for Markov chains with atom. *Stochastic Process. Appl.* **95** 203–217.

Douc, R., Fort, G., Moulines, E. and Soulier, P. (2004). Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.* **14** 1353–1377.

Fort, G. and Moulines, E. (2000). V-subgeometric ergodicity for a Hastings-Metropolis algorithm. *Statist. Probab. Lett.* **49** 401–410.

Fuk, D. H. and Nagaev, S. V. (1971). Probabilistic inequalities for sums of independent random variables. *Teor. Verojatnost. i Primenen.* **16**.

JARNER, S. and ROBERTS, G. O. (2001). Polynomial convergence rates of Markov chains. *Annals of Applied Probability* **12** 224–247.

JONES, G. and HOBERT, J. (2001). Honest exploration of intractable probability distributions via Markov chain Monte Carlo. *Statist. Sci.* **16** 312–334.

LEDOUX, M. (1992). Sur les déviations modérées des sommes de variables aléatoires vectorielles indépendantes de même loi. *Ann. Inst. H. Poincaré* **28** 267–280.

MEYN, S. P. and TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.

NUMMELIN, E. (1984). *General Irreducible Markov Chains and Non-Negative Operators*. Cambridge University Press.

NUMMELIN, E. and TUOMINEN, P. (1983). The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewal theory. *Stochastic Processes and Their Applications* **15** 295–311.

ROBERTS, G. O. and TWEEDIE, R. L. (1999). Bounds on regeneration times and convergence rates for Markov chains. *Stochastic Processes and Their Applications* **80** 211–229.

TUOMINEN, P. and TWEEDIE, R. (1994). Subgeometric rates of convergence of f-ergodic Markov Chains. *Advances in Applied Probability* **26** 775–798.

RANDAL DOUC (CORRESPONDING AUTHOR), CMAP, ECOLE POLYTECHNIQUE, 91128 PALAISEAU CEDEX, FRANCE

E-mail address: douc@cmapx.polytechnique.fr

ARNAUD GUILLIN, CEREMADE, UNIVERSITÉ PARIS IX DAUPHINE, PLACE DU MARÉCHAL DE LATTRE DE TASSIGNY 75775 PARIS CEDEX 16

E-mail address: guillin@ceremade.dauphine.fr

ERIC MOULINES, ECOLE NATIONALE SUPÉRIEURE DES TÉLÉCOMMUNICATIONS, DÉPARTEMENT DE TRAITEMENT DU SIGNAL ET DES IMAGES, 46, RUE BARRAULT, 75634 PARIS CÉDEX 13, FRANCE

E-mail address: moulines@tsi.enst.fr